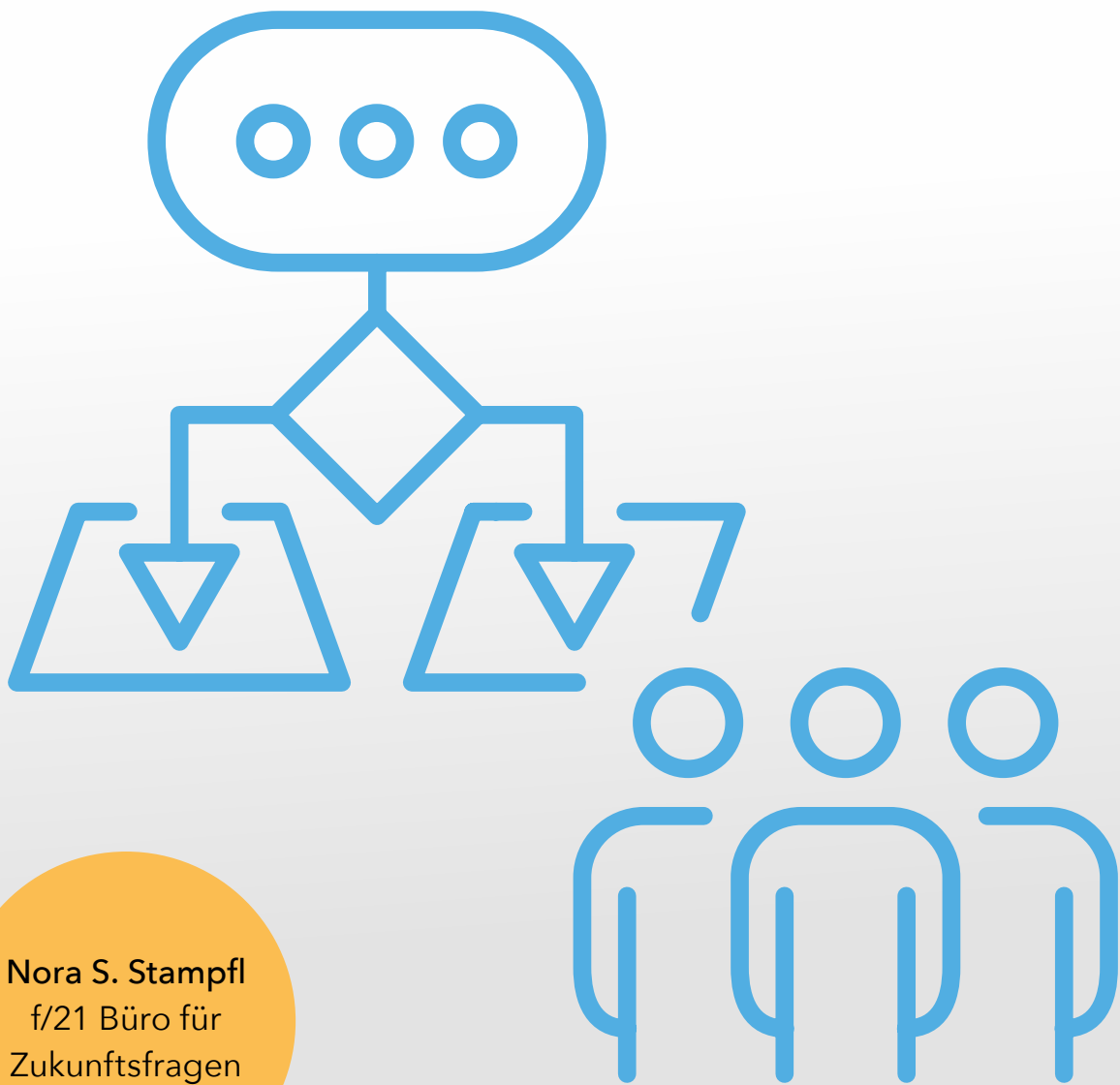


WIE VERÄNDERT KI DIE SOZIALVERSICHERUNG?

Einsatzfelder, Treiber und Erfahrungen



Nora S. Stampfl
f/21 Büro für
Zukunftsfragen

INHALTSVERZEICHNIS

1	Einleitung	4
2	Grundlagen der Künstlichen Intelligenz	7
2.1	Begriffsklärung und Historie	7
2.2	Einsatzfelder der künstlichen Intelligenz	14
2.3	Aktuelle Treiber der Entwicklung von künstlicher Intelligenz	15
2.4	Verbreitung und Akzeptanz von künstlicher Intelligenz	17
3	Einsatz von KI im Bereich der Sozialversicherung	26
3.1	Mögliche Anwendungsfelder	27
3.1.1	Leistungserstellungsprozess der Sozialversicherung	27
3.1.2	Potentielle Anwendungen von KI entlang des Leistungserbringungsprozesses	31
3.1.3	Einsatz von KI im Gesundheitssektor	42
3.2	Konkrete Beispiele des KI-Einsatzes in der Sozialversicherung	46
3.3	Voraussetzungen und Erfolgsfaktoren eines effektiven Einsatzes von KI in Organisationen der sozialen Sicherung	53
4	Chancen und Herausforderungen des KI-Einsatzes in der sozialen Sicherung	60
5	Leitlinien für den Einsatz von KI in den sozialen Sicherungssystemen	72
6	Fazit und Ausblick	81
7	Interviews: Perspektiven auf die KI-Nutzung in der Sozialversicherung	85
7.1	Ellen Hellmann (BG BAU)	85
7.2	Jens Klostermann und Manuela Makiola (DRV Bund)	89
7.3	Dr. Henrike Stephani (Fraunhofer ITWM)	93
8	Quellenverzeichnis	97

1 EINLEITUNG

Die Wahrnehmung, dass unsere Gesellschaft immer komplexer wird, ist weit verbreitet. Im Jahr 1983 erklärt Niklas Luhmann in einem Vortrag in Bielefeld die Komplexitätszunahme damit, dass „im Vergleich zu älteren Gesellschaften dem Handeln heute viel mehr Entscheidungen vorgelagert sind“. Dabei kommt es „zu hochkomplexen Entscheidungsverkettungen: Entscheidungen über Entscheidungen, Mitentscheidungen, Entscheidungen darüber, ob man entscheiden oder nicht entscheiden soll, Entscheidungen über Beteiligung an Entscheidungen usw.“ (Luhmann, 2022, S. 247; Hervorhebung im Original) Gerade in Bezug auf den modernen Sozialstaat ist es einleuchtend, Gesellschaft mit Luhmann (2022, S. 248) als „Konglomerat aus Entscheidungen“ zu begreifen und Bürokratie als einen der möglichen Wege zu sehen, mit der Vielzahl an Entscheidungen umzugehen. Wann immer man als Bürger mit dem Sozialstaat in Berührung kommt, mag man Bürokratie zwar als Ärgernis empfinden, nichtsdestotrotz bleibt nur, Luhmann (2022, S. 248) zuzustimmen, wenn er Bürokratie zu den „Unvermeidlichkeiten der modernen Gesellschaft“ zählt.

Der moderne Sozialstaat ist in der Tat ohne Bürokratie nicht denkbar, allein schon aufgrund der Bindung an Gesetze. In der Tat wäre ein nicht-bürokratischer Sozialstaat eine Horrorvorstellung. Der bürokratische Sozialstaat, wie wir ihn heute kennen, ist überdies nur auf Basis einer umfangreichen Computerisierung möglich. Daten zu sammeln, vorzuhalten, zu verarbeiten und verschiedene Datenbestände miteinander zu verknüpfen, ist die Grundlage jeglicher Entscheidung und ebnet damit erst den Weg zur Erbringung von Sozialleistungen. Im Zuge der Digitalisierung des öffentlichen Sektors hat generell die Bedeutung persönlicher Daten für den modernen Sozialstaat zugenommen. (vgl. Gantchev, 2019, S. 6) Kasper (2020, S. 413) zeigt für den Bereich der Rentenversicherung, dass in der Bundesrepublik Deutschland die Computerisierung und die Ausdifferenzierung des Sozialrechts „das Zusammenspiel zwischen komplexerem Sozialsystem und intensiverer Computernutzung [begünstigten]“. Rentenpolitische Entscheidungen erforderten eine immer größere Menge an Informationen der Bürger, sodass an die Datenerfassung und -verarbeitung immer höhere Anforderungen gestellt wurden. Analog zur Rentenversicherung darf man auch für die anderen Säulen der Sozialversicherung annehmen, dass Informations- und Kommunikationstechnologie nicht nur Rationalisierungswerkzeug war und ist, sondern von Beginn an als Schrittmacher eines komplexeren Sozialstaats fungierte.

Kehren wir nochmals kurz zu Niklas Luhmann zurück: Wenn der Soziologe davon spricht, dass es angesichts der hochkomplexen Verkettungen von Entscheidungen so aussähe, „als ob die Gesellschaft sich eine zweite Komplexität schafft, um ihre erste Komplexität zu ertragen, und dann diese zweite nicht mehr ertragen kann“ (Luhmann, 2022, S. 248), dann hat die Frage, welche Folge ein solcher Befund nun hat, nichts an Aktualität verloren. An dieser Stelle kommt künstliche Intelligenz (KI) ins Spiel: Weil durch die Nutzung von KI für die Erkennung von Wirkungszusammenhängen in einer unüberschaubaren Menge von Daten heute die Beherrschung von Komplexität besser gelingen kann.

Mit künstlicher Intelligenz steht nun also eine Technologie zur Verfügung, zu deren Versprechungen stets gehörte, mit Komplexität in einer Weise fertig zu werden, zu der natürliche Intelligenz niemals imstande wäre. Vor diesem Hintergrund kann es kaum verwundern, dass der aktuelle Entwicklungssprung der KI-Technologie, in dessen Zuge praxistaugliche Anwendungen ins Leben gerufen wurden, zu deren starker Verbreitung in Organisationen führte. Die Nutzung von KI bietet auch im Rahmen der bürokratischen Aufgaben des Sozialstaats eine Antwort auf die Fragen nach dem Umgang mit den überbordenden Datenmassen in einem Umfeld, das durch notorische Budgetknappheit gekennzeichnet ist. Vor dem Hintergrund dieser Aussichten und weil digitale Technologien in sämtlichen unserer Lebensbereiche zur Selbstverständlichkeit geworden sind, überrascht es nicht, dass KI auch in Sozialversicherungsorganisationen Einzug gehalten hat.

Zunehmend ist es möglich, über Chatbots mit Behörden in Kontakt zu treten. Der Einsatz künstlicher Intelligenz erlaubt der Verwaltung, Profile von Bürgern zu erstellen und auf dieser Basis eine differenzierte, personalisierte oder irgendwie „besondere“ Behandlung von Fällen. Daten können mit anderen staatlichen und nichtstaatlichen Datensätzen abgeglichen werden, um die Einhaltung von Vorschriften sicherzustellen, was die Grundlage für eine automatisierte Prüfung von Anspruchsberechtigungen, die Aussetzung und Streichung von Leistungen sowie die Verhängung von Sanktionen darstellt. Dies sind nur einige Beispiele, inwiefern der Einsatz von künstlicher Intelligenz dazu beitragen kann, dass die Sozialverwaltung ganz gravierend ihr Gesicht verändert. Nicht zuletzt die vorgebrachten Rechtfertigungen für das Einbringen der neuen Technologie in die Organisationen der Sozialverwaltung sind aufschlussreich: Effizienzgewinne, Kostensenkungen und Personaleinsparungen wurden zwar seit eh und je mit der Einführung digitaler Technologien in Aussicht gestellt. Im Falle von KI-Einführungen gesellten sich neue Rechtfertigungsgründe wie etwa ein verbesserter Kundenservice, Personalisierung oder die Aufdeckung von Überzahlung und Betrug hinzu.

Noch weniger als ein Computer stellt die „Blackbox“ KI ein bloßes Werkzeug dar, mit dem die gewohnten Aufgaben schlicht auf andere Weise ausgeführt werden. Vielmehr liegt die Vermutung nahe, dass die Nutzung von künstlicher Intelligenz in der Sozialverwaltung weitreichendere Konsequenzen für Individuen und Gesellschaft hat. Vor diesem Hintergrund muss die Frage erlaubt sein, ob nicht die großen Potenziale, die künstliche Intelligenz durch ihr Prinzip der Mustererkennung in großen Datenmengen eröffnet, auf die Inhalte und die Steuerung der sozialen Sicherung zurückwirken.

Die Nutzung von künstlicher Intelligenz im Bereich der sozialen Sicherung ist ein noch weitgehend unkartiertes Terrain. Die vorliegende Studie möchte einen Beitrag leisten, das Feld zu erkunden, indem nach einem kurzen Überblick über die Grundlagen künstlicher Intelligenz grundlegende Möglichkeiten eines Einsatzes von KI in der Sozialverwaltung dargestellt und konkrete Beispiele – in Deutschland sowie international – vorgestellt werden. Diskutiert werden darüber hinaus auch Chancen sowie vor allem die Risiken, welche mit dem Einsatz dieser Technologie für den modernen Sozialstaat verknüpft sind, um schließlich kritisch auf Bestrebungen diverser Akteure einzugehen, die die Nutzung von KI innerhalb wünschenswerten Schranken einhegen sollen.

2 GRUNDLAGEN DER KÜNSTLICHEN INTELLIGENZ

Künstliche Intelligenz ist das Schlagwort der Stunde. Eine vorübergehende Modeerscheinung ist KI jedoch nicht – das dürfte nach den Innovationsprüngen der jüngsten Zeit und der immer weiteren Verbreitung praktischer Anwendungen sicher sein. Die Technologie hinterlässt unverkennbare Spuren in der Art und Weise, wie wir leben und arbeiten. Das Zusammentreffen einer Reihe von Entwicklungen – beginnend bei einer exponentiell anwachsenden Rechenleistung bis hin zu einer sich stets ausweitenden Masse verfügbarer Daten – hat eine Innovationswelle ausgelöst, die keine Anzeichen zeigt, bald wieder abzuflachen.

Um die nachfolgenden Ausführungen zu den Potenzialen des Einsatzes von KI im Bereich der sozialen Sicherung auf eine solide Basis zu stellen, soll im Folgenden zunächst kurz umrissen werden, wovon eigentlich die Rede ist, wenn von künstlicher Intelligenz gesprochen wird.

2.1 BEGRIFFSKLÄRUNG UND HISTORIE

Will man also der Frage auf den Grund gehen, welchen Einfluss KI auf die Systeme der sozialen Sicherung hat und haben wird, so steht am Beginn die Klärung dessen, was denn unter KI eigentlich zu verstehen ist. Der Untersuchungsgegenstand macht es jedoch alles andere als einfach, den nachfolgenden Ausführungen schlichtweg eine zufriedenstellende Definition voranzustellen. Dies liegt nicht zuletzt daran, dass noch nicht einmal unumstritten ist, was man sich denn exakt unter natürlicher Intelligenz, um deren Nachahmung es einer wie auch immer gearteten künstlichen Intelligenz ja letzten Endes geht, vorzustellen hat. Menschliche Intelligenz ist ein derart vielschichtiges Konzept, dass allgemeingültige Definitionen schlicht nicht zu finden sind. So unterscheidet etwa der Psychologe Howard Gardner (1983) mehrere Formen der Intelligenz: linguistisch, musikalisch, logisch-mathematisch, räumlich, körperlich-kinästhetisch, interpersonal und intrapersonell. Verfolgt man die historische Entwicklung von KI und den damit einhergehenden Diskurs, so erscheint die von Rich et al. (2009, S. 3) vorgeschlagene Definition – so vage und flexibel sie auch bleiben mag – daher zunächst ein guter Ausweg aus dieser Zwickmühle: „*Artificial intelligence (AI) is the study of how to make computers do things which, at the moment, people do better.*“

Die Kennzeichnung von KI als die Erforschung dessen, Maschinen dazu zu bringen, Dinge zu tun, die zum jeweiligen Zeitpunkt von Menschen besser ausgeführt werden können, macht auf zweierlei aufmerksam: Auf der einen Seite verdeutlicht eine solche Definition, dass im Bereich von KI die technische Fortentwicklung nicht zu einem Ende kommt, vielmehr die Grenzen des Machbaren stets weiter verschoben werden. KI wird zu jeder Zeit abgesteckt durch die menschlichen Fähigkeiten, die von Computern nachgeahmt werden sollen: Galt der Schach spielende Computer einst als großer Durchbruch in der KI-Forschung, so erscheint heute die

Bewältigung derart formaler Aufgaben als Messlatte zu niedrig. Von einer künstlichen Intelligenz, die diesen Namen verdient, wurde bald schon mehr erwartet als ein Computer, der über eine Art Inselbegabung verfügt und ein Brettspiel gewinnt – sich im echten Leben zurechtzufinden, mit Menschen auf Augenhöhe zu kommunizieren, sollte fortan der Maßstab sein. Dabei begleitet die KI-Forschung ein Dilemma, das als Moravecs Paradox bekannt wurde: Hans Moravecs (1988, S. 15), ein Pionier auf dem Feld der Erforschung künstlicher Intelligenz, gelangte zur Erkenntnis, dass es relativ einfach ist, Computern Leistungen abzurufen, für die hochrangiges Denken erforderlich ist, wie etwa Schachspielen, während es ungleich schwieriger ist, ihnen Fähigkeiten zu verleihen, wie sie schon für Kleinkinder in Bezug auf Wahrnehmung und Mobilität selbstverständlich sind. Auf der anderen Seite legen diese Beobachtungen in Zusammenhang mit obiger Definition nahe, dass der Weg zu einem echten Verständnis von KI nur über die Auseinandersetzung mit deren Entwicklung führen kann. Im Folgenden sei daher ein kurzer Abriss der Forschungsgeschichte künstlicher Intelligenz gegeben.

Als Geburtsstunde der KI als Forschungsgebiet gilt die Dartmouth Conference am Dartmouth College in Hanover im US-amerikanischen Bundesstaat New Hampshire im Sommer 1956. Der von dem Informatiker John McCarthy organisierte mehrwöchige Workshop namens „Dartmouth Summer Research Project on Artificial Intelligence“ versammelte Forscher, darunter Marvin Minsky, Nathaniel Rochester und Claude Elwood Shannon, um gemeinsam an der Frage zu arbeiten, wie eine Maschine zu konstruieren sei, die sich auf eine Art und Weise verhält, welche man bei Menschen als intelligent bezeichnen würde.

Die Idee, menschliches Denken zu automatisieren oder zu mechanisieren, ist jedoch bereits viel älter. Als frühe Quelle einer Maschine, die auf gewisse Weise intelligentes Verhalten zeigt, kann Julien Offray de La Mettrie mit seinem 1748 veröffentlichten Werk *L'HOMME MACHINE* gelten. Auch die Vorstellung des nach dem französischen Mathematiker, Physiker und Astronomen Pierre-Simon Laplace benannten Laplaceschen Dämons ist insofern mit den KI-Forschungsbemühungen verflochten, als Laplace eine Intelligenz skizziert, die Ausdruck eines absolut deterministischen Weltbilds ist: Weil das Universum wie ein Uhrwerk funktioniert, meinte Laplace, müsse sein Dämon bloß den Zustand in allen Details kennen und könnte daraufhin die weitere Entwicklung bis in alle Zeit vorhersagen. Unübersehbar sind die heutigen Vorstellungen ungeahnter Erkenntnisgewinnung durch den Zugang zu großen Datenmengen und besonders zu neuen Techniken ihrer Analyse („Big Data“) ein Wiederläufer von Laplaces Gedankenexperiment. Selbstredend kann auch Gottfried Wilhelm Leibniz mit der der künstlichen Intelligenz zugrundeliegenden Annahme, der Prozess menschlichen Denkens könne formalisiert werden, in Verbindung gebracht werden. Mit seinem Projekt der Erschaffung einer Universalsprache als Grundlage einer Universalwissenschaft kann Leibniz insofern als theoretischer Wegbereiter von KI gelten, als es ihm darum ging, eine „Denkmaschine“ zu entwickeln, mit deren Hilfe die Lösung aller wissenschaftlichen Probleme auf den Umgang mit Symbolen reduziert werden könne. Es ist diese Auffassung von menschlichem Denken als Symbolmanipulation, die auch für die frühe Phase der KI-Entwicklung bestimmend war.

In den auf die Dartmouth Conference folgenden Jahren feierte die KI-Forschung mit Programmen, die einfache Aufgaben lösen konnten, frühe Erfolge. War diese Phase der KI-Forschung grundsätzlich gekennzeichnet von einem großen Enthusiasmus und beachtlichen Fortschritten, so stießen die Forscher jedoch an praktische Grenzen, die vor allem durch die damalige Computertechnik markiert wurden. Die ambitionierten Visionen übertrafen die realen Möglichkeiten der verfügbaren Rechenleistung. In der Folge dämpften sich die großen Erwartungen der Anfangsjahre und man gelangte zu einer realistischeren Einschätzung in Bezug auf die Möglichkeiten der KI. Es folgte eine Zeit der Ernüchterung und Enttäuschung, oftmals als KI-Winter bezeichnet, in der das Interesse an und die Finanzierung der KI-Forschung abebbten. Ab etwa 1980 folgte eine weitere Boom-Welle, die durch die Entwicklung regelbasierter Expertensysteme ausgelöst wurde. Weil diese durch Abbildung von Domänenwissen durch die Befragung von Experten gekennzeichneten Systeme allerdings die an sie gestellten Erwartungen nicht einlösen konnten, trat die KI-Entwicklung ab den 1990er Jahren in ihre zweite Winterphase ein.

Der zweite KI-Winter endete um die Jahrtausendwende mit einem Paradigmenwechsel: Mehr und mehr wurden regelbasierte durch statistische und maschinelle Lernverfahren ersetzt. Insbesondere die durch Cloud-Computing und die zunehmende Verfügbarkeit von Daten angetriebene Deep-Learning-Revolution läutete einen neuerlichen KI-Frühling ein: Als spezieller Zweig des maschinellen Lernens hat sich das sogenannte Deep Learning als außerordentlich leistungsfähiges Verfahren im Zusammenhang mit dem Lernen aus Daten herauskristallisiert. Deep Learning basiert auf Berechnungsmodellen, den sogenannten neuronalen Netzen, die ursprünglich von den Mechanismen des Lernens und der Informationsverarbeitung im menschlichen Gehirn inspiriert wurden. In den künstlichen neuronalen Netzen verarbeiten kleine Recheneinheiten (künstliche Neuronen) Inputs nach bestimmten Regeln zu Outputs, welche wiederum in ein oder mehrere Neuronen eingehen. Solche künstlichen neuronalen Netze werden durch komplizierte Algorithmen gesteuert und können mit Hilfe riesiger Datenmengen trainiert werden, sodass sie in einem gewissen Sinne „lernen“. Neuronale Netze erwiesen sich in vielen Anwendungsbereichen als überlegenes Verfahren und trugen entscheidend dazu bei, dass die KI-Forschung etwa in den Bereichen der Bilderkennung oder der Sprachverarbeitung Erfolge feiern konnte. Entwicklungsbemühungen richten sich heute in erster Linie auf KI-Technologien, die wahrnehmen, lernen und entscheiden können. Die großen Erfolge auf Basis von Deep Learning und der immensen Datenexplosion als entscheidenden Antriebsfaktoren führten zu einem regelrechten KI-Rausch, wodurch sich der Fokus heute stark auf die Entwicklung und Implementierung von Alltagsanwendungen richtet.

WIE VERÄNDERT KI DIE SOZIALVERSICHERUNG?

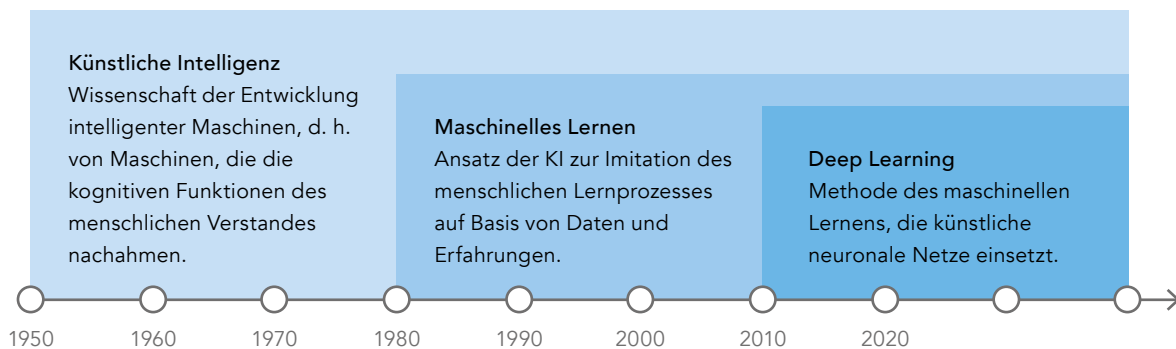


Abbildung 1: Künstliche Intelligenz, Maschinenlernen und Deep Learning;
Quelle: Eigene Darstellung, in Anlehnung an Hassan et al., 2022, S. 4

Generative KI, die auf hochentwickelten Modellen für maschinelles Lernen beruht, brachte das Thema der künstlichen Intelligenz weltweit in die Schlagzeilen. Insbesondere die Einführung des Chatbots ChatGPT durch das US-amerikanische Unternehmen OpenAI im Jahr 2022 hatte einen kaum zu überschätzenden Anteil daran, das Thema KI auch Privatanwendern näher zu bringen. Generative KI lernt aus vorhandenen Artefakten, um auf die Anfrage eines Benutzers neue, realistische Originalinhalte wie Text, Bilder, Video, Audio oder Softwarecode zu generieren, die die Merkmale der Trainingsdaten widerspiegeln, ohne diese zu wiederholen. Indem dieser Typus von KI Inhalte auf Anfragen in natürlicher Sprache erstellt, ohne dem Anwender besondere Kenntnisse oder gar die Eingabe von Code abzuverlangen, schreibt generative KI ein völlig neues Kapitel in der langen Geschichte der KI-Entwicklung. Sowohl für Individuen als auch Organisationen verspricht generative KI immense Produktivitätsgewinne und hat ein beispielloses Innovationspotenzial.

Dementsprechend gehen einer weltweiten Umfrage zufolge die meisten Menschen davon aus, dass sie aufgrund von KI künftig weniger Zeit für die Erledigung von Dingen benötigen werden. 54 Prozent der Befragten (weltweit; 18 Jahre und älter) erwarten in den kommenden Jahren eine Zeitersparnis. 28 Prozent rechnen hingegen nicht damit, dass sich die Effizienz der Zeitnutzung verbessern werde und gehen davon aus, dass in dieser Hinsicht alles beim Alten bleiben werde. Zugleich erwarten 36 Prozent der Befragten Verschlechterungen durch KI-Einsatz auf dem Arbeitsmarkt. (vgl. Thormundsson, 2024) In Unternehmen sind mit der Einführung von generativer KI hohe Erwartungen verknüpft – vor allem im Hinblick auf die Verbesserung von Effizienz und Produktivität, Kostenreduktion und das Ankurbeln von Innovation und Wachstum werden positive Impulse erwartet. (vgl. Deloitte, 2024, S. 11-12) In der deutschen Wirtschaft halten 68 Prozent der Unternehmen KI für die wichtigste Zukunftstechnologie. Dem stehen 29 Prozent gegenüber, die KI für einen massiv überschätzten Hype halten. (vgl. Bitkom Research, 2023, S. 2)

WIE VERÄNDERT KI DIE SOZIALVERSICHERUNG?

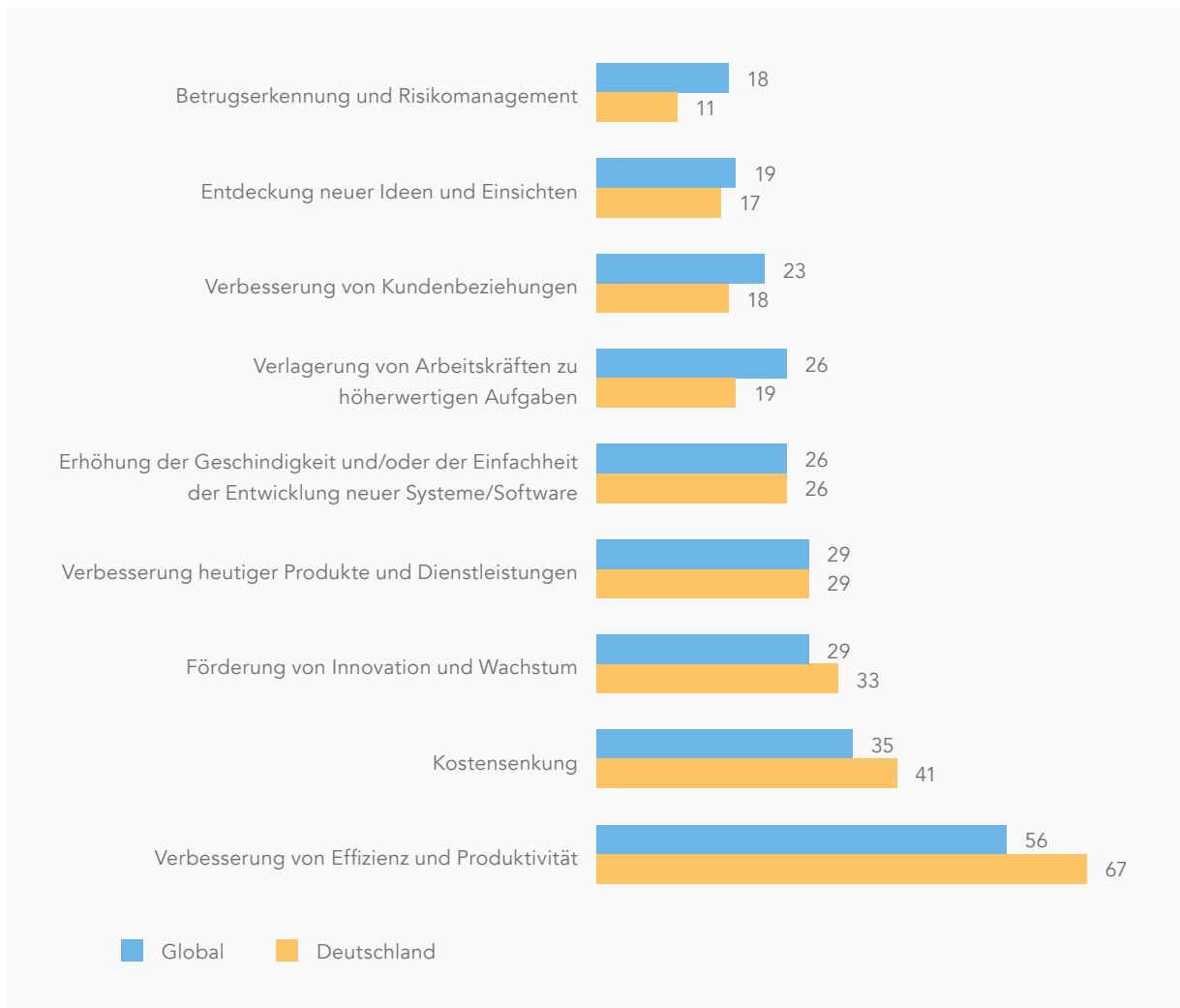


Abbildung 2: Nutzen durch generative KI; Quelle: Eigene Darstellung, in Anlehnung an Deloitte, 2024, S. 12

Vor dem Hintergrund der historischen Entwicklung hin zu immer leistungsfähigeren KI-Systemen, die von der maßgeblichen Frage „Können Maschinen denken?“ begleitet wurde, sind unterschiedliche Ansätze in Bezug auf die Definition von KI-Zielen zu differenzieren. Weit verbreitet ist die auf John R. Searle (1980) zurückgehende Unterscheidung von schwacher und starker KI:

- **Schwache KI** bezeichnet solche Systeme, die darauf abzielen, menschliche Intelligenz zu simulieren bzw. zu imitieren, um bestimmte, klar definierte Aufgaben mit einer festgelegten Methodik zu lösen. Die Übertragung bekannter Lösungsmuster auf andere Aufgaben ist solchen Systemen jedoch nicht möglich. Schwacher KI fehlt Intentionalität, also die Fähigkeit, sich auf die Dinge der Welt zu beziehen.
- **Starke KI** ist hingegen eine universell einsetzbare KI, die menschliches Denken nicht nur simuliert, vielmehr ist ihr Denkfähigkeit zuzuschreiben. In dieser Zukunftsvision kann KI als autonome, d. h. aus sich selbst heraus zwecksetzende Technik Aufgabenstellungen eigenständig erkennen, Probleme analysieren, um zu angemessenen Lösungen zu gelangen. Hierbei sind auch neue oder kreative Lösungen möglich. Eine solche KI würde dem Menschen in nichts nachstehen oder diesen sogar übertreffen. In diesem Fall ist von Super-Intelligenz die Rede.

Vor dem Hintergrund dieser Differenzierung ist nochmals – für den Zweck dieser Untersuchung – auf eine definitorische Bestimmung dessen, worauf im Folgenden fokussiert werden soll, zurückzukommen: Starke KI ist – zumindest derzeit – eine Zukunftsvision und spielt allenfalls in der Science Fiction eine Rolle. Für den vorliegenden Zweck, KI im Zusammenhang mit Systemen der sozialen Sicherung zu untersuchen, geht es um Anwendungen, die dem Feld der schwachen KI zuzuordnen sind. Der Fokus liegt auf der Lösung bestimmter Aufgaben – wenngleich unbestritten ist, dass die KI-Systemen überantworteten Aufgaben im Laufe der Zeit stets komplexer werden. In diesem Kontext soll im Folgenden der vom Europäischen Parlament veröffentlichten Definition von KI als „zukunftsweisender Technologie“ gefolgt werden:

“Künstliche Intelligenz ist die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren. KI ermöglicht es technischen Systemen, ihre Umwelt wahrzunehmen, mit dem Wahrgenommenen umzugehen und Probleme zu lösen, um ein bestimmtes Ziel zu erreichen. [...] KI-Systeme sind in der Lage, ihr Handeln anzupassen, indem sie die Folgen früherer Aktionen analysieren und autonom arbeiten.“ (Europäisches Parlament, 2023)

Diese Beschreibung lehnt sich an die Phasen der Wahrnehmung, der Schlussfolgerung und des Handelns an, wie sie von Menschen bei der Lösung von Problemen typischerweise durchlaufen werden. Zunächst erfolgt die Erfassung der für die jeweilige Aufgabe relevanten Daten aus der Umgebung mittels Sensoren. Der Kern des KI-Systems besteht in der Transformation der Inputdaten in einen Output, d. h. es wird die vorgegebene Aufgabe gelöst oder eine Handlungsentscheidung getroffen. Daraufhin erfolgt die zur Umsetzung erforderliche Aktion mit Hilfe der zur Verfügung stehenden Aktuatoren.

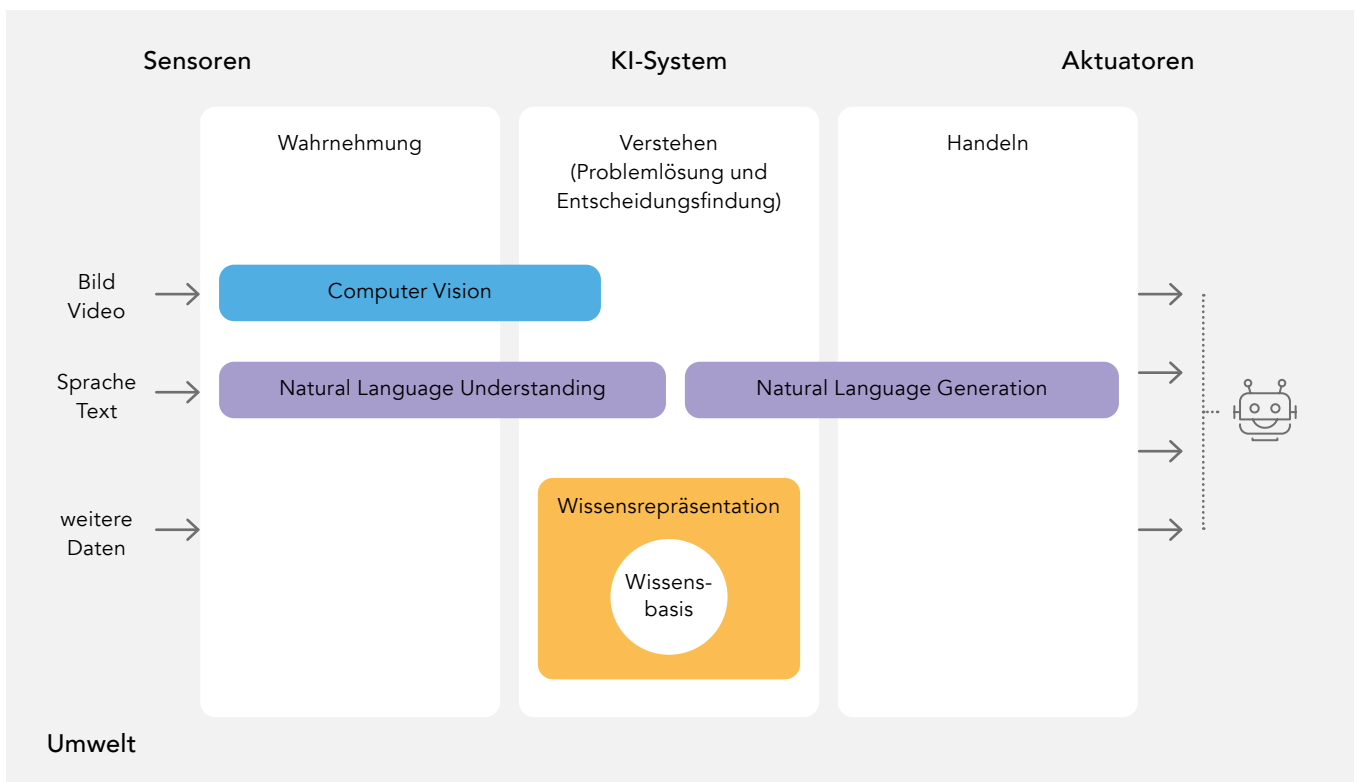


Abbildung 3: Schematische Darstellung eines KI-Systems; Quelle: Eigene Darstellung, in Anlehnung an Lanquillon & Schacht, 2023, S. 16

2.2 EINSATZFELDER DER KÜNSTLICHEN INTELLIGENZ

Wenn man KI grob als das Vermögen von Computersystemen auffasst, Aufgaben zu lösen, die gewöhnlich menschliche Intelligenz voraussetzen, dann lässt sich eine Reihe von Fähigkeiten (z. B. Wahrnehmen, Kommunizieren, Lernen, Wissen, Denken, Handeln) identifizieren, die zur Strukturierung der Einsatzfelder von KI herangezogen werden können. (vgl. Humm et al., 2022, S. 16) Demnach werden häufig die folgenden Teilbereiche der KI genannt:

- **Computer Vision (CV):** Das menschliche Sehvermögen imitierend dient Computer Vision dazu, aussagekräftige Informationen aus digitalen Bildern, Videos und anderen visuellen Eingaben zu gewinnen. „Computerbasiertes Sehen“ basiert darauf, Maschinen zu trainieren, Objekte zu unterscheiden und zu vermessen, Distanzen einzuschätzen, Bewegung festzustellen sowie Auffälligkeiten in Bildern zu erkennen. Hierbei kommen anstelle des menschlichen Sehapparats Kameras, Daten und Algorithmen zum Einsatz. Typische Anwendungsbereiche sind etwa die autonome Navigation von Robotern, das Erkennen visueller Informationen im Straßenverkehr bei Fahrerassistenz oder selbstfahrenden Fahrzeugen, die Inspektion von Produkten zur Feststellung von Qualitätsmängeln sowie die Überwachung von Produktionsanlagen.
- **Natural Language Processing (NLP):** Durch die Kombination von Computerlinguistik mit statistischen und maschinellen Lernmodellen werden Maschinen befähigt, Text und Sprache zu erkennen, zu verstehen und zu generieren. KI erlaubt ein Verständnis der Sprache, das auch den Kontext sowie Absicht und Stimmung des Sprechers bzw. Autors umfasst. NLP bildet den Kern von Anwendungen, die Übersetzungsaufgaben übernehmen, Benutzer anhand von Spracheingaben erkennen oder authentifizieren, auf eingegebene oder gesprochene Befehle oder Fragen reagieren bzw. antworten, Texte zusammenfassen, Absichten oder Stimmungen von Text oder Sprache bewerten und auf Anfrage Inhalte generieren. Im Alltag weit verbreitet ist NLP bereits etwa in Form von Chatbots im Kundenservice, digitalen Assistenten oder sprachgesteuerten GPS-Systemen.
- **Expertensysteme:** Maschinen fungieren gewissermaßen als Experten, indem sie das Wissen eines Fachbereichs abbilden und Menschen bei der Lösung von Problemstellungen in dem jeweiligen Fachbereich unterstützen. Die zur Verfügung gestellten Problemlösungen und Handlungsempfehlungen werden aus einer Wissensbasis abgeleitet, die das formalisierte Expertenwissen in der Regel in Form von Wenn-dann-Regeln repräsentiert. Hierbei erlaubt die KI, dass die Expertensysteme das aus Fakten und Regeln bestehende Wissen interpretieren und eigene Schlussfolgerungen ableiten können. Als typische Aufgabenstellungen von Expertensystemen kommen das Erkennen von Fehlerursachen und die Reduzierung von Arbeitsfehlern, die Vorhersage von Ereignissen auf Basis bestimmter Sachverhalte oder die dialogorientierte, fachspezifische Beratung von Menschen in Betracht.

- **Robotik:** Die Funktionalität traditioneller Roboter, die ohne KI-basierte Steuerung auskommen, ist auf vorprogrammierte Abläufe begrenzt. Die Verbindung von KI und Robotik schafft die Grundlage für die Entwicklung von Software zur Steuerung von Robotern, die diesen erlaubt, über vorgegebene Routinen hinauszugehen und auf unvorhergesehene Ereignisse eigenständig zu reagieren. Damit wird die Ausführung komplexerer Aufgaben, die Anpassung an neue Umgebungen und das Treffen intelligenter Entscheidungen möglich. Anwendungsbereiche KI-basierter Roboter sind etwa autonome Fahrzeuge, die autonom durch den Verkehr navigieren; präzise arbeitende Assistenzsysteme für chirurgische Eingriffe; Produktionsroboter, die Arbeitsprozesse effizienzsteigernd optimieren; Pflegeroboter, die u.a. soziale Interaktionen imitieren.

2.3 AKTUELLE TREIBER DER ENTWICKLUNG VON KÜNSTLICHER INTELLIGENZ

Die jüngsten Erfolge im Einsatz von künstlicher Intelligenz geben Anlass, das Erreichen einer neuen Stufe in der KI-Entwicklung zu konstatieren. Künstliche Intelligenz ist keine Utopie mehr, sondern zählt für viele Menschen – oftmals unbemerkt – zum Alltag. Die weitere technologische Entwicklung und das Potenzial der Nutzung von KI sind nur in groben Umrissen vorauszusehen. Was aber als sicher angenommen werden darf: Die Technologie wird kontinuierlich leistungsfähiger werden und in immer neue Anwendungsbereiche vordringen. Dies ist vor dem Hintergrund der jahrzehntelangen Geschichte von KI freilich keine überraschende Erkenntnis. Doch geben die neuesten Entwicklungen auf dem Feld der künstlichen Intelligenz durchaus Anlass dazu, von einem neuerlichen KI-Frühling zu sprechen: Die rasanten Entwicklungen Anfang der 2020er Jahre führten dazu, dass das Thema KI von den Medien aufgegriffen und der gesellschaftliche Diskurs angefacht wurde. Insbesondere generative KI-Systeme tragen seitdem zur immensen Medienpräsenz von KI bei und lösten einen wahren KI-Boom aus, weil solche Systeme einer breiten Öffentlichkeit kostenlos zugänglich gemacht wurden. Die Möglichkeit, in natürlicher Sprache mit Chatbots wie etwa ChatGPT zu kommunizieren oder KI-Systeme wie beispielsweise Dall-E basierend auf Texteingaben der Nutzer Bilder generieren zu lassen sowie Musik, Stimmen und Videos KI-basiert zu erstellen, übt nicht nur große Faszination aus, sondern wirft darüber hinaus große gesellschaftliche Fragen auf – von Urheberrecht über die Zukunft geistig-kreativer Berufe bis hin zur Verbreitung von Deep Fakes.

Dass die aktuellen Entwicklungen im Bereich der KI zu grundlegenden Veränderungen führen, wird mit der geläufigen Rede von der „KI-Revolution“ auf den Begriff gebracht. Dabei werden die jüngsten Umwälzungen in eine Reihe mit früheren industriellen Revolutionen gestellt, in deren Zuge ebenfalls jeweils eine Technologie radikal und nachhaltig veränderte, wie die Menschheit lebt und arbeitet. War es in der Vergangenheit die körperliche Leistungsfähigkeit des Menschen, so sind es im KI-Zeitalter die geistigen Fähigkeiten des Menschen, auf deren Verbesserung, Ergänzung und Erweiterung Maschinen zielen. (vgl. Bughin & Hazan, 2017)

Der gewaltige Sprung in der Entwicklung künstlicher Intelligenz in den zurückliegenden Jahren hat nicht eine einzige Ursache, sondern ist auf ein ganzes Bündel von Faktoren zurückzuführen. Es ist das Zusammenwirken verschiedener sich gegenseitig verstärkender Entwicklungen, das für den gegenwärtigen Höhenflug der Technologie verantwortlich zeichnet. Hervorzuheben sind insbesondere die folgenden Antriebskräfte:

- Die exponentielle Steigerung von Rechenleistung bietet erst die technischen Voraussetzungen für die Durchführung der im Rahmen von KI zu leistenden komplexen Berechnungen. Zudem sind die Kosten für die Rechenleistung im Laufe der Zeit gesunken, wodurch es erschwinglicher wird, mit KI zu experimentieren und intelligente Anwendungen zu entwickeln. Die gesteigerte Rechenleistung führte zudem zu optimierten Algorithmen für maschinelles Lernen und hat damit wesentlich zu einer Steigerung der Leistungsfähigkeit von KI-Systemen beigetragen. Hoffnungen liegen aktuell auf neuen Computerparadigmen, die einen weiteren Anstieg der Leistungsfähigkeit und Schnelligkeit von Rechnern mit sich bringen: So soll etwa mittels Neuromorphic Computing die selbstorganisierende und selbstlernende Natur des menschlichen Gehirns nachgebildet werden. Quantencomputing soll durch Nutzung der Gesetze der Quantenmechanik die Rechenzeiten für komplexe Aufgaben erheblich beschleunigen.
- Einen unerlässlichen Treibstoff für die weitere KI-Entwicklung bilden Daten. Maschinelles Lernen ist darauf angewiesen, mit Daten „gefüttert“ und auf dieser Basis trainiert zu werden. Digitalisierung und Vernetzung führen zu einem gravierenden Anstieg der Menge an generierten Daten. Das als Big Data bezeichnete Phänomen ist durch eine ständig wachsende Menge an Daten gekennzeichnet, die mit immer größerer Geschwindigkeit erzeugt werden und in verschiedenen Formaten vorliegen. KI-Algorithmen haben also nicht nur immer mehr Daten zur Verfügung, ebenso können vielfältigere Datenquellen genutzt werden. So stehen etwa unterschiedlichste Daten durch die weite Verbreitung von Sensoren sowohl in der industriellen Produktion als auch in Alltagsgegenständen bereit.
- Ein Anstieg der Investitionen in künstliche Intelligenz sowie eine rege Forschungstätigkeit in diesem Bereich tun ihr Übriges, um die Entwicklung und Verbreitung dieser Technologie weiter voranzutreiben. Der von der Stanford University herausgegebene AI INDEX REPORT 2024 zeigt ein eindruckliches Wachstum vieler Indikatoren: So hat sich etwa die Anzahl der wissenschaftlichen Publikationen zum Thema KI zwischen 2010 und 2022 auf weltweit 240.000 beinahe verdreifacht. (vgl. Maslej et al., 2024, S. 31) Ebenso lässt sich insbesondere in den letzten Jahren ein bedeutender Anstieg der KI-Patente verzeichnen – allein von 2021 auf 2022 stieg deren Zahl weltweit um 62,7 Prozent. (vgl. Maslej et al., 2024, S. 38) Ebenso hat sich die Anzahl der KI-Startups und damit einhergehend die Höhe der Investitionen im zurückliegenden Jahrzehnt vervielfacht. (vgl. Maslej et al., 2024, S. 245) Zudem sorgen die großen Cloud-Anbieter wie etwa Microsoft, Amazon, IBM oder Google für eine Senkung der Einstiegschancen, indem sie kostengünstige Entwicklungswerkzeuge für KI auf SaaS-Basis anbieten.

Es existieren jedoch auch eine Reihe von Faktoren, die sich hemmend auf die weitere Entwicklung von künstlicher Intelligenz auswirken können. Dazu zählen etwa das Fehlen KI-spezifischer rechtlicher Rahmenbedingungen und die damit Hand in Hand gehende fehlende Rechtssicherheit. Ebenso können sich Sorgen und Befürchtungen im Hinblick auf KI bremsend auf die weitere Verbreitung auswirken. So äußern Menschen in Deutschland etwa die folgenden Sorgen bezüglich KI: Bedrohung der Demokratie durch gefälschte bzw. verzerrte Inhalte (29 Prozent), die Ersetzung von menschlicher Arbeitskraft (28 Prozent), Auswirkungen auf Privatsphäre und Datenschutz (14 Prozent) und die ungeklärte Haftung bei Schäden bzw. Unfällen (9 Prozent). (vgl. Zandt, 2023) Solchen Befürchtungen negativer Auswirkungen eines KI-Einsatzes kann zwar durch Regulierungsbemühungen wie etwa dem 2023 verabschiedeten AI-Act auf EU-Ebene entgegengewirkt werden, doch rufen einschränkende Maßnahmen regelmäßig von Unternehmensseite immer auch Warnungen vor Überregulierung und damit verknüpftem administrativen Zusatzaufwand sowie Innovationshemmnissen hervor.

Eine vom Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO durchgeführte Befragung¹ (vgl. Feike et al., 2024) hat ergeben, dass neben rechtlichen Risiken vor allem ein Mangel an Zeit Unternehmen in der Praxis davon abhält, den KI-Einsatz voranzutreiben. Neben vielen weiteren genannten Hemmnissen führen ein Mangel an geeigneten Daten und ein Mangel an Personal und technischem Know-how die Liste der Herausforderungen an. „KI-Beginner“, also solche Unternehmen, die sich noch nicht aktiv mit der Thematik auseinandergesetzt haben, werden zudem stark von einem Mangel an Anwendungsfällen ausgebremst.

Ebenso ist ins Kalkül zu ziehen, dass ein positiver Zusammenhang zwischen erfolgreicher KI-Implementierung und Digitalisierungsgrad von Unternehmen besteht. (vgl. Engels, 2023) Der Grad der Digitalisierung von Unternehmen wiederum ist auch von unternehmensexternen Faktoren, wie beispielsweise der Verfügbarkeit von Internet mit hohen Bandbreiten, abhängig.

2.4 VERBREITUNG UND AKZEPTANZ VON KÜNSTLICHER INTELLIGENZ

Eine ganze Reihe von KI-Anwendungen führt vor Augen, dass wir im gewöhnlichen Alltagsleben bereits öfter mit Künstlicher Intelligenz zu tun haben als dies den meisten Menschen bewusst sein dürfte. Wer sich von Sprachassistenten wie Amazons Alexa, Apples Siri oder Google Assistant das Licht an- und ausschalten, Musik abspielen und Fragen beantworten lässt, für automatische Übersetzungen auf Google Translate oder DeepL zurückgreift oder Gesichtserkennung auf Smartphones nutzt, profitiert von KI. Einen erheblichen Schub hat die Bedeutung von KI in der jüngeren Vergangenheit durch die Einführung leicht zugänglicher generativer KI-Anwendungen erfahren – die Veröffentlichung der Chatbot-Anwendung ChatGPT

¹ Die Befragung wurde in der Region Heilbronn-Franken durchgeführt, in der sich derzeit ein Innovationsökosystem entwickelt, welches sich zum Ziel gesetzt hat, Forschung und Entwicklung im Bereich der künstlichen Intelligenz massiv voranzutreiben.

durch das Unternehmen OpenAI im November 2022 rückte das KI-Thema in den Mittelpunkt des öffentlichen Interesses und löste einen regelrechten KI-Boom aus: Innerhalb der ersten fünf Tage erreichte der Chatbot und virtuelle Assistent eine Million Benutzer (vgl. Buchholz, 2023) und brachte es zu einer bemerkenswerten Präsenz in den Medien. Dementsprechend ist dem Großteil der Internetnutzer in Deutschland (81 Prozent) generative KI für die Erstellung von Texten, Bildern oder Programmcode ein Begriff. Mindestens einmal genutzt haben solche Anwendungen jedoch nur ein gutes Drittel dieses Personenkreises (36 Prozent). Dabei fällt die Nutzung bei Erwerbstätigen (40 Prozent) deutlich höher aus als bei Personen ohne Erwerbstätigkeit (27 Prozent). (vgl. Schlude et al., 2023)

Auch im Unternehmensbereich, die Vermutung liegt nahe, hat die maßgeblich durch ChatGPT auf die KI-Technologie gelenkte Aufmerksamkeit zu einer Ausweitung von Akzeptanz und Verbreitung geführt. Auch kleine und mittelständische Unternehmen experimentieren mit der Technologie und schlagen neue Wege ein: So entstehen beispielsweise Bäckereifilialen, die gänzlich ohne Mitarbeiter auskommen, weil Kunden sich selbst bedienen und die KI die gekauften Waren erkennt und via App eine Rechnung schickt. Darüber hinaus hilft KI bei einer akkuraten Mengenplanung, indem sie historische Verkaufszahlen, aber auch Daten zum Wetter oder zu Feiertagen einbezieht und auf dieser Basis relativ genaue Vorhersagen treffen kann. Den Einsatzzwecken sind kaum Grenzen gesetzt – selbst Rezepte werden mittels KI optimiert. Selbst in das als kaum automatisierbar geltende Baugewerbe hat KI Einzug gehalten: Getüftelt wird an Anwendungen, die bei der Abrechnung oder Planung helfen oder die Sortierung von Abfällen unterstützen. Im Malerhandwerk ist zu besichtigen, wie Mensch und Maschine effizient Hand in Hand arbeiten: Der Malerroboter streicht die Wandflächen, der menschliche Maler übernimmt die Details und versorgt den Roboter, füllt Farbe nach, wechselt den Akku. (vgl. Braun, 2024)

Doch wie sieht die Verbreitung von KI in Deutschlands Unternehmenslandschaft insgesamt aus?² Die Zahlen und Statistiken sprechen eine deutliche Sprache: Die Entwicklung der KI-Nutzung in der deutschen Wirtschaft ist seit mehreren Jahren äußerst dynamisch. Es lässt sich ein deutlicher Zuwachs an Unternehmen feststellen, die KI nutzen – und dieser Trend dürfte sich fortsetzen. Immerhin gilt vor dem Hintergrund kontinuierlich steigender Rechenleistung und der rasant zunehmenden Verfügbarkeit großer Datenmengen KI als kaum zu überschätzender Faktor im Wettbewerb. Der Einsatz von KI-Anwendungen eröffnet für Unternehmen ein großes Potenzial zur Produktivitätssteigerung, weil mit Hilfe der Technologie die Erledigung diverser Aufgaben automatisiert sowie Entscheidungsverfahren beschleunigt werden können. (vgl. Damioli et al., 2021; Rammer et al., 2022) Untersuchungen ergeben darüber hinaus, dass Unternehmen, die KI einsetzen, durchschnittlich innovativer sind. (vgl. Cockburn et al., 2019)

2 Eine Vielzahl von Untersuchungen widmet sich in den letzten Jahren dem Thema KI und deren Verbreitung. Die Ergebnisse dieser Studien weichen teilweise drastisch voneinander ab. Dies mag – unter anderen Gründen – auf unterschiedliche Auffassungen des Begriffs künstlicher Intelligenz zurückzuführen sein. Die im Rahmen der vorliegenden Studie präsentierten Resultate sind daher mit Vorsicht zu genießen. Was die Zahlen jedoch zweifelsfrei leisten können – in der Tendenz sind sich die Studienergebnisse einig – ist die Vermittlung eines Stimmungsbildes in Bezug auf die Neigung eines KI-Einsatzes.

Die Ergebnisse diverser Studien der letzten Jahre zur Verbreitung von KI in der deutschen Wirtschaft lassen nur den Schluss einer sehr raschen Diffusion von KI-Technologien zu. Nutzten im Jahr 2019 5,8 Prozent der Unternehmen KI, so stieg dieser Wert in nur zwei Jahren um 4,3 Prozentpunkte auf 10,1 Prozent an. (vgl. Rammer, 2022, S. 11) Zarges et al. (2023, S. 42) zeigen, dass diese Dynamik einer beinahe Verdoppelung der Anzahl KI-nutzender Unternehmen beibehalten werden konnte: Im Jahr 2023 beträgt der erhobene Anteil 19,8 Prozent – jedes fünfte Unternehmen in Deutschland nutzt also aktuell KI. Die Autoren schließen aus diesem rasanten Anstieg der KI-Nutzung, dass „[d]as Potenzial, welches KI für betriebliche Innovationen sowie neue Geschäftsmodelle birgt, [...] somit jedes Jahr für mehr Unternehmen greifbarer“ (Zarges et al., 2023, S. 43) wird. Gemeinsam mit weiteren ca. 36 Prozent der Unternehmen, die angeben, einen KI-Einsatz zu planen, stellt künstliche Intelligenz mithin für die Hälfte der Unternehmen ein bedeutsames Thema dar. Demgegenüber sehen nur ca. 15 Prozent keine Relevanz der KI-Thematik für das eigene Unternehmen und ca. 27 Prozent nutzen KI aktuell nicht und planen auch keinen Einsatz für die Zukunft. (vgl. Zarges et al., 2023, S. 43)

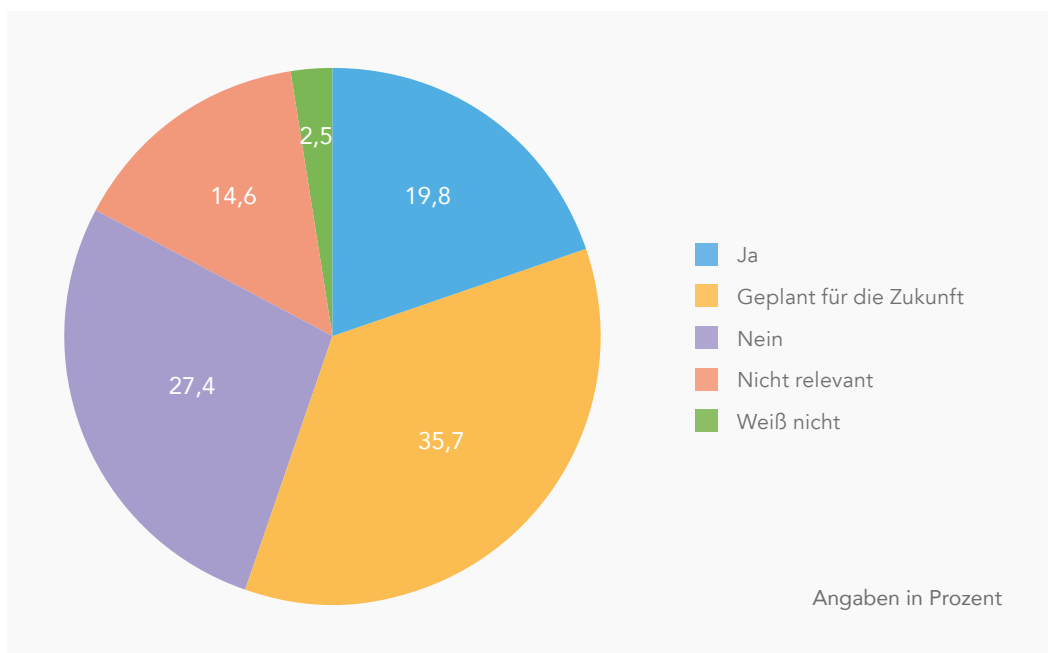


Abbildung 4: Nutzung von KI; Quelle: Eigene Darstellung, in Anlehnung an Zarges et al., 2023, S. 43

Dabei wird KI deutlich häufiger in Großunternehmen eingesetzt. Während sich im Kreis der Unternehmen mit mehr als 250 Beschäftigten bereits 30,1 Prozent finden, die KI nutzen, beträgt dieser Anteil bei Unternehmen mit 50 bis maximal 249 Beschäftigten nur knapp die Hälfte (14,7 Prozent). Mit zunehmender Unternehmensgröße scheint die Relevanz des Themas KI zu steigen: Mehr als ein Viertel (27,3 Prozent) der kleinsten Betriebe mit maximal neun Beschäftigten hält KI-Nutzung für nicht relevant für das eigene Unternehmen. Auch in Bezug auf die Planung des KI-Einsatzes scheint ein Zusammenhang mit der Unternehmensgröße gegeben zu sein: Von den Unternehmen mit maximal neun Mitarbeitern befassen sich mehr als 15 Prozent mit Vorbereitungen im Hinblick auf einen KI-Einsatz, wohingegen dieser Anteil bei Unternehmen mit über 250 Beschäftigten knapp 45 Prozent beträgt. Insgesamt lässt sich also sagen, dass mit steigender Zahl der Beschäftigten in einem Unternehmen ein KI-Einsatz eher stattfindet oder zumindest geplant wird. (vgl. Zarges et al., 2023, S. 44)

Im Vergleich der Wirtschaftssektoren zeigen sich deutliche Unterschiede in der Haltung gegenüber KI. Weit abgeschlagen bildet das Baugewerbe mit 3,5 Prozent von KI-einsetzenden Unternehmen das Schlusslicht. Der Anteil der Unternehmen dieses Sektors, die keine Relevanz für KI sehen (21,9 Prozent) ist hingegen höher als in den übrigen Sektoren. Mehr als jedes fünfte Unternehmen des produzierenden Gewerbes sowie der Unternehmensdienstleistungen nutzt aktuell bereits KI. Eine dynamische Entwicklung verspricht dabei insbesondere der Bereich des produzierenden Gewerbes: Mehr als 41 Prozent planen einen KI-Einsatz und nur etwa jedes zehnte Unternehmen betrachtet KI als irrelevant. Der Sektor der übrigen Dienstleister³ ist Spitzenreiter bei der KI-Nutzung: Ein Viertel hat bereits KI im Einsatz und fast 38 Prozent planen einen solchen. Diese hohen Werte sind in erster Linie auf den Informations- und Kommunikationsbereich (51 Prozent Nutzung; 37,6 Prozent geplante Nutzung) und das Finanz- und Versicherungswesen (30,5 Prozent Nutzung; 58,6 Prozent geplante Nutzung) zurückzuführen. (vgl. Zarges et al., 2023, S. 45)

3 Zu den übrigen Dienstleistern zählen die folgenden Branchen: Finanz- und Versicherungsdienstleistungen; Informations- und Kommunikationsdienstleistungen; Kunst, Unterhaltung und Erholung; Grundstücks- und Wohnungswesen; Sonstige wirtschaftliche Dienstleister; Verkehr und Lagerei.

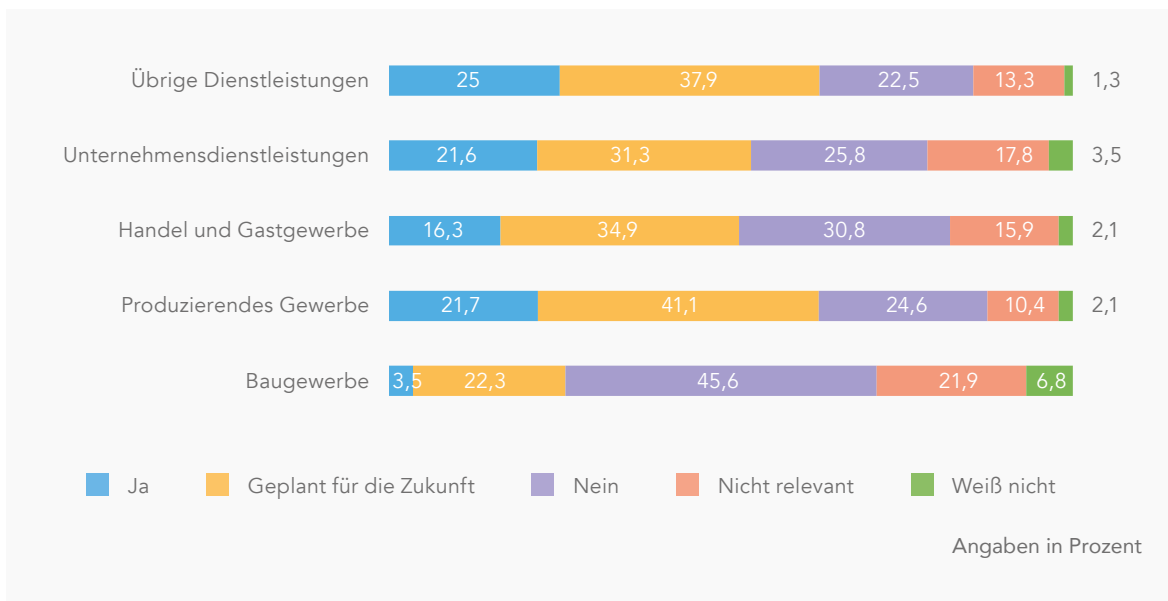


Abbildung 5: Nutzung von KI nach Wirtschaftssektoren; Quelle: Eigene Darstellung, in Anlehnung an Zarges et al., 2023, S. 45

Diese Ergebnisse sind im Lichte des Digitalisierungsgrads von Unternehmen zu interpretieren. Engels (2023, S. 525) weist richtigerweise darauf hin, dass „[e]ntscheidend für eine erfolgreiche Implementierung von KI ist, wie digital das Unternehmen ist, in dem diese implementiert werden soll“. Die Automatisierung durch KI setzt voraus, dass Unternehmensbereiche digital vernetzt und Unternehmensprozesse digital umgesetzt sind. Darüber hinaus liegt die Annahme auf der Hand, dass sich KI-Anwendungen einfacher in digital reife Unternehmen eingliedern lassen. Der jährlich im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz erhobene Digitalisierungsindex zeigt auf, dass die verschiedenen Branchen und Unternehmensgrößen im Hinblick auf Digitalisierung sehr unterschiedliche Reifegrade aufweisen. Die aktuellen Ergebnisse (vgl. Büchel et al., 2024) deuten zwar darauf hin, dass sich die Lücken schließen und die deutsche Wirtschaft hinsichtlich Digitalisierung homogener wird, doch nichtsdestotrotz bleiben Digitalisierungsunterschiede: Spitzenreiter bei der Digitalisierung ist die Branche Informations- und Kommunikationstechnologie, während die Baubranche im Jahr 2023 zwar Digitalisierungszuwächse verzeichnete, jedoch weiterhin das Schlusslicht bildet. Auch unternehmensnahe Dienstleistungen schneiden im Hinblick auf die Digitalisierung der Unternehmen gut ab. Zudem ist ein klarer Zusammenhang zwischen Digitalisierungsgrad und Unternehmensgröße erkennbar. Je größer Unternehmen sind, desto höher fällt deren Digitalisierungsindex aus.

Neben dem Digitalisierungsgrad, darauf weist Engels (2023, S. 527) hin, spielt auch der Umgang mit Daten im Unternehmen eine entscheidende Rolle für die erfolgreiche Implementierung von KI. Weil Daten das Fundament der Entwicklung und Ausführung von KI-Anwendungen darstellen, kommt es maßgeblich auf deren umfassende Speicherung, effizientes Management und intelligente Nutzung an. Auch betreffend ihrer sogenannten „Data Economy Readiness“⁴ weisen große Unternehmen einen Vorsprung auf. Während im Jahr 2022 77 Prozent der großen Unternehmen (mindestens 250 Beschäftigte) als „data economy ready“ gelten, nimmt dieser Wert mit der Unternehmensgröße sukzessive ab: 58 Prozent der mittleren (50 bis 249 Beschäftigte) und 30 Prozent der kleinen (bis 49 Beschäftigte) Unternehmen wird das Prädikat „data economy ready“ verliehen. (vgl. Büchel & Engels, 2022, S. 2) Betrachtet man mit Engels (2023, S. 527) Data Economy Readiness als Indikator dafür, ob Unternehmen die Voraussetzungen einer erfolgreichen KI-Implementierung und Nutzung aufweisen, so bleibt nur der Schluss: Große Unternehmen sind deutlich besser aufgestellt, um KI erfolgreich zu nutzen. Insgesamt stellt der Umgang mit Daten eine nicht zu vernachlässigende Stellschraube dar, damit KI ihr Potenzial in der deutschen Wirtschaft vollständig entfalten kann.

Liegt der Fokus „traditioneller“ KI auf dem Unternehmensbereich, so lässt sich weltweit beobachten, dass sich generative KI ihren Weg auch in das Privatleben von Menschen bahnt. Nicht nur steigt die Wahrscheinlichkeit, mit der Menschen KI in ihrem Berufsleben nutzen, ebenso wird es immer wahrscheinlicher, dass sie die Technologie sowohl beruflich als auch privat einsetzen. (vgl. McKinsey, 2024)

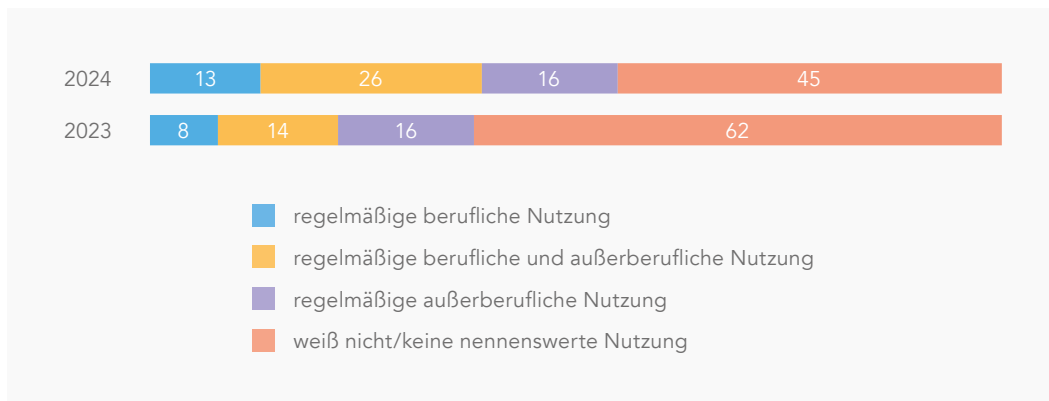


Abbildung 6: Die Erfahrung mit generativer KI nimmt zu; Quelle: Eigene Darstellung, in Anlehnung an McKinsey, 2024

4 Als „data economy ready“ bezeichnen Büchel & Engels (2022, S. 1) solche „Unternehmen, die die Voraussetzungen erfüllen, um Daten effizient zu bewirtschaften“. Diese Unternehmen erfüllen durchschnittlich mehr als die Hälfte der als relevant erachteten Aspekte aus den Bereichen Speicherung (z. B. Umfang der gespeicherten Prozess- oder Lieferantendaten), Management (z. B. Qualitätsprüfung, interner Datenaustausch über standardisierte und permanente Schnittstellen) und Nutzung (z. B. Daten zur Prozessautomatisierung, zur Produktentwicklung, zur Marktprognose) von Daten. Wird Unternehmen die Data Economy Readiness bescheinigt, so „haben sie die Voraussetzungen dafür, das theoretische Potenzial, das Datenmengen bieten, zu realisieren und damit die Vorteile der Digitalisierung für Innovation und nachhaltiges Wachstum zu nutzen“ (Büchel & Engels, 2022, S. 1).

Auch ein Blick auf die Marktdaten zeigt deutlich, dass künstliche Intelligenz nach ihrer bewegten Entwicklungsgeschichte den Schritt aus dem Forschungslabor in die kommerzielle Praxis geschafft hat. Dabei stehen die Zeichen für die weitere Entwicklung von KI ganz klar auf Wachstum: Weltweit wird für die wichtigsten Einsatzfelder von KI in den kommenden Jahren ein Anstieg der Marktgröße vorausgesagt. Zeigten die ökonomisch turbulenten Jahre Anfang der 2020er Jahre große Marktschwankungen, so wird ab 2022 ein stetiges Wachstum vorhergesagt.

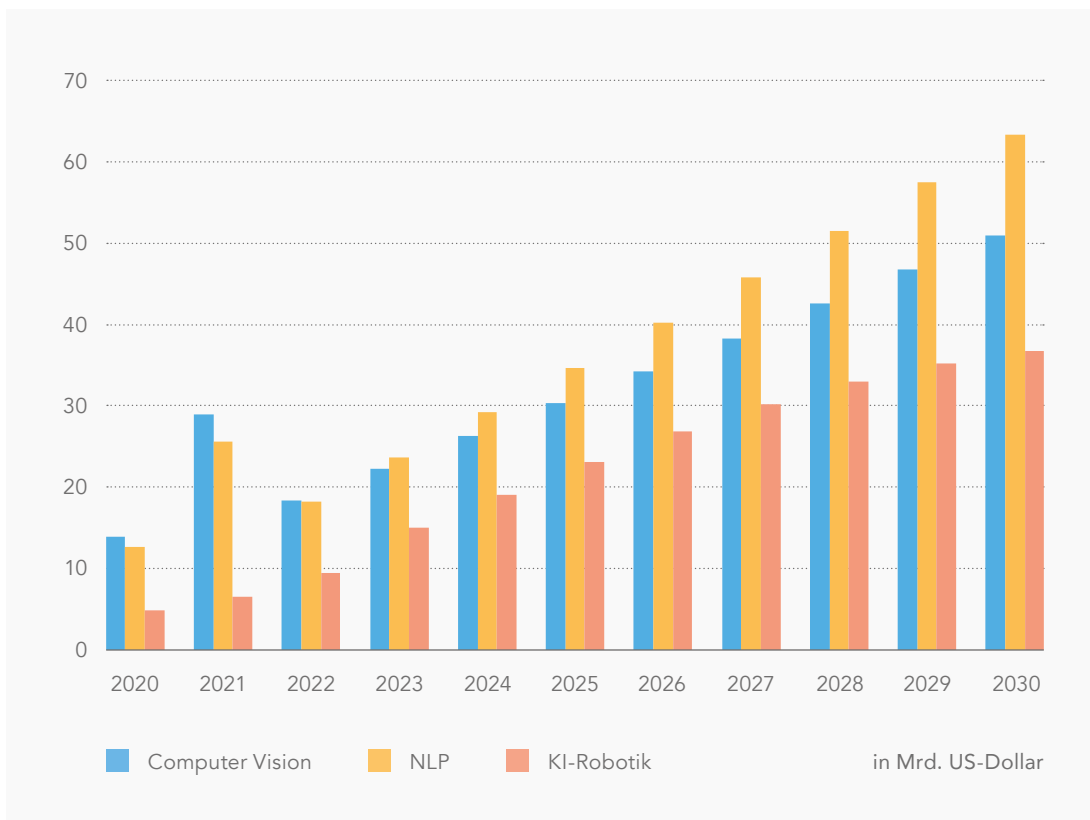


Abbildung 7: Der KI-Markt wächst stetig; Quelle: Eigene Darstellung, in Anlehnung an Statista, 2024b, 2024c, 2024e

Auch der Blick auf Schätzungen bezüglich der Marktgröße von generativer KI zeigt ein stetes Wachstum für die kommenden Jahre. Verdoppelte sich die Marktgröße Anfang der 2020er Jahre jeweils jährlich, so wird ab 2023 ein Wachstum der Größe des Markts für generative KI um jeweils 20 Mrd. US-Dollar (weltweit) bzw. 6 Mrd. US-Dollar (Europa) geschätzt.

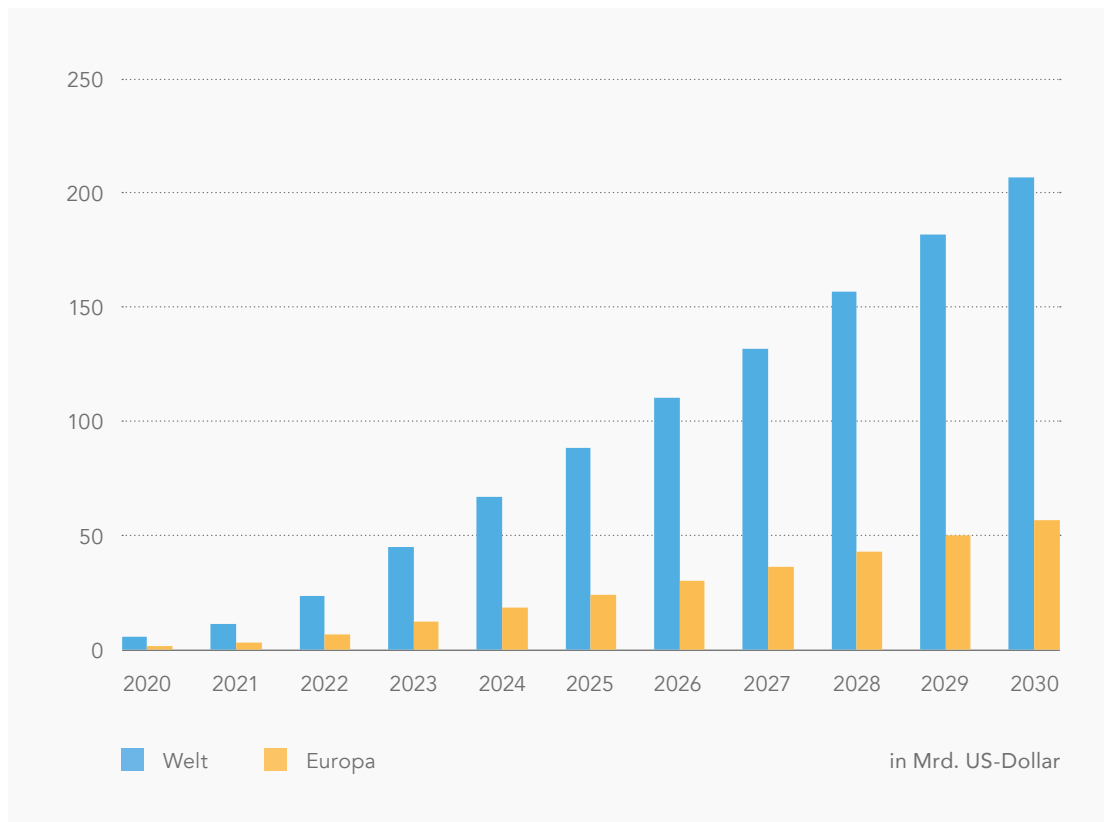


Abbildung 8: Der Markt für generative KI weitet sich stetig aus; Quelle: Eigene Darstellung, in Anlehnung an Statista, 2024a, 2024d

Es ist davon auszugehen, dass künstliche Intelligenz weiter an Verbreitung und Akzeptanz gewinnen wird, zumal auf politischer Ebene Übereinstimmung herrscht, dass KI eine förderungswürdige Technologie darstellt. In Deutschland forciert die Bundesregierung mit ihrer 2018 ins Leben gerufenen und 2020 fortgeschriebenen Nationalen KI-Strategie den Weg Deutschlands zum führenden Standort für die Entwicklung und Anwendung von künstlicher Intelligenz. Die Strategie zielt darüber hinaus auf einen verantwortungsvollen und gemeinwohlorientierten Umgang mit der Technologie sowie deren ethische, rechtliche, kulturelle und institutionelle Einbettung in die Gesellschaft im Rahmen eines breiten gesellschaftlichen Dialogs und einer aktiven politischen Gestaltung.

Ebenso wie die deutsche Bundesregierung haben auch Regierungen anderer Länder in den vergangenen Jahren spezifische Strategien oder Eckpunktepapiere zum Umgang mit KI ausgearbeitet. Die Europäische Kommission veröffentlichte im April 2018 eine AGENDA ZUR FÖRDERUNG KÜNSTLICHER INTELLIGENZ (Europäische Kommission, 2018) in Europa. Anfang 2020 folgte mit dem WEIßBUCH ZUR KÜNSTLICHEN INTELLIGENZ (Europäische Kommission, 2020) eine Diskussionsgrundlage für einen auf Exzellenz und Vertrauen basierenden Rahmen für vertrauenswürdige KI. Daraufhin folgte im April 2021 mit der Veröffentlichung des

KOORDINIERTEN PLANS FÜR KÜNSTLICHE INTELLIGENZ (Europäische Kommission, 2021) die Darstellung der nächsten Schritte der EU zu globaler Führerschaft im Bereich vertrauenswürdiger KI. Mit dem gleichzeitig durch die EU-Kommission veröffentlichten Entwurf zur Regulierung riskanter KI-Systeme war der Anstoß zu einem europäischen ARTIFICIAL INTELLIGENCE (AI) ACT gegeben, der zu einer vertrauenswürdigen KI beitragen soll. Eine endgültige Version des EU AI ACTS (Europäisches Parlament, 2024) wurde im Dezember 2023 nach den erfolgreich abgeschlossenen Trilog-Verhandlungen zwischen den EU-Institutionen beschlossen. Damit haben die EU-Mitgliedstaaten das weltweit erste Gesetz zur Regulierung von KI beschlossen. Der AI Act wird in Deutschland Gesetzescharakter haben. (vgl. Kap. 5)

3 EINSATZ VON KI IM BEREICH DER SOZIALVERSICHERUNG

Im Bereich der sozialen Sicherung haben die für deren Durchführung verantwortlichen Organisationen entscheidende Schritte unternommen, um KI in ihre Kern-tätigkeiten zu integrieren. Dabei lässt sich die Rolle der KI im Bereich der sozialen Sicherung nicht trennen von der fortschreitenden Digitalisierung in sämtlichen Lebensbereichen und nur vor dem Hintergrund der historischen Entwicklung des IT-Einsatzes in der öffentlichen Verwaltung im Allgemeinen und im Bereich der sozialen Sicherung im Besonderen bewerten. Der Einsatz künstlicher Intelligenz im Bereich der sozialen Sicherung kann als Teil einer datengesteuerten Strategie verstanden werden, um die Verwaltungseffizienz in der sozialen Sicherung zu erhöhen. Dabei ist zu beobachten, wie sich der Computereinsatz zur Unterstützung der Aufgaben der sozialen Sicherung insofern wandelte, als es in den Anfangstagen stärker um Entlastung in Bezug auf Routinetätigkeiten ging und die Computerisierung danach in Bereiche einzog, in denen ein differenziertes und professionelles Urteilsvermögen erforderlich ist. (vgl. Henman, 2022, 267f.) Hand in Hand mit der zunehmenden Verbreitung von Computern im Verwaltungshandeln geht die steigende Verfügbarkeit von digitalen Daten, die als Grundlage für Statistik und Data Analytics dienen. Mit dem technischen Fortschritt entstanden sogenannte Entscheidungsunterstützungssysteme und Expertensysteme, welche Entscheidungsträgern in komplexeren Fällen im Bereich der sozialen Sicherung zur Seite stehen sollten. Solche Systeme haben sich in Verbindung mit Big Data und Maschinenlernen weiterentwickelt und erlauben zunehmend differenzierte, zielgerichtete und personalisierte Ansätze in ihrem jeweiligen Einsatzbereich.

Dabei soll die Technisierung des Bereichs der sozialen Sicherung nicht nur dazu beitragen, umfassenden Sozialschutz zu gewährleisten, sondern ebenso die Kostenexplosion einzudämmen. Letztlich geht es darum, durch eine verbesserte, effizientere Ausführung der Aufgaben der Sozialversicherungsorganisationen zu einer verbesserten Leistung zu gelangen, wodurch sich insgesamt der Einsatz künstlicher Intelligenz im Rahmen der sozialen Sicherungssysteme als gesellschaftlich nützlich erweisen könnte.

Im Folgenden werden zunächst mögliche Einsatzfelder künstlicher Intelligenz in der sozialen Sicherung im Allgemeinen vorgeschlagen, um daran anschließend konkrete nationale und internationale Beispiele der KI-Nutzung in der Sozialverwaltung vorzustellen. Schließlich werden Voraussetzungen und Erfolgsfaktoren eines effektiven Einsatzes von KI in Organisationen der sozialen Sicherung umrissen.

3.1 MÖGLICHE ANWENDUNGSFELDER

Das derzeitige Interesse an künstlicher Intelligenz als Instrument der Effizienzsteigerung im öffentlichen Sektor war nie zuvor größer. Aufgrund der immensen Datenmassen, mit denen es im Bereich der sozialen Sicherungssysteme umzugehen gilt, sind Organisationen der Sozialverwaltung ein natürlicher Kandidat für die Nutzung von KI-Technologie. Tatsächlich gibt es viele Anknüpfungspunkte für eine KINutzung in den Sozialversicherungsorganisationen. Um die Potenziale für KI in den Sozialversicherungsorganisationen zu identifizieren, soll aus prozessorientierter Perspektive auf die Aufgabe der sozialen Sicherung geblickt werden, um ausgehend von einem generischen Leistungserstellungsprozess der Sozialversicherung beispielhaft Anwendungsmöglichkeiten für KI zu skizzieren.

3.1.1 Leistungserstellungsprozess der Sozialversicherung

Dem einzelnen Menschen Schutz für die Wechselfälle des Lebens zu bieten, ist Zweck des Sozialversicherungssystems als in Deutschland wichtigstem Teil der Aufgabe der sozialen Sicherung. Als gesetzliches Versicherungssystem umfasst die Sozialversicherung die Absicherung gegen allgemeine Lebensrisiken wie Krankheit, Unfall, Arbeitslosigkeit und Pflegebedürftigkeit sowie für das Alter. Die verschiedenen Arten von Lebensrisiken strukturieren das Sozialversicherungssystem in verschiedene Zweige – dementsprechend gliedert sich die deutsche Sozialversicherung in eine Kranken-, Unfall-, Renten-, Arbeitslosen- und Pflegeversicherung. Eine solche Strukturierung drängt sich auf, weil die verschiedenen Risiken im Falle ihres Eintretens jeweils Anknüpfungspunkt unterschiedlicher Ansprüche und Leistungen sind. Auch in organisatorischer Hinsicht wird zwischen verschiedenen Trägern der Sozialversicherung unterschieden, die jeweils für die Ausführung der Aufgaben des jeweiligen Sozialversicherungszweigs zuständig sind.

Aus einer organisationstheoretischen Perspektive bringt es diese strikte Trennung in Versicherungszweige mit sich, dass die Gemeinsamkeiten zwischen den Versicherungsträgern im Schatten bleiben. Auf den ersten Blick scheint der gemeinsame Nenner, auf den die Sozialversicherungsorganisationen zu bringen sind, ihr gemeinsamer sozialer Zweck zu sein. Im Vordergrund steht zumeist das „Was“, d. h. welche Ziele verfolgt und Aufgaben erfüllt werden, unterbelichtet bleibt hingegen das „Wie“, d. h. die Prozesse und Methoden, mit Hilfe derer Sozialversicherungsträger ihrem gesetzlichen Auftrag nachkommen. Die Interna dieser Organisationen sind nur selten Gegenstand der Betrachtung. So sind Brussig (2017, S. 590) zufolge die Ablaufstrukturen der Sozialversicherungsorganisationen im Vergleich zu anderen Organisationstypen wissenschaftlich nur wenig ausgeleuchtet. Allenfalls werden interne Prozesse von Rechnungshöfen oder von den Sozialversicherungsorganisationen selbst in den Blick genommen.

Für den vorliegenden Zweck der Ausleuchtung des KI-Einsatzes im Bereich der sozialen Sicherung erscheint es angebracht, Sozialversicherungsorganisationen als eigenständigen Organisationstyp in den Fokus zu rücken und hierbei weniger auf die Unterschiede zwischen den einzelnen Zweigen der Sozialversicherungen zu blicken als vielmehr auf deren Gemeinsamkeiten, die zugleich Differenzierungsmerkmale zu anderen Organisationstypen darstellen. Hervorstechendstes

Merkmal von Sozialversicherungsorganisationen ist bestimmt ihr hoher Institutionalisierungsgrad; ihr Bestand ist gesetzlich gesichert und sie setzen einen hoheitlichen Auftrag um. Mit der „Zwangsmitgliedschaft“ ist eine weitere Besonderheit gegeben: Versicherte werden als „Mitglied“ beschrieben, denen allerdings die Möglichkeit zum Austritt als typisches Merkmal einer Organisationsmitgliedschaft fehlt. Brussig (2017, S. 596) führt aus, dass diese Begrifflichkeit nahelegt, auch die Versicherten als Bestandteil der Organisation zu betrachten, doch anders als die Beschäftigten sind die Versicherten nicht zur organisatorischen Aufgabenerfüllung nötig. Zugleich führt die Redeweise von Versicherten als „Kunden“ in die Irre, wie dies insbesondere von den Arbeitsagenturen praktiziert wird, weil Versicherten eine Konsumentensouveränität nur in sehr beschränktem Maße zukommt. Zugleich unterliegen die Versicherten Mitwirkungspflichten. Darüber hinaus sind ein starker hierarchischer Aufbau mit eindeutigen Aufgabenzuweisungen und – aufgrund der engen Bindung an einen rechtlichen Auftrag – vorhersehbare Entscheidungen und Abläufe charakteristisch für Sozialversicherungsorganisationen. (vgl. Brussig, 2017, 592ff.)

Insbesondere der letztgenannte Aspekt, die Art und Weise der Ausführung der Sozialversicherungsaufgaben, ist – zumindest aus einer übergeordneten Perspektive – ein verbindendes Element zwischen den Sozialversicherungsorganisationen der verschiedenen Zweige, das für den Gang unserer Untersuchung von Interesse ist. In diesem Rahmen erscheint es durchaus sinnvoll, sich auf die Suche nach prozessualen Gemeinsamkeiten zwischen den Organisationen der verschiedenen Sozialversicherungszweige zu machen. Eine nähere Betrachtung der betreffenden Organisationen fördert ein höheres Maß an Gemeinsamkeiten hinsichtlich der operationalen Aspekte der Sozialversicherung zutage als dies auf den ersten Blick zu vermuten ist. Tatsächlich weisen die unterschiedlichen Sozialversicherungsleistungen allesamt einen ähnlichen Leistungserstellungsprozess auf.

Mit einer solcherart prozessorientierten Sicht auf die Organisationen der Sozialverwaltung soll zum Ausdruck gebracht werden, dass sämtliche Aktivitäten innerhalb der Sozialversicherungsorganisationen in einer organisationstheoretischen Sichtweise als Geschäftsprozesse angesehen werden können. In einer stark abstrahierenden Perspektive lässt sich feststellen, dass innerhalb der Organisationen der verschiedenen Sozialversicherungszweige vergleichbare Aufgaben durchgeführt werden, die in eine Reihe von Geschäftsprozessen eingebettet sind und in ihrer Gesamtheit einen Leistungserstellungsprozess ergeben, der für alle Zweige und mithin für die Sozialversicherung Gültigkeit hat. Im Sinne der Prozessorientierung in der Organisationstheorie wird der Fokus hierbei auf die Tätigkeiten und Aufgaben, also das „Wie“, anstatt auf die Leistung, also das „Was“, gerichtet. Dabei zeichnen sich Geschäftsprozesse dadurch aus, dass sie aus einem oder mehreren Inputs einen Output generieren, der dessen Adressaten einen Mehrwert stiftet. Auf den Bereich der Sozialversicherung angewandt bedeutet dies nichts anderes, als dass durch die Transformation bestimmter Informationen Sozialleistungen erbracht werden.

Eine solche auf die Informationseingabe- und -ausgabe-Beziehungen fokussierende Geschäftsprozessperspektive auf die Organisationen der sozialen Sicherung liefert das Grundgerüst, um jene Prozessschritte zu identifizieren, die sich für eine Nutzung künstlicher Intelligenz eignen. Hierbei sind im Bereich der Sozialversicherung einige Besonderheiten von Interesse: In Bezug auf die Abläufe ist herauszustellen, dass sich die Durchführung der Sozialversicherung durch eine sehr hohe Zahl von Geschäftsvorgängen auszeichnet. Dabei ist es eine Vielzahl verschiedener Daten aus den unterschiedlichsten Quellen, die im Bereich der sozialen Sicherung zum Tragen kommt. Die bei den Geschäftsvorgängen in der Sozialversicherung anfallenden Daten unterscheiden sich von Daten in anderen Bereichen allein schon dadurch, dass die gesetzliche Sozialversicherung für die meisten Beteiligten eine Pflichtversicherung darstellt und sie sich dem System deshalb nicht entziehen können. Weil sich der Umfang des Datenaustauschs aus den betreffenden rechtlichen Bestimmungen ergibt, hat der Einzelne keinen Einfluss darauf, welche seiner Daten von welcher Stelle in welcher Form an wen weitergeleitet werden. Neben dem Zwangscharakter des Systems ergibt sich auch aus dem Umstand, dass sensible Daten (z. B. Einkommen, Krankheiten, familiäre Situation) verarbeitet werden, eine besondere Schutzwürdigkeit der Daten. Eine weitere Besonderheit liegt im hohen Grad der gesetzlichen Reglementierung: Das Sozialgesetzbuch enthält detaillierte Vorschriften zum Umgang mit Daten⁵ und verweist damit den Einsatz von KI-Technologie in enge Schranken. (vgl. Geißler, 2017, 2f.)

Um nachfolgend mögliche Einsatzfelder von künstlicher Intelligenz im Bereich der Sozialversicherungssysteme aufzuzeigen, sollen zunächst die einzelnen Prozesse des schematischen Leistungserstellungsprozesses der Sozialversicherung als Anknüpfungspunkte von KI-Anwendungen umrissen werden:⁶ (vgl. etwa Barca & Hebbbar, 2021; Lindert et al., 2020)

- **Information & Kommunikation:** Grundlegend für das Funktionieren einer sozialen Absicherung ist, dass Menschen über Informationen hinsichtlich der Ziele, Funktionsweise und Regeln, Prozesse und Anspruchskriterien der einzelnen Sozialversicherungszweige und deren Leistungen verfügen. Es muss sichergestellt sein, dass schutzbedürftige Gruppen erreicht und die Informationsbedarfe von (potentiell) Begünstigten erfüllt werden. Als Startpunkt des Leistungserstellungsprozesses legen Informationsmaßnahmen den Grundstein für die Effektivität der Sozialversicherung. Gleichermaßen sind Informations- und Kommunikationsmaßnahmen jedoch im Zuge aller nachfolgenden Phasen der Leistungserstellung von Bedeutung.

5 So sind etwa personenbezogene Daten gemäß § 67a SGB X grundsätzlich bei den betroffenen Personen zu erheben (Direkterhebungsgrundsatz). Ausnahmsweise dürfen Daten jedoch auch ohne Mitwirkung der Betroffenen erhoben werden. Für den Bereich der Rentenversicherung gilt beispielsweise, dass Zeiten, in denen Personen beschäftigt, arbeitslos oder krank sind, von den zuständigen Stellen, d. h. dem Arbeitgeber, der Arbeitsagentur bzw. der gesetzlichen Krankenkasse, auf elektronischem Weg an die Deutsche Rentenversicherung gemeldet werden, die die Angaben dann dem Versicherungskonto zuordnet. Soweit dies die Erfüllung bestimmter Aufgaben erfordert, können Daten auch direkt bei den zuständigen Stellen erfragt werden.

6 Im Rahmen dieser Untersuchung wird auf die Bereitstellung der Sozialversicherungsleistungen für die Versicherten fokussiert, während die Finanzierungsseite unberücksichtigt bleibt. Eine vollständige Prozesskette, die sämtliche Aktivitäten der Sozialversicherung in den Blick nimmt, müsste selbstredend auch die mit Beitragseinzahlung und -verwaltung befassten Aufgabenschritte umfassen, wie dies etwa La Salle (2021) vorsieht.

- **Datenerfassung:** Da für den gesamten weiteren Prozess die Verfügbarkeit von Daten eine große Rolle für reibungslose Abläufe spielt, sind zunächst im Rahmen einer Registrierung der betreffenden Personen initiale Informationen einzuholen. Gewöhnlich erhalten die (potentiell) Begünstigten nach der Registrierung eine eindeutige Kennung („Sozialversicherungsnummer“). Die Daten können unterschiedlichen Quellen entstammen: entweder werden sie von den Versicherten erfragt oder sie sind bereits in sonstigen staatlichen Datenquellen enthalten und werden dort entnommen oder aber sie werden von Unternehmen im Zuge der Anmeldung zur Sozialversicherung anlässlich der Arbeitsaufnahme übermittelt. In jedem Fall umfasst dieser Prozessschritt eine ausreichende Verifizierung und Validierung der Daten zu deren Nutzung für die nachfolgenden Prozesse.
- **Anspruchsprüfung:** Die Prüfung der Anspruchsberechtigung umfasst die Gegenüberstellung der den einzelnen Sozialversicherungszweigen zugrundeliegenden Qualifikationskriterien (z. B. demographische und sozioökonomische Faktoren, Beschäftigungsstatus und -historie, Versicherungsbeiträge) sowie der im Zuge der Registrierung erstellten Personenprofile. Viele Leistungen basieren auf einer Kombination von Kriterien mit Schwellenwerten für die Anspruchsberechtigung in Bezug auf beispielsweise Einkommen, Vermögen etc.
- **Antragsmanagement:** Um tatsächlich Unterstützungsleistungen zu erhalten, ist ein formaler Akt des Leistungsabrufs, etwa in Form eines Antrags, erforderlich. Daraufhin erfolgt die Bestimmung der Art bzw. Höhe der konkreten Unterstützung. Dieser Prozessschritt kann je nach Versicherungszweig irgendeine Form der initialen Benachrichtigung des Versicherten sowie zusätzliche Maßnahmen der Datenerfassung (zwecks praktischer Durchführung, z. B. Kontoinformationen) umfassen, sodass der Datensatz des Versicherten aktualisiert werden kann.
- **Leistungserbringung:** Die Aufgaben in der Phase der Erbringung der Unterstützungsleistung hängen stark vom jeweiligen Versicherungszweig ab. Grundsätzlich kann es sich entweder um Geldzahlungen oder verschiedene Arten von Dienstleistungen handeln. Der Prozessschritt der Leistungserbringung ist häufig diejenige Phase mit dem intensivsten Kontakt zwischen Sozialversicherungsorganisation und Begünstigten.
- **Monitoring & Qualitätsmanagement:** Schließlich gilt es, die Interaktionen mit den Versicherten begleitend zur eigentlichen Aufgabenausführung zu überwachen und zu verwalten. Die Sicherung der Qualität der Abläufe erfordert ein solides und reaktionsschnelles Verfahren für die Bearbeitung von Beschwerden, Klagen und Fehlern. Zudem muss die Einhaltung von Bedingungen, die gegebenenfalls an den Leistungsbezug geknüpft wurden, überwacht werden. Ebenso ist fortlaufend zu prüfen, ob weiterhin überhaupt ein Anspruch auf die jeweilige Unterstützungsleistung besteht. Neben dem Monitoring der administrativen Aspekte umfasst dieser Prozessschritt auch eine Überwachung der Leistungsfähigkeit der Prozessdurchführung in den einzelnen Sozialversicherungszweigen,

um sicherzustellen, dass die gewünschten Ergebnisse erreicht werden. Dies umfasst, Prozesse fortlaufend zu verbessern und Bereiche mit Optimierungspotenzialen zu identifizieren.

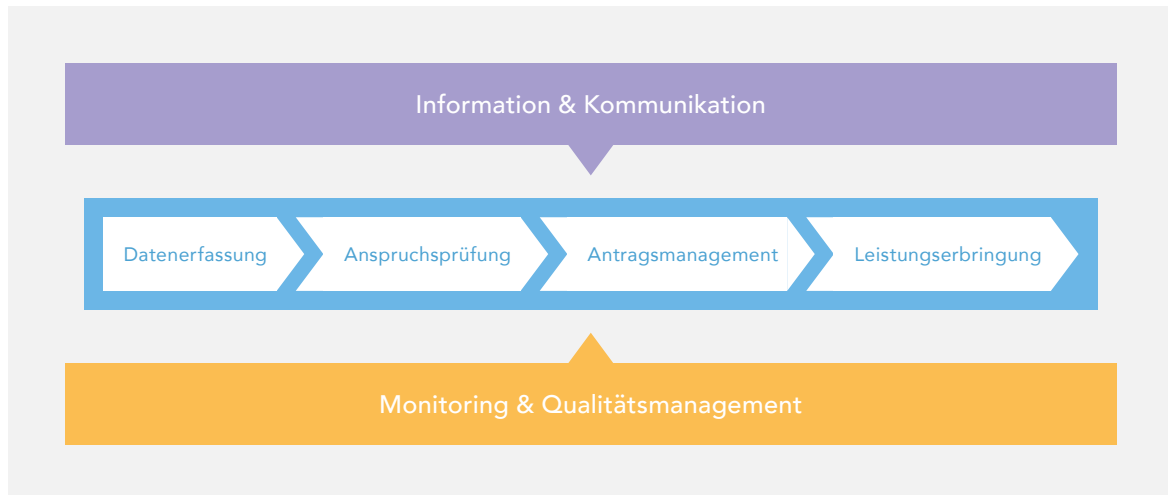


Abbildung 9: Schematischer Leistungserstellungsprozess der Sozialversicherung;
Quelle: Eigene Darstellung

3.1.2 Potentielle Anwendungen von KI entlang des Leistungserbringungsprozesses

Grundsätzlich ist es dem Wesen künstlicher Intelligenz entsprechend nun möglich, Sozialversicherungsorganisationen in jedem der oben beschriebenen Prozessschritte in mehrerlei Hinsicht zu unterstützen. Zunächst kann KI immer dort zum Einsatz gebracht werden, wo es um die Identifikation und Kategorisierung von Informationen geht. Mit Hilfe von KI-Anwendungen lassen sich spezifische Objekte, Merkmale oder auch Datenmuster nicht nur viel schneller erkennen, sondern entdeckt die Maschine auch Sachverhalte, die dem Menschen verborgen bleiben. Darüber hinaus kann KI zur Verarbeitung von Informationen eingesetzt werden und übernimmt dann die automatisierte Fallbearbeitung. Nicht nur die internen Bearbeitungsschritte können KI überlassen werden, ebenso können Kommunikationsaufgaben automatisiert ausgeführt werden, sodass Interaktion mit Menschen möglich wird.

Die Vorhersage von Technologieeinsätzen bleibt natürlich stets spekulativ – so auch an dieser Stelle in Bezug auf den Einsatz von KI. Stets ist im konkreten Fall von den spezifischen Voraussetzungen und Rahmenbedingungen auszugehen und KI sodann passgenau zu implementieren. Die im Folgenden vorgenommene Auflistung potentieller Anwendungsfälle künstlicher Intelligenz ist daher nicht als zuverlässige Prognose der Entwicklung des Sozialversicherungsbereichs im Hinblick auf künstliche Intelligenz zu verstehen, auch ist mit einer Nennung nichts über die Wünschbarkeit einer spezifischen Anwendung ausgesagt. Zudem ist mit dem Auf-

zeigen möglicher Einsatzbereiche für künstliche Intelligenz keine Bewertung dahingehend verknüpft, inwieweit der skizzierte Anwendungsfall im Hinblick auf rechtliche Rahmenbedingungen umsetzbar ist. Hierzu sind jeweils Einzelfallprüfungen im Hinblick auf die jeweilig geltenden Rechtsgrundlagen für die spezifische Situation erforderlich. Vielmehr geht es im Folgenden darum, den Möglichkeitsraum abzustecken und aus der Vielzahl möglicher Anwendungsbereiche von KI diejenigen zu skizzieren, die für das Sozialversicherungswesen sachlich geeignet erscheinen. Nicht zuletzt geht es damit auch darum, einen Beitrag zur Debatte rund um KI zu leisten. So jung diese Debatte – insbesondere auf dem Feld des Sozialversicherungswesens – noch ist, so dringlich ist sie im Angesicht einer sich mit immensem Tempo weiterentwickelnden Technologie, die das Potential hat, tiefgreifende Umwälzungen zu bewirken.

Im Folgenden sollen beispielhaft Anwendungsfälle für KI-Nutzung entlang der Prozesskette von Sozialversicherungsorganisationen aufgeführt werden. Dabei soll an dieser Stelle vorausgeschickt werden, dass sich eine nicht geringzuschätzende Einschränkung für die Nutzung von KI im Bereich der Sozialversicherung schon allein daraus ergibt, dass in diesem Bereich auf wesentliche Elemente von Privatversicherungen, insbesondere auf eine am individuellen Risiko orientierte Festsetzung von Prämien oder einen strikt versicherungsmathematischen Zusammenhang zwischen Prämien und Leistungen verzichtet wird. Durch die Geltung der Individualäquivalenz in der privaten Versicherungswirtschaft, d. h. einer am individuellen Erwartungsschaden orientierten Prämienberechnung, eröffnet sich beispielsweise ein bedeutendes Einsatzfeld für KI-basierte Prognosen.

INFORMATION & KOMMUNIKATION

Die Darbietung von Informationen zu den jeweiligen Versicherungszweigen, deren Leistungsangebot und Anspruchsvoraussetzungen sollte aus Gründen der Ressourceneffizienz so zielgenau wie möglich erfolgen. Kundenkommunikation und Werbung im Wirtschaftsbereich haben von KI-Einsatz in den letzten Jahren dadurch profitiert, dass Kommunikationsmaßnahmen personalisierter und zielgenauer gestaltet werden können, indem so grundlegende Funktionen wie die Identifikation von Adressateninformationen, die Medienplanung, der Einkauf und die Erstellung von Werbung sowie die Bewertung ihrer Wirksamkeit automatisiert wurden. (vgl. Ford et al., 2023) So stellt etwa Targeting, also das möglichst präzise Ausrichten einer Botschaft auf die Zielgruppe, einen entscheidenden Anwendungsbereich von KI im Werbesektor dar. KI unterstützt dabei, effektiv die richtige Zielgruppe mit den richtigen Informationen zum richtigen Zeitpunkt zu erreichen. Eine Analyse von relevanten Daten führt hierbei zur Identifikation verschiedener Adressatengruppen und deren spezifischen Bedürfnissen und Präferenzen. Eine adressatengerechte Ansprache könnte sich bestimmt auch in der Kommunikation bezüglich Sozialleistungen vorteilhaft auswirken, bedenkt man beispielsweise, dass suboptimale Alterssicherung immer wieder auf Informations- und Motivationsdefizite zurückgeführt wird.

Hand in Hand mit dem Targeting geht die Personalisierung von Kommunikationsbotschaften. Basierend auf Erkenntnissen durch die Analyse von demographischen Informationen, Gewohnheiten und Präferenzen, um potentiell interessierte Personenkreise zu identifizieren, werden einzelnen Personen jeweils die für sie relevantesten Inhalte offeriert. In diesem Zusammenhang können beispielsweise KI-basierte Empfehlungssysteme oder virtuelle Agenten zum Einsatz kommen, die Inhalte zur Verfügung stellen, welche auf die individuellen Interessen der Adressaten zugeschnitten sind oder Fragen beantworten. So ist etwa denkbar, Informationen nach dem Prinzip bereitzustellen, wie es im Onlinehandel bereits gang und gäbe ist: „Kunden, die X gekauft haben, kauften auch Y.“ Oder aber Informationen werden in Anknüpfung an vorangegangene Informationsbedürfnisse und Verhaltensweisen verfügbar gemacht.

Sollen in die Kommunikation mit Interessenten nicht nur eigene, sondern auch externe Informationsquellen einbezogen werden, so kann KI bei der Kuratierung von Inhalten unterstützen. Durch das Angebot lediglich solcher Informationen, die im konkreten Fall für einen bestimmten Nutzer gerade relevant sind, lässt sich nicht nur die Effektivität von Informations- und Kommunikationsmaßnahmen steigern, ebenso ist von einer erhöhten Zufriedenheit der Informationssuchenden auszugehen. (vgl. Gao et al., 2023, 9f.)

In Bezug auf eingehende Anfragen können diese KI-basiert validiert, kategorisiert und geeigneten weiteren Bearbeitungsschritten zugeordnet werden. Darüber hinaus erlaubt KI eine kontinuierlich effektiver werdende Kommunikation mit Hilfe von NLP-Technologie, die es Chatbots erlaubt, natürlichsprachliche Dialoge zu simulieren, in deren Zuge die Intentionen und Bedarfe der Nutzer analysiert und daraufhin angemessen geantwortet wird. Damit wären Sozialversicherungsorganisationen imstande, nicht nur generelle Informationen bereitzustellen, sondern automatisiert auf individuelle Informationsbelange einzugehen und rund um die Uhr qualitativ hochwertige Dienstleistungen bereitzustellen. Es steht zu erwarten, dass Chatbots einen immer wichtigeren Baustein der Nutzerkommunikation von Institutionen der sozialen Sicherung ausmachen werden. Dabei kann man davon ausgehen, dass die KI-basierten Dialogsysteme die herkömmlichen digitalen und persönlichen Kanäle ergänzen werden: Während einige herkömmliche Kommunikationskanäle durch sie ersetzt werden können, werden sie andere verbessern und wiederum andere werden bestehen bleiben, weil der Einsatz von Chatbots etwa aus rechtlichen Gründen nicht umfassend möglich sein wird. (vgl. IVSS, 2022b) Die Personalisierung von Nachrichten wird dabei mit Hilfe von generativer KI auf die Spitze getrieben, indem neuer, origineller Inhalt in Echtzeit erschaffen wird. Auf Basis von individuellen Adressatendaten und -präferenzen ist es möglich, einzelnen Personen personalisierte und ansprechende Inhalte in diversen Medienformaten (z. B. Text, Bild, Video) zuzuspielen. Mit Hilfe einer Sentimentanalyse können Nachrichten zudem an die Stimmungslage der betreffenden Personen angepasst werden.

Darüber hinaus ist ein KI-Einsatz denkbar, der Informationen nach dem „push-Prinzip“ an interessierte Personen versendet. Als „Contextual Targeting“ (vgl. Gao et al., 2023, S. 9) wird in der Marketingsprache das Anzeigen von Onlinewerbung in einem inhaltlichen Umfeld bezeichnet, das dafür optimal geeignet ist. Eine solche Form der Kommunikation folgt der Grundidee, dass die Ansprache von potentiellen Interessenten auf Webseiten stattfindet, die in erster Linie von Personen frequentiert werden, die sich für die entsprechenden Angebote grundsätzlich interessieren. So werden beispielsweise aller Voraussicht nach Informationen zur Altersabsicherung auf Portalen für Finanzinformationen auf vergleichsweise hohes Interesse stoßen. Mit Predictive Customer Service (vgl. Nair & Gupta, 2021, S. 325) lässt sich die prädiktive Kraft von KI im Bereich des Kundenservices nutzen, um zukünftige Bedarfe und Präferenzen von Kunden zu antizipieren. Dies ebnet Sozialversicherungsorganisationen den Weg, über die Analyse von beispielsweise Kundendaten, bisherigen Anfragen und sonstigen Interaktionen sowie dem Browsing-Verhalten Service zu bieten, bevor überhaupt eine Kontaktaufnahme seitens der Kunden stattgefunden hat.

Zwecks Unterstützung von Versicherten sind virtuelle Assistenten vorstellbar, die während der verschiedenen Lebenslagen durch die unterschiedlichen Leistungen des Sozialwesens leiten und mit Rat und Tat zur Seite stehen. Sie können etwa bei der Beantragung von Rente oder Pflegegeld behilflich sein, bei der Recherche von Job- und Ausbildungsmöglichkeiten helfen oder Fragen bezüglich der Altersvorsorge beantworten. Im Idealfall geben solche KI-Agenten nicht nur Auskunft, sondern helfen bei der Erledigung der konkreten Anliegen: Beispielsweise könnte KI in Antragsverfahren unterstützen, indem beim Auffinden der notwendigen Formulare unterstützt und sogleich auch das Antragsverfahren vereinfacht wird. Es ist denkbar, Daten aus der natürlichen Sprache direkt zu übernehmen und nach einer Prüfung auf Plausibilität und Richtigkeit in das Formular einzutragen. Sollte dies erforderlich sein, kann der KI-Agent auch gezielte Nachfragen stellen.

DATENERFASSUNG

Der tagtäglich zu verzeichnende Dokumenteneingang bei Sozialversicherungsorganisationen (z. B. Briefe, Uploads über die Webseite, Anträge) muss bearbeitet werden, d. h. Dokumente unterschiedlicher Formate müssen eingelesen und validiert werden; zudem muss der Inhalt jedes einzelnen Dokuments erkannt werden, um die Art des Dokuments zu erfassen und diese dann in den richtigen fachlichen Prozess einzuleiten. Diese Kategorisierungsaufgabe ist mit Hilfe von KI voll- oder teilautomatisiert zu erledigen, sodass Aufgaben insgesamt schneller erledigt werden.

Über eine gute Datenqualität zu verfügen, stellt nicht nur die Grundlage für die Durchführung der verschiedenen Aufgaben der Sozialversicherungsorganisationen dar, zugleich ist die Sicherung der Datenqualität entscheidend für den Erfolg von KI-Projekten. Probleme im Hinblick auf die Datenqualität können auf unterschiedliche Gründe zurückzuführen sein. So können Daten etwa aufgrund menschlicher Eingabefehler fehlerhaft sein oder aber Daten fehlen, wodurch die Entwicklung

robuster und aussagekräftiger KI-Modelle beeinträchtigt sein kann. Auch das Vorhandensein zu vieler Daten, also beispielsweise irrelevanter und redundanter Daten, kann sich nachteilig auf die Leistungsfähigkeit und Zuverlässigkeit von KI-Systemen auswirken. Vor diesem Hintergrund ist von Bedeutung, dass Anwendungen künstlicher Intelligenz nicht nur von einem qualitativ hochwertigen Datenbestand profitieren, sondern selbst eingesetzt werden können, um aktiv zu einer Erhöhung der Datenqualität beizutragen, indem sie Ungenauigkeiten, Lücken, Duplikate und Inkonsistenzen in den Daten automatisch identifizieren und korrigieren.

Mit Hilfe von KI-basierten Anwendungen können „Ausreißer“, also ungewöhnliche Datenpunkte, die deutlich von der Norm abweichen, identifiziert werden (vgl. Hassan et al., 2022; Sikder & Batarseh, 2023). Die sogenannte „Outlier Detection“ dient dazu, in der weiteren Folge solche Ausnahmefälle einer näheren Analyse zu unterziehen sowie Bereiche zu erkennen, wo derartige Ereignisse wiederholt geschehen. Beispielsweise könnte ein solches System einen Ausreißer hinsichtlich einer Person identifizieren, die ein sehr hohes Einkommen, aber keinerlei Vermögensgüter aufweist. Durch nachfolgende Prüfung wäre in einem solchen Fall herauszufinden, ob es sich um einen Fehler oder schlicht eine Abweichung von der Norm handelt.

Weil Algorithmen maschinellen Lernens auf vollständige Datensätze angewiesen sind, ist die Bereinigung im Hinblick auf fehlende Daten von großer Bedeutung. Auch im Bereich der automatischen Vervollständigung des Datenbestands kann KI zum Einsatz kommen. Bertsimas et al. (2018) zeigen, dass eine KI-basierte prädiktive Methode zur Vervollständigung fehlender Daten Vorteile gegenüber anderen Methoden (z. B. Mittelwert) aufweist.

ANSPRUCHSPRÜFUNG

Zu den offensichtlichsten Anwendungsmöglichkeiten von KI im Bereich der Sozialversicherung gehören sicherlich die Prüfung von und die Entscheidung über die Anspruchsberechtigung sowie das Leistungsniveau. Insbesondere Methoden des überwachten Lernens⁷ eignen sich für die Erledigung dieser Aufgaben. Hierbei werden Entscheidungsmuster aus Trainingsdaten erlernt, die entweder dem manuellen Vorgängersystem entstammen können oder speziell für das Training erstellt werden. Es besteht auch die Möglichkeit, die genannten Aufgaben nicht vollständig an ein KI-System abzutreten, sondern dieses in Ergänzung zur menschlichen Aufgabendurchführung einzusetzen. Dabei kann mit Hilfe einer Modellunsicherheitsanalyse die Zuverlässigkeit der bereitgestellten Ergebnisse des KI-Modells bewertet werden, um anschließend eine Aufgabenteilung vorzunehmen: Offensichtliche Fälle werden von der KI bearbeitet und weniger eindeutige Entscheidungen werden an einen menschlichen Mitarbeiter zur Überprüfung weitergereicht. Um auf diesem Weg langfristig zu guten Entscheidungen zu kommen, ist es rat-

7 Supervised Learning (überwachtes Lernen) beschreibt den Prozess des maschinellen Lernens, bei dem ein Algorithmus aus einem Datensatz mit bekannten Ergebnissen lernt. Hierbei wird ein Modell erstellt, indem dem System Input und gewünschter Output mitgeteilt werden. Als Algorithmen kommen beispielsweise Lineare Regression oder Entscheidungsbäume in Betracht. Beim Unsupervised Learning (unüberwachtes Lernen) hingegen wird Deep Learning genutzt, um anhand von unmarkierten Trainingsdaten zu Schlussfolgerungen und Mustererkennung zu gelangen.

sam, die Unsicherheitseinschätzungen regelmäßig zu überprüfen, um sicherzustellen, dass die Vorauswahl erwartungsgemäß funktioniert.

Zudem lässt sich die Last der Fallbearbeitung durch Expertensysteme und virtuelle Agenten lindern. Solche Hilfestellungen lassen sich beispielsweise einsetzen, um Mitarbeitern dabei zu helfen, mit den zahlreichen und komplexen Gesetzen und Vorschriften zurechtzukommen sowie bei Rechercheaufgaben zu unterstützen, indem relevante Informationen und Quellen gesucht und Inhalte zusammengefasst werden. Auch bei Übersetzungsaufgaben kann KI wertvolle Dienste leisten.

Eine weitere Einsatzmöglichkeit ist im Bereich der Bewertung von Arbeits- und Erwerbsfähigkeit als Eckpfeiler von Sozialschutzsystemen gegeben – hier könnte KI neue datenbasierte Wege weisen: Für die Beurteilung, ob und in welchem Ausmaß eine Person sich auf dem Arbeitsmarkt engagieren kann, könnte künstliche Intelligenz für Ärzte ein unterstützendes Instrument sein, das neue Einsichten beisteuert, die im Ergebnis zu treffsichereren Entscheidungen sowie gezielteren Integrationsmaßnahmen führen.

Zudem stehen Sozialversicherungsorganisationen vor der Herausforderung, Bewertungen hinsichtlich der Anspruchsberechtigung kontinuierlich durchzuführen, damit diese die jeweils aktuellen Umstände reflektieren. Auch in diesem Zusammenhang kann KI eingesetzt werden, um zu gewährleisten, dass das System der sozialen Sicherung möglichst flexibel auf veränderte Umstände reagiert. Legt die Aktualisierung von personenbezogenen Daten eine Neuüberprüfung des Sozialversicherungsanspruchs nahe, so könnte eine solche automatisch angestoßen und eine aktualisierte Bewertung durchgeführt werden.

Die Fähigkeit der verschiedenen Zweige der Sozialversicherung, auf veränderte Gegebenheiten angemessen zu reagieren, kann auch durch Zeitreihenprognosen im Hinblick auf die Inanspruchnahme von Sozialversicherungsleistungen sowie des Beitragsaufkommens erhöht werden. Zu diesem Zweck können KI-basierte Prognoseverfahren eingesetzt werden, die beispielsweise durch Verknüpfung von sequentiellen Daten zu Anspruchsniveaus mit sozioökonomischen Trends und anderen zukunftsgerichteten Informationen zu Vorhersagen gelangen, die es Sozialversicherungsorganisationen erlauben, stärker proaktiv anstatt reaktiv Anpassungen vorzunehmen. So erforscht beispielsweise das vom Innovationsausschuss beim Gemeinsamen Bundesausschuss geförderte Projekt KI-THRUST (Laufzeit von 07/2021 bis 06/2024) die Potenziale von KI-gestützten Vorhersageverfahren, durch Auswertungen von Routinedaten der gesetzlichen Krankenversicherungen (z. B. Informationen aus der ambulanten medizinischen Versorgung und Aufhalten in Krankenhäusern und Rehabilitationseinrichtungen) Versorgungsbedarfe früher und besser zu erkennen.

ANTRAGSMANAGEMENT

Ein wichtiger Anwendungsbereich mit einem großen Potenzial für Effizienzgewinne liegt im Antragsmanagement. Sozialversicherungsorganisationen erhalten eine hohe Anzahl von Anträgen auf Erbringung bzw. Auszahlung der jeweiligen Sozialversicherungsleistungen, deren Abarbeitung schnelle und präzise Entscheidungen erfordert. Unzweifelhaft kann eine Automatisierung mit Hilfe von KI in diesem Bereich zu erheblichen Zeit- und Kostenersparnissen beitragen. Um treffsichere Entscheidungen sowie Effizienzgewinne zu realisieren, bietet sich auch in diesem Bereich der vorhin beschriebene Ansatz an, Fälle einzuteilen in einerseits solche, die eindeutig und daher rein maschinell zu bearbeiten sind, und andererseits solche, die weniger eindeutig sind und daher einer menschlichen Überprüfung zugänglich gemacht werden sollten.

So ist etwa die Abrechnungsprüfung als der vom deutschen Gesetzgeber vorgehene Möglichkeit für Krankenkassen, Abrechnungen der Krankenhäuser dahingehend zu prüfen, ob diese nur medizinisch notwendige Leistungen erbringen und diese nachvollziehbar dokumentieren, ein in der Praxis komplexer und aufwendiger Prozess. Dabei zeigt die Abrechnungsqualität seit Jahren ein ernüchterndes Bild: Rund 50 Prozent der Abrechnungen sind inkorrekt. (vgl. GKV-Spitzenverband, 2024) Hehner et al. (2017) weisen darauf hin, dass die deutschen Krankenkassen aufgrund der Fehlerhaftigkeit der Abrechnung von Leistungen mit aufwendigen und kostenintensiven Prüfungen befasst sind. Der Einsatz einer KI, die aus vergangenen Fällen lernt und sich kontinuierlich verbessert, verspricht in diesem Zusammenhang eine effizientere Überprüfung von Fällen. Die gegenwärtige Praxis zeigt, dass ein großer Teil der Anträge auf Kostenerstattung irgendwie auffällig, d. h. potentiell fehlerhaft ist, woraufhin eine detaillierte Prüfung und gegebenenfalls eine Intervention erfolgt. Solche Einsprüche führen den Autoren zufolge nur in zehn Prozent aller auffälligen Fälle zum Erfolg. Mit Hilfe von KI kann es gelingen, so die Autoren, aus der großen Zahl der Anträge auf Kostenerstattung, zuverlässig solche Fälle zu identifizieren, die tatsächlich fehlerhaft sind. Smarte Algorithmen würden in einem solchen Szenario auf das Aufspüren von Fällen mit einer hohen Wahrscheinlichkeit eines erfolgreichen Einspruchs fokussieren. Umgekehrt würden unbedenkliche Fälle und solche, bei denen eine erfolgreiche Intervention unwahrscheinlich ist, einer vollautomatischen Hintergrundverarbeitung zugeführt werden, sodass menschliche Kapazitäten auf jene Fälle konzentriert werden können, die eine Überprüfung erfordern.

KI-Anwendungen können darüber hinaus im Bereich der Betrugserkennung eingesetzt werden, um diejenigen Ansprüche herauszufiltern, die in betrügerischer Absicht beantragt werden und die Versichertengemeinschaft unnötig Geld kosten. (vgl. IVSS, 2022a; Johnson & Khoshgoftaar, 2023) Als ultimatives Ziel eines KI-Einsatzes definieren Farbmacher et al. (2022) die Entwicklung eines Vorhersagemodells, das betrügerische Ansprüche ausfiltert und nicht betrügerische Ansprüche automatisch auszahlen kann. Aufgrund des Ungleichgewichts zwischen Betrugs- und Nichtbetrugsfällen hat es die Entwicklung von KI-Modellen in diesem

Bereich allerdings mit besonderen Herausforderungen zu tun. Da nur ein geringer Anteil aller Fälle betrügerischer Natur ist, wovon wiederum nur ein kleiner Anteil als Betrug erkannt wurde, sind Methoden des „Supervised Learning“ in diesem Bereich nur eingeschränkt nutzbar, weil Daten nicht in jedem Fall als Betrugsfall oder Nichtbetrugsfall markiert sind. Zur Aufdeckung von Betrugsfällen kommen daher Methoden des „Unsupervised Learning“ zur Anwendung, die nach ungewöhnlichen, d. h. in bestimmten Dimensionen stark abweichenden, Fällen suchen, um diese durch menschliche Mitarbeiter überprüfen zu lassen. Nach erfolgter Prüfung kann die Markierung des betreffenden Falls dann gegebenenfalls von „Nichtbetrug“ auf „Betrug“ abgeändert werden. (vgl. Spindler & Kögel, 2020, S. 5) Dennoch gleicht die KI-basierte Aufdeckung von Betrugsfällen dem Wettlauf zwischen Hase und Igel. Da die Betrugstaktiken kontinuierlich angepasst werden, sind extrem anpassungsfähige Systeme vonnöten, um die jeweils aktuellen Betrugsmuster zu erkennen.

LEISTUNGSERBRINGUNG

Die Vorhersage von Ereignissen, die zu einer wahrscheinlichen Inanspruchnahme der Sozialversicherung führen, ist ein vielversprechender Anwendungsfall für KI, weil sich durch solche Prognosen die Möglichkeit auftut, durch frühzeitige Maßnahmen den Schaden von vornherein zu minimieren. Vorhersagen können sich auf ökonomische Entwicklungen, Naturkatastrophen, Verwerfungen auf dem Arbeitsmarkt etc. beziehen, um die Folgen der Prognosen in den einzelnen Sozialversicherungszweigen zu antizipieren und entsprechende Vorkehrungen zu treffen.

Insbesondere der Krankenversicherungszweig kann von Vorhersagen zum Zwecke der Prävention profitieren. Mit Hilfe von künstlicher Intelligenz ist vorstellbar, Personen oder Personengruppen zu identifizieren, die mit großer Wahrscheinlichkeit Sozialversicherungsleistungen beziehen werden. Auf Basis dieser Erkenntnisse ist es dann möglich, Maßnahmen zu treffen, um die zur Anspruchsberechtigung führenden Fälle von vornherein zu vermeiden. So kann etwa auf Basis von Krankenversicherungsdaten, demographischen Informationen, Daten von Wearables etc. die zukünftige Gesundheit einer Person vorausgesagt werden. (vgl. Badidi, 2023) Beispielsweise kann KI eingesetzt werden, um chronische Krankheiten frühzeitig zu erkennen. (vgl. Patel et al., 2022) Rashid et al. (2022) schlagen eine KI-basierte Vorgehensweise zur Vorhersage von weit verbreiteten chronischen Erkrankungen wie Brustkrebs, Diabetes, Herzinfarkt, Hepatitis und Nierenerkrankungen auf Basis verfügbarer Patientendaten vor, um frühzeitig auf die Veränderung von Lebensgewohnheiten hinzuwirken sowie Behandlungsmaßnahmen zu beginnen. Die frühzeitige Erkennung solcher Krankheiten führt nicht nur zu einer Senkung der Sterblichkeitsrate, sondern entlastet auch das Gesundheitssystem.

Für den Bereich der Arbeitslosenversicherung sind KI-basierte Vorhersagen über die Vermittlungswahrscheinlichkeit Arbeitssuchender möglich. So zeigen etwa van den Berg et al. (2024), dass Systeme auf Basis maschineller Lernverfahren zu akkuraten Ergebnissen im Hinblick auf die Einschätzung der Wahrscheinlichkeit einer Wiederbeschäftigung innerhalb von sechs Monaten nach Eintritt in die Arbeits-

losigkeit gelangen. Es konnte empirisch gezeigt werden, dass die KI-basierten Vorhersagen sowohl die selbstwahrgenommene Wiederbeschäftigungswahrscheinlichkeit durch die Arbeitslosen als auch Einschätzungen durch Vermittlungsfachkräfte übertrafen.

Zur Entlastung der Arbeitslosenversicherung kann in den Arbeitsagenturen KI-basiertes Job-Matching zur Anwendung kommen, um die Dauer der Arbeitslosigkeit zu verkürzen, indem Arbeitsuchende schneller vermittelt werden. Als Vorbild dienen hierbei KI-basierte Anwendungen, wie sie im betrieblichen Recruitingprozess immer weitere Verbreitung finden, um für Unternehmen geeignete Mitarbeiter zu identifizieren, indem durch Algorithmen die Passgenauigkeit zwischen Jobsuchenden und Jobanbietern optimiert wird. Mit Hilfe von Algorithmen lassen sich automatisiert Jobbeschreibungen und Bewerbungsunterlagen abgleichen und ideale Kandidaten ausfindig machen. (vgl. Rezaeipourfarsangi & Milios, 2023; Rojas-Galeano et al., 2022)

Als Instrumente zur Realisierung einer zielgerichteten Ressourcennutzung und Personalisierung von Leistungen könnte KI außerdem zur Klassifizierung und Profilbildung eingesetzt werden. Hierbei werden (potentiell) Begünstigte und Situationen klassifiziert, kategorisiert und bewertet – oftmals unter Risikogesichtspunkten. Henman (2022, S. 270) weist auf die mit dieser Herangehensweise erheblichen rechtlichen, politischen, ethischen und sozialen Herausforderungen hin. So wird Profiling- und Klassifizierungsalgorithmen immer wieder vorgeworfen, soziale Vorurteile und Diskriminierung zu reproduzieren und zu verschärfen.

Auch im Zuge der Abrechnung der Leistungen kann KI zunehmend eine Rolle spielen. Hausmann & Lämmel (2021) berichtet etwa von Leistungsabrechnungen der Versicherten in den privaten Krankenversicherungen, die sich gut für eine automatische Verarbeitung – vom Scannen eingehender Papierdokumente über deren Weiterverarbeitung bis hin zur Zahlungsanweisung – eignen. Dies setzt voraus, die in der Gebührenordnung für Ärzte (GOÄ) für jede ärztliche Leistung festgelegten Schlüssel, Kurzbezeichnungen als auch Geldbeträge sicher zu erkennen und zu „verstehen“.

MONITORING & QUALITÄTSMANAGEMENT

Mit der Überwachung und Bearbeitung von Beschwerden tut sich ein weites Einsatzfeld für KI auf. Die Bedürfnisse, Erfahrungen und die Zufriedenheit der (potentiellen) Begünstigten kann im Rahmen eines Beschwerdemanagements laufend ermittelt und analysiert werden. KI unterstützt auf diesem Weg eine hohe Dienstleistungsqualität und faire Bereitstellung der Sozialversicherungsleistungen. Im Hinblick auf das Beschwerdemanagement sind zwei Dimensionen zu unterscheiden, in denen KI auf jeweils unterschiedliche Weise zum Einsatz kommen kann: Zum einen geht es im Rahmen der Außenbeziehungen zu Versicherten um die Annahme, Bearbeitung und Reaktion auf Beschwerden mit dem Ziel, individuelle Anliegen zu prüfen und zu lösen. Zum anderen geht es außerdem darum, die aus dem Beschwerdemanagementprozess stammenden Daten zu erheben und zu analysieren, um daraus Hinweise auf Planungs-, Prozess- und Leistungsmängel ab-

zuleiten. Die Analyse des eingehenden Feedbacks dient hierbei als Grundlage, um beispielsweise Abweichungen von Normen zu erkennen und abzustellen oder Maßnahmen zur Qualitätsverbesserung einzuleiten.

Der Einsatz von KI in der Beschwerdebearbeitung ermöglicht eine schnelle und effiziente Verarbeitung einer großen Zahl von Reklamationen und Anliegen. Weil KI-Anwendungen innerhalb kürzester Zeit eine Vielzahl von Beschwerden analysieren und passende Lösungen oder Antworten generieren können, verkürzen sich die Bearbeitungszeiten und es erfolgt eine schnellere Reaktion auf Anfragen. Selbst die Generierung von kontextbezogenen Antworten, welche mit Hilfe von Sentimenterkennung die jeweilige Stimmung des Anfragenden in der Antwort berücksichtigt, ist möglich. Auch wenn der KI-Einsatz in diesem Bereich die menschliche Interaktion nicht vollständig zu ersetzen vermag, so erscheinen hybride Lösungen ideal: Während KI-Modelle wiederkehrend auftretende Fragen und Probleme identifizieren und automatisch in geeigneter Weise beantworten bzw. lösen können, werden hierdurch die Kapazitäten menschlicher Mitarbeiter geschont, um sich auf komplexere Anfragen konzentrieren zu können. KI-Systeme bereiten dies effizient vor, indem sie die große Menge strukturierter und unstrukturierter Informationen kategorisieren und sie dann schnell an die geeigneten Servicemitarbeiter weiterleiten. (vgl. Forster & Entrup, 2017)

Gerade im Gesundheitsbereich besteht die Möglichkeit, Patientenbeschwerden als Indikator für mögliche Qualitätsmängel in der Gesundheitsversorgung zu nutzen. Beschwerden von Patienten können Probleme in der Gesundheitsversorgung aufdecken, die durch andere Sicherheits- und Qualitätsüberwachungssysteme nicht erfasst werden. Von besonderer Relevanz sind die von Patienten und deren Angehörigen vorgebrachten Beschwerden schon allein deshalb, weil diese Personen über einen privilegierten Zugang zu Informationen über die Gesundheitsversorgung verfügen. Ständen bisher keine ausreichenden Möglichkeiten zur Verfügung, solche unstrukturierten Datenmengen einer Analyse zuzuführen, so wäre dieses Problem mit geeigneten KI-Anwendungen zu umgehen. Die hierdurch gewonnenen Erkenntnisse können genutzt werden, um Probleme und blinde Flecken in der Gesundheitsversorgung aufzudecken. (vgl. Birkeland et al., 2024)

Auch KI-basierte Chatbots können im Bereich des Beschwerdemanagements eingesetzt werden, um Effizienzgewinne durch die automatische Aufnahme und Bearbeitung von Beschwerden herbeizuführen. Den Vorteilen einer Rund-um-die-Uhr-Verfügbarkeit, einer Prozessstandardisierung sowie einer mehrsprachigen Unterstützung stehen die – gerade im Bereich des Beschwerdewesens nicht zu vernachlässigenden – Nachteile gegenüber, beim Beschwerdeführer Frustration hervorzurufen, weil Chatbots nicht ausreichend in der Lage sind, Nuancen zu interpretieren und ohne menschliches Zutun eine zufriedenstellende Lösung zu bewirken. Auch in diesem Bereich lässt sich bestimmt am ehesten mit hybriden Ansätzen die optimale Balance von Nutzen und Risiken finden. (vgl. Ismailbekovich, 2024)

Die Aufdeckung von Betrugsfällen ist ein wichtiger Bereich des Monitorings. KI kann zu diesem Zweck nicht nur im Zuge der Antragsprüfung eingesetzt werden, sondern spielt ebenso über die gesamte Phase der Leistungserbringung eine Rolle, um Täuschungen und bewusst unwahre Behauptungen der Versicherten zu identifizieren, die auf einen unrechtmäßigen Erhalt von Leistungen zielen. KI kann dabei unterstützen, ungewöhnliche Auszahlungsmuster zu erkennen, die auf Manipulationen hindeuten. So sieht etwa die Deutsche Sozialversicherung Arbeitsgemeinschaft Europa e.V.⁸ in „der Ermittlung von Betrugsmustern, der Verfolgung von systematischem Missbrauch sozialer Versorgungsleistungen oder der Aufdeckung von organisierten kriminellen Fehlverhaltensstrukturen“ (Deutsche Sozialversicherung, 2020, S. 4) ein beispielhaftes Einsatzfeld für künstliche Intelligenz im Bereich Sozialversicherung.

Des Weiteren kann KI nicht nur zur Entdeckung kurzfristiger Anomalien, sondern auch zur Identifizierung langfristiger Veränderungen eingesetzt werden. So kann KI etwa im Rahmen von Bedarfsprognosen eine bedeutende Rolle spielen, um beispielsweise durch Verknüpfung von Zeitreihendaten mit externen sozioökonomischen Trends Vorhersagen zum Beitragsaufkommen oder Versorgungsbedarf zu erstellen. Auf diese Weise können administrative Anpassungen an veränderte Leistungs- oder Servicevolumina rechtzeitig vorgenommen werden.

Abgesehen von den oben dargestellten Einsatzmöglichkeiten künstlicher Intelligenz im Rahmen der Leistungserstellungsprozesse der Sozialversicherungsorganisationen sind auch Anwendungen auf einer übergeordneten Ebene denkbar: Eine KI-basierte Datenanalyse ist imstande, eine bessere Informationsbasis für politische Entscheidungen betreffend die Weiterentwicklung der grundsätzlichen Herangehensweisen an die soziale Sicherung und deren Leistungen zu schaffen. Vorhersagen und Einsichten im Hinblick auf die Auswirkungen, die Wirksamkeit und die besten Praktiken des Sozialschutzes können dazu beitragen, die Grundlage für evidenzbasierte Entscheidungen zu festigen. Ebenso spielen in diesem Zusammenhang Vorhersagen eine große Rolle sowie darauf basierende Simulationen für die langfristige Planung von politischen Maßnahmen.

Zusammenfassend lassen sich zwei grundsätzliche Einsatzmöglichkeiten von KI im Rahmen des Leistungserstellungsprozesses der Sozialverwaltung unterscheiden: Künstliche Intelligenz kann entweder im Front-End-Bereich unterstützen, um die Interaktion mit den Versicherten zu optimieren oder sie wird im Back-End-Bereich zur Automatisierung der Sachbearbeitung im Rahmen der internen Prozesse eingesetzt.

⁸ Die Deutsche Rentenversicherung Bund (DRV Bund), die Deutsche Gesetzliche Unfallversicherung (DGUV), der GKV-Spitzenverband und die Verbände der gesetzlichen Kranken- und Pflegekassen auf Bundesebene haben sich zwecks Vertretung ihrer gemeinsamen europapolitischen Interessen zur „Deutschen Sozialversicherung Arbeitsgemeinschaft Europa e.V.“ zusammengeschlossen.

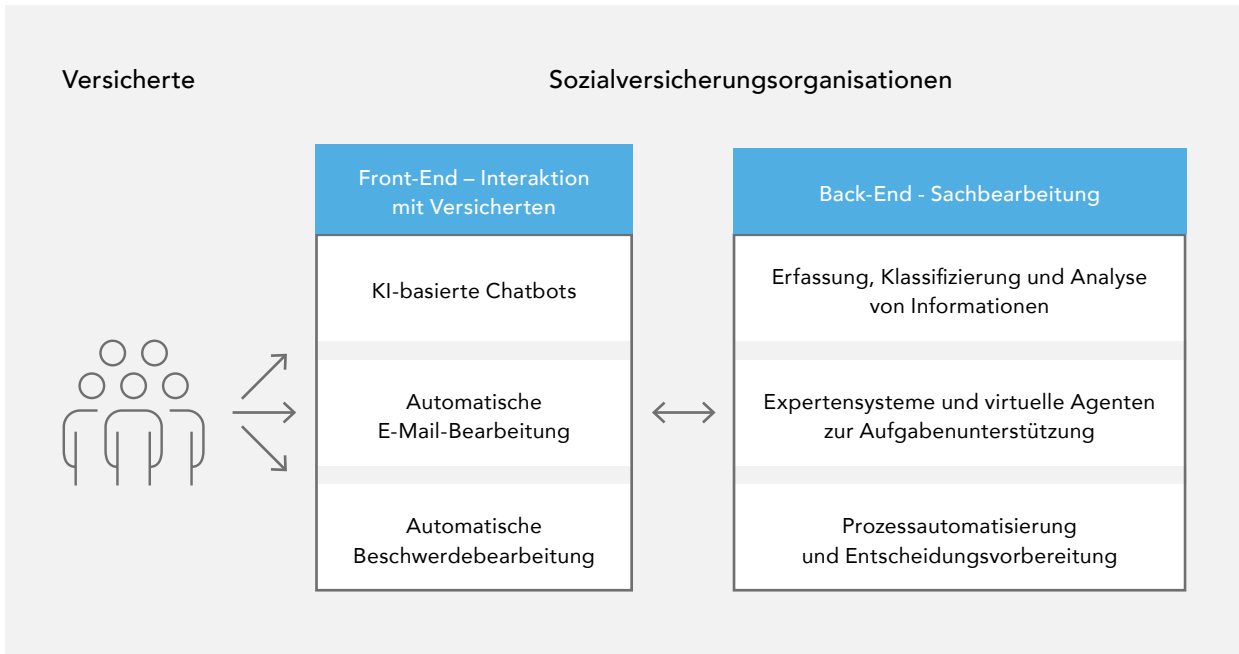


Abbildung 10: Einsatzmöglichkeiten von KI in der sozialen Sicherung; Quelle: Eigene Darstellung

3.1.3 Einsatz von KI im Gesundheitssektor

Weil der Einsatz von künstlicher Intelligenz im Gesundheitswesen Auswirkungen auf die Krankenkassen erwarten lässt, soll im Folgenden ein Schlaglicht auf KI-Anwendungen im Bereich der medizinischen Leistungserbringung geworfen werden. Durch eine gegebenenfalls verbesserte Gesundheitsversorgung sowie die Schaffung von optimierten Vorsorgemöglichkeiten durch die prognostischen Fähigkeiten von KI sind erhebliche Potenziale gegeben, das System der Krankenversicherung finanziell zu entlasten. Einen Blick auf die Leistungserbringung im Gesundheitssektor zu werfen, ist auch deshalb lohnend, weil es in zahlreichen Feldern des Gesundheitswesens bereits KI-Anwendungen gibt, die sich entweder noch in der Entwicklung befinden oder als Prototypen existieren oder bereits eingesetzt werden.

Zugleich ist gerade im Gesundheitsbereich unmittelbar einleuchtend, dass die rechtlichen und ethischen Fragen im Zusammenhang mit dem Einsatz künstlicher Intelligenz nicht ignoriert werden können. Dabei eignet sich der Gesundheitsbereich auch, die Dilemmata aufzuzeigen, mit denen sich eine Gesellschaft durch die KI-Technologie konfrontiert sieht. Kann KI auf der einen Seite einen bislang nie dagewesenen Vorteil im Schutz von Gesundheit und Kampf gegen Krankheiten erbringen, so sind gleichzeitig mögliche negative Konsequenzen der Datennutzung, wie sie durch das Schlagwort vom „gläsernen Patienten“ auf den Begriff gebracht werden, nicht außer Acht zu lassen.

Im Folgenden sollen einige KI-Anwendungen, wie sie im Gesundheitsbereich bereits existieren, überblickshaft dargestellt werden. Die Einsatzgebiete sind vielfältig und die Entwicklung ausgesprochen dynamisch, weshalb die Darstellung nicht abschließend ist, vielmehr sollen die Potentiale veranschaulicht werden, die sich aus den Anwendungen für das Gesundheitswesen und in weiterer Folge für die Systeme der sozialen Sicherung auf tun.

- **Bild- und Sprachanalysen zur Früherkennung:** In der medizinischen Bildgebung ist KI bereits Teil des medizinischen Alltags. Die Entwicklung von Deep-Learning-Algorithmen hat zu erheblichen Fortschritten in der Bilderkennung und -analyse geführt und unterstützt bei der Analyse von medizinischen Bildern wie Röntgenbildern, CT-Scans und MRTs zur Früherkennung von Krankheiten. In vergleichbarer Weise wird KI in Verfahren der Spracherkennung, etwa zwecks Diagnose von Alzheimer, eingesetzt. Solche Verfahren zielen auf eine Erhöhung der Genauigkeit von Diagnosen, weil automatisiert Muster und Abnormalitäten in Bildern bzw. Auffälligkeiten in der Sprache identifiziert werden können, die für Menschen nur schwer zu erkennen sind. So wird beispielsweise in der Hautkrebsdiagnostik von KI-basierten Verfahren bereits eine vergleichbare oder sogar überlegene Leistung gegenüber der menschlichen Beurteilung durch Dermatologen erzielt. Zugleich muss allerdings einschränkend hinzugefügt werden, dass die Klassifikation von Bildern dort weniger gut ausfällt, wo die KI auf seltenere Entitäten mit unzureichenden Trainingsbildern trifft oder Bildartefakte zu Fehldiagnosen führen. Die besten Ergebnisse werden durch eine Kooperation von Mensch und Maschine erzielt. (vgl. Winkler & Haenssle, 2022)
- **Versorgung und Monitoring:** Weil die Überwachung des Gesundheitszustandes von Patienten auf der regelmäßigen Erhebung und Beobachtung medizinischer Daten beruht, bietet das Aufgabenfeld der Überwachung auf Auffälligkeiten ideale Voraussetzungen für die Automatisierung durch KI-Systeme. Beispielsweise wird KI in der Intensivstation eingesetzt, um postoperative Nachblutungen als gefährlichste und am häufigsten auftretende Komplikation in der Herzchirurgie durch Datenanalyse in Echtzeit vorausszusagen und hierdurch negative Folgen für Patienten zu verhindern. (vgl. Hohnstein, 2021) Gerade das Internet der Dinge eröffnet im Bereich der medizinischen Überwachung eine Vielfalt neuer Möglichkeiten: Geräte – vom Smartphone bis zu Maschinen in der Intensivstation – tauschen gesundheitsrelevante Informationen aus, um sie in Echtzeit zu evaluieren und gegebenenfalls Handlungsempfehlungen an das medizinische Personal zu senden.
- **Personalisierte Therapien:** Weil nicht alle Patienten auf eine Therapie gleichermaßen gut ansprechen, zielt der Ansatz der personalisierten Therapie darauf, ein tiefergehendes Verständnis der individuellen Krankheitsbilder zu entwickeln und durch die Analyse von medizinischen Daten, wie genetischer Informationen, Lebensgewohnheiten und klinischen Befunden, maßgeschneiderte Behandlungsempfehlungen zu geben. Solcherart maßgeschneiderte Lösungen führen zu einer gezielteren und effektiveren Therapie und minimieren das Risiko

unerwünschter Nebenwirkungen. Vor allem in der Krebsmedizin ist die personalisierte Therapie ein großer Hoffnungsträger. Eine auf den individuellen Patienten zugeschnittene Therapie setzt ein gutes Verständnis des Tumors voraus, wofür wiederum eine aufwändige Analyse von Daten und deren Interpretation notwendig ist.

- **Pflege:** In der Pflege kommt künstliche Intelligenz etwa in Form von Service-robotern (assistive Systeme wie Hebe- oder Putzhilfen, Monitoringsysteme wie Sturzüberwachung, Erinnerungsmodule für Bewegung, Medikamente und Nahrung), im Rahmen der Spracherfassung zur Pflegedokumentation und in Form von sozialen Robotern als Companion-Systeme zum Einsatz. (vgl. Nass & Schneider, 2022, S. 313) Als häufigste Einsatzfelder von KI in der (ambulant) Pflege nennt Peters (2024, S. 53–54) zum einen Tracking und Monitoring: So wird KI etwa zur Analyse von Bewegungs- und Vitaldaten eingesetzt, um Zustandsveränderungen zu erkennen bzw. zu klassifizieren, um daraufhin Abweichungen zu identifizieren und Frühwarnungen zu generieren. Mit Hilfe von Gangbildanalysen können Stürze nicht nur im Nachhinein erkannt werden, sondern es lässt sich anhand von verschiedenen Parametern wie etwa Schrittlänge, -höhe, -geschwindigkeit das Sturzrisiko erkennen, um präventive Maßnahmen einzuleiten. Zum anderen ist KI häufig im Bereich der Koordination und Kommunikation anzutreffen, um beispielsweise die Planung von Versorgungs- und Arbeitsprozessen zu unterstützen sowie den Informationsfluss zu optimieren. Dazu zählen etwa Aufgaben wie die Touren- und Personalplanung. Hierbei geht es darum, vorhandene Ressourcen bestmöglich bei gleichbleibender oder verbesserter Pflegequalität zu nutzen, indem Algorithmen „nicht nur individuelle Versorgungswünsche, Qualifikation und Verfügbarkeit des Personals, sondern auch Umgebungsdaten, wie Verkehrsaufkommen, Bedarfe und Möglichkeiten von Angehörigen und weiteren an der Versorgung beteiligten Berufsgruppen“ (Peters, 2024, S. 54) kombinieren.
- **Gesundheitsapps:** KI-basierte Anwendungen kommen im Alltag etwa zur kontinuierlichen Prüfung von Symptomen zum Einsatz. Dabei werden die KI-basierten Gesundheitsanwendungen von ihren Nutzern mit Daten gefüttert, ebenso werden Fragen zu eventuellen Beschwerden beantwortet, woraufhin ein Diagnosevorschlag ausgegeben wird. Weitere Anwendungen in KI-basierten Gesundheitsanwendungen können beispielsweise den Schlaf-Wach-Rhythmus überwachen und Verbesserungsmaßnahmen vorschlagen, das Sturzrisiko anhand von Videoaufzeichnungen bewerten oder Einschätzungen zu Hauterkrankungen geben. Zudem verbreiten sich Apps, die Menschen dabei unterstützen, die eigenen Gesundheitsdaten im Blick zu behalten und gesünder zu leben: Hierbei unterstützen Wearables wie etwa Schrittzähler, Pulsmesser, Smart Watches.

- **Entwicklung von Medikamenten:** KI kommt schließlich auch im Rahmen des sehr aufwändigen und folglich sehr teuren Entwicklungsprozesses von Medikamenten zur Anwendung, indem vor allem die analytischen Vorgänge dieses Unterfangens deutlich effizienter durchgeführt werden können. Für die Pharmaindustrie eröffnet KI große Potenziale, sowohl im Bereich der Wirkstoffentwicklung als auch bei der Umsetzung klinischer Studien Prozesse zu verkürzen und Kosten zu senken. (vgl. Schweighöfer & Pfannstiel, 2022)

Die aufgeführten Beispiele zeigen klar die großen Potenziale von KI für das Gesundheitswesen auf. Deren volle Ausschöpfung könnte erheblich zu einem Paradigmenwechsel beitragen, der mit der Ablösung der bisherigen, auf hohen Fallzahlen beruhenden Medizin durch ein value-basiertes System vollzogen wäre. KI verspricht verbesserte Behandlungsergebnisse, eine frühzeitige Erkennung von Krankheiten, die Eindämmung von Diagnose- und Behandlungsfehlern sowie eine Entlastung von Fachpersonal bei Routinetätigkeiten im medizinischen und pflegerischen Bereich.

Die gesetzliche Krankenversicherung hat qua Gesetz die Aufgabe, „die Gesundheit der Versicherten zu erhalten, wiederherzustellen oder ihren Gesundheitszustand zu verbessern“ (§1 SGB V). Während der KI-Einsatz im Bereich der medizinischen Leistungserbringung unmittelbar auf diesen Zweck hin ausgerichtet ist, können die durch die medizinische Leistungserbringung generierten Daten auch genutzt werden, um mittelbar auf diesen Zweck hinzuwirken. Insgesamt hält der KI-Einsatz damit große Potenziale bereit, Verbesserungen sowohl auf individueller Ebene des Gesundheitsschutzes herbeizuführen als auch auf gesellschaftlicher Ebene zu einer effizienteren Mittelverwendung beizutragen. Für Krankenkassen ist dies von Bedeutung, beruht deren Leistungsfähigkeit doch auf der Balance von Gesundheitsvorsorge sowie dem wirtschaftlichen Einsatz von Ressourcen.

So kann etwa die Analyse der Behandlungsdaten Aufschluss über die Wirksamkeit von Behandlungsmethoden und Medikamenten geben. Mit Hilfe von Clusteranalysen lassen sich individuelle Krankheitsverläufe vergleichen und daraus Informationen für eine beschleunigte Genesung gewinnen. Die Patientendaten können aber auch Hinweise zu einem bevorstehenden Ausbruch von Epidemien geben, wodurch frühzeitig Gegenmaßnahmen eingeleitet werden können. Ganz grundlegend kann KI zur Public-Health-Forschung ihren Beitrag leisten, um die Verbreitung und Verhinderung von Krankheiten in der Bevölkerung besser zu verstehen. Außerdem kann KI zwecks Optimierung der Versicherungsleistungen eingesetzt werden, um auf Basis der Behandlungsdaten die Wirtschaftlichkeit von medizinischen Behandlungen zu analysieren. In der Versorgungsforschung besteht die Möglichkeit, durch eine Analyse von Behandlungspfaden Erkenntnisse zur Verbesserung der Behandlungsqualität sowie der Effizienz und Wirtschaftlichkeit bestimmter Behandlungsmethoden zu gewinnen. Dies ist nicht nur für den Gesundheitsbereich, sondern ebenso den Pflegebereich von großem Nutzen. Auch im Rahmen der Entwicklung von Vorsorgeprogrammen kann eine KI-unterstützte Analyse anonymisierter Krankheitsdaten von Versicherten herangezogen

werden. Frühere Diagnosen oder individuelle Prognosen werden durch die Analyse von Risikofaktoren für schwerer erkennbare Krankheiten ermöglicht, wodurch ein Ausbruch von Krankheiten vermieden oder die Behandlung zu einem früheren Zeitpunkt begonnen werden kann. Zudem lassen sich auf Basis der riesigen den Krankenkassen verfügbaren Datenmengen individuell optimale Behandlungsstrategien identifizieren. KI-Verfahren sind zudem zur Verbesserung der Diagnose von seltenen Erkrankungen (Orphan Diseases) und zur Erkennung und Vermeidung von Medikamenten-Wechselwirkungen einsetzbar. (vgl. Wissenschaftlicher Beirat für Digitale Transformation der AOK Nordost, 2018, 7ff.)

3.2 KONKRETE BEISPIELE DES KI-EINSATZES IN DER SOZIALVERSICHERUNG

Obwohl die mediale Omnipräsenz des KI-Themas anderes vermuten lässt, muss an dieser Stelle vorausgeschickt werden, dass die Anzahl der Anwendungsfälle von KI im Bereich der Sozialversicherung sowohl global als auch in Deutschland noch äußerst begrenzt ist und die bereits existierenden Beispiele des KI-Einsatzes relativ jung sind und oftmals experimentellen Charakter haben. In der Literatur lassen sich zwar Berichte über KI-Anwendungen in einer Reihe von Ländern ausfindig machen, diese sind jedoch auf punktuelle Einsätze entlang des Leistungserstellungsprozesses der Sozialversicherung beschränkt. Zugleich deuten eine Vielzahl von Forschungsvorhaben und Bekenntnisse zu neuen digitalen Technologien einschließlich KI seitens Regierungen und Sozialversicherungsorganisationen jedoch darauf hin, dass wir erst am Anfang einer Entwicklung stehen – gerade in den letzten Jahren war eine Dynamik im Einsatz neuer KI-basierter Lösungen im Bereich der sozialen Sicherung zu beobachten, die in der nahen Zukunft eine Vielzahl von Innovationen verspricht.

Die folgende Darstellung von KI-Anwendungen soll daher nicht als abschließender Katalog aller existierender Anwendungsfälle verstanden werden, vielmehr handelt es sich um einen Streifzug durch die KI-Landschaft im Bereich des Sozialversicherungswesens, wobei schlaglichtartig paradigmatische Fälle herausgegriffen und kurz dargestellt werden sollen. Dies ist zunächst der praktischen Einschränkung geschuldet, dass Beispiele für praktische Einsatzfälle nicht offen zutage liegen, vielmehr bloß auf dem Wege von Veröffentlichungen auf dem Suchradar erscheinen. Im Sinne eines Ausblicks soll durch eine solche Rundschau nicht zuletzt aufgezeigt werden, welche Möglichkeiten die Anwendung von KI im Bereich der sozialen Sicherung grundsätzlich eröffnet. Aufgrund der Neuheit des Phänomens und des damit verbundenen Mangels an entsprechenden Daten bleibt eine empirische Bewertung, inwieweit die präsentierten Anwendungsfälle tatsächlich zur Realisierung der KI-Potenziale beitragen, an dieser Stelle außer Acht.

Relativ weit verbreitet sind im Front-End-Bereich bereits Chatbots zur Optimierung der Interaktion mit Versicherten. Einen gewaltigen Schub erfuhr der Einsatz von KI-basierten Dialogsystemen durch den verstärkten Informationsbedarf seitens der Bevölkerung im Zuge der COVID-19-Pandemie.

So sollte etwa in Belgien Chatbot Ori zur Bewältigung der gestiegenen Anzahl von Telefonanrufen beim Kontaktzentrum des Landesamts für Arbeit (ONEM) beitragen. Eine Vorläuferversion namens Marc startete im Jahr 2020 mit der Ausführung bloß einer einzigen Aufgabe: Bürger konnten auf diesem Weg schnell eine zur Erstellung der Steuererklärung nötige Kopie ihrer Lohnsteuerbescheinigung erhalten. Nachfolger Ori wurde erweitert, um ab 2021 eine Vielzahl verschiedener Fragen zu den Themen Arbeitslosigkeit und Unterbrechung der Erwerbstätigkeit zu beantworten. Darüber hinaus unterstützt der Chatbot bei der Navigation auf der ONEM-Webseite. In Abhängigkeit von den Nutzeranfragen werden die vom Chatbot abgedeckten Themen regelmäßig aktualisiert. (vgl. IVSS, 2022b; Zaber et al., 2024, S. 24) Um die komplexen administrativen und rechtlichen Informationen in verständlicher Weise bereitzustellen, stellte das belgische ONEM mit Kundendienstexperten, die den Chatbot laufend testeten, sicher, dass typische Nutzer den Chatbot verstehen. Weil eine dauerhaft hohe Dialogqualität und Antwortgenauigkeit wichtige Voraussetzungen für eine anhaltend hohe Nutzerakzeptanz darstellen, legte das ONEM ein besonderes Augenmerk auf die Identifikation fehlerhafter Antworten, um diese rasch korrigieren zu können. (vgl. IVSS, 2022b)

Auch in Norwegen behalf man sich, um der hohen Nachfrage nach Informationen und Sozialleistungen während der pandemiebedingten Lockdowns zu begegnen, indem konversationelle KI eingesetzt wurde. Mit dem die Norwegische Arbeitsmarkt- und Wohlfahrtsbehörde (NAV) unterstützenden Chatbot namens Frida war es möglich, eintreffende Anfragen schnell zu bearbeiten, ohne für diese Übergangszeit übermäßig Personal aufstocken zu müssen. Der Chatbot wird mit einer täglich aktualisierten Wissensdatenbank trainiert und durch einen menschlichen Experten kontrolliert. (vgl. IVSS, 2020)

Österreich setzt in der Arbeitsmarktverwaltung mit dem „Berufsinformat“ (www.ams.at/berufsinformat) auf künstliche Intelligenz, um den Beratungsprozess bei der Arbeitssuche zu erleichtern. Der Chatbot steht auf der Internetseite des Arbeitsmarktservice (AMS) zur Verfügung, um Fragen rund um Berufsbilder, Ausbildungsmöglichkeiten, Gehaltsniveaus, Weiterbildungen und ähnliche berufsbezogene Themen zu beantworten. Dabei greift das KI-Tool auf die wesentlichen Datenbanken des AMS zu und wird mit Hilfe des abgegebenen Feedbacks laufend aktualisiert. Der Zweck der Anwendung liegt darin, schnelle Antworten zu liefern und den Nutzen des Beratungsgesprächs durch Wegfall repetitiver Tätigkeiten und Recherchen zu erhöhen. Dementsprechend können nur berufsspezifische, öffentlich verfügbare Informationen abgerufen werden. Es besteht kein Zugang zu personenbezogenen Daten. Folglich ist der „Berufsinformat“ nicht für die Jobvermittlung oder Auskünfte zu Ansprüchen auf Arbeitslosengeld sowie Förderungen geeignet. In Finnland ist mit dem Dialogsystem Kela-Kelpo ein mehrsprachiger Chatbot im Einsatz, der Nutzern bei der Informationssuche über Leistungen im Selbstbedienungswebportal der finnischen Sozialversicherungsanstalt (KELA) hilft und personalisierte Auskünfte gibt. Mit Hilfe von NLP gelingt eine möglichst natürlich anmutende Interaktion. (vgl. Zaber et al., 2024, S. 25) Um die Akzeptanz der Nutzer hoch zu halten, kümmerten sich um Fragen, die der Chatbot häufig nicht beantworten konnte, Mitarbeiter sofort online. (vgl. IVSS, 2022b)

Neben Chatbots im Front-End-Bereich finden sich KI-Anwendungen zunehmend auch im Back-End-Bereich. Dabei geht es in erster Linie darum, Arbeitsentlastungen durch Effizienzsteigerungen zu erzielen, um Kapazitäten von Mitarbeitern für andere Aufgaben zu schaffen bzw. Kapazitäten abzubauen. Auch steht im Fokus, Vorgänge zu beschleunigen, um die Kundenzufriedenheit zu erhöhen. Zudem geht es in vielen Fällen um die KI-basierte Vorbereitung von danach menschlich zu treffenden Entscheidungen. Der Mehrwert der Technologie liegt dann vor allem darin, große Datenmengen zu sichten und Muster zu erkennen, um darauf aufbauend Mitarbeitern eine informierte Entscheidung zu ermöglichen. In vielen Fällen werden im Vergleich zu den manuellen Prozessen mehr Daten hinzugezogen, um Entscheidungen auf eine breitere Basis zu stellen und dadurch insgesamt bessere Entscheidungen auf effizienterem Weg zu treffen.

Seit Anfang 2019 ist bei der deutschen Bundesagentur für Arbeit ein KI-basiertes System zur Klassifikation und Informationsextraktion von Studienbescheinigungen für die Familienkasse im Einsatz, um Antragsverfahren zu beschleunigen. Weil zur Beantragung des Kindergeldes zweimal jährlich die aktuelle Studienbescheinigung des betreffenden Kindes vorgelegt werden muss, entsteht ein hohes Aufkommen manuell zu prüfender Dokumente. Erschwerend tritt hinzu, dass die Hochschulen kein einheitliches Formular benutzen. Der KI-basierte Lösungsansatz sieht vor, dass der Text der von den Antragstellern auf einer entsprechenden Webseite hochgeladenen Studienbescheinigungen mittels OCR (Optical Character Recognition) extrahiert und nachfolgend klassifiziert wird. Dabei wird geprüft, ob und mit welcher Wahrscheinlichkeit das hochgeladene Dokument eine Studienbescheinigung ist, ob es sich dabei um das richtige Kind, den richtigen Gültigkeitszeitraum und um eine bekannte deutsche Hochschule handelt. Der von der KI generierte Output wird den Mitarbeitern im Fachverfahren der Familienkasse angezeigt, wobei das Ergebnis der Prüfung „Studienbescheinigung ja/nein“ aktiv zu akzeptieren ist. Das System basiert auf dem Verfahren des überwachten Lernens, wobei semi-automatisiert annotierte und anonymisierte Studienbescheinigungen Trainingszwecken dienen, und wird regelmäßig trainiert. Vorteilhaft wirkt sich in der Bundesagentur für Arbeit durch den KI-Einsatz aus, dass Mitarbeiter von fehleranfälligen, monotonen und einfachen Massenaufgaben befreit werden, dadurch eine große Arbeitsentlastung erfahren und ihre Zeitressourcen für den Prüfaufwand an anderer Stelle einsetzen können. Zugleich wird es möglich, Versicherten schneller eine Rückmeldung zu geben und die Bearbeitungszeit zu verkürzen. (vgl. Bundesagentur für Arbeit, 2024; Engelmann & Puntschuh, 2020; Hofmann, 2021)

Die Gründung eines Kompetenzzentrums für KI und Machine Learning im IT-Systemhaus der Bundesagentur für Arbeit unterstreicht den hohen Stellenwert, den künstliche Intelligenz für die Behörde einnimmt. KI dient innerhalb der Bundesagentur für Arbeit neben der Automatisierung von Dokumentenerkennung zur Bewältigung der täglichen Flut eingehender Dokumente auch der Unterstützung von Analysen und Prognosen, indem KI kontextbasierte Informationsauswertungen rund um das Arbeitsmarktgeschehen liefert. Mit Stea-BERT ist darüber hinaus eine KI-Anwendung gegeben, die den Kern der Aufgabenerfüllung der Bundesagentur

für Arbeit betrifft. Die Arbeitsvermittlungs-KI hilft dabei, Informationen aus Stellenanzeigen zu extrahieren. Mittels Machine Learning wird aus den unstrukturiert vorliegenden Informationen eine Reihe von Attributen extrahiert (z. B. Voll-/Teilzeit, Befristung, Branche, geforderte Kompetenzen), um das Erstellen und Publizieren von Stellenanzeigen und schließlich das Matching zwischen Arbeitsuchenden und offenen Stellen zu erleichtern. (vgl. Bundesagentur für Arbeit, o.D.; Hofmann, 2021)

Darüber hinaus setzt die Bundesagentur für Arbeit eigenen Angaben zufolge künstliche Intelligenz in einer Reihe weiterer Bereiche ein: Im Arbeitgeberservice der Bundesagentur für Arbeit werden die zahlreichen Informationen aus Stellenmeldeformularen und Mails von Arbeitgebern KI-basiert strukturiert und automatisch Stellenangebote generiert. Für die Mitarbeiter ergibt sich daraus eine enorme Zeitersparnis, weil sie die erstellten Vorschläge nur noch prüfen müssen. (vgl. Bundesagentur für Arbeit, 2024)

Um der Bedeutung des KI-Themas Rechnung zu tragen, ist am Institut für Arbeitsschutz (IFA) der Deutschen Gesetzlichen Unfallversicherung (DGUV) als Spitzenverband der Berufsgenossenschaften das Kompetenzzentrum Künstliche Intelligenz und Big Data (KKI) angesiedelt, das fachliches Knowhow zu diesen Zukunftstechnologien im Zusammenhang mit Sicherheit und Gesundheit bei der Arbeit bündeln und eine zentrale Anlaufstelle für die DGUV und ihre Träger darstellen soll.

So wird KI etwa zur Risikoprävention eingesetzt, um im Vergleich zu manuellen Verfahren zu besseren Einschätzungen im Hinblick auf Gefahrenlagen zu gelangen und Versicherungsfälle erst gar nicht entstehen zu lassen. Die Berufsgenossenschaft der Bauwirtschaft (BG BAU), Träger der gesetzlichen Unfallversicherung für die Bauwirtschaft und baunahe Dienstleistungen, nutzt KI beispielsweise zur automatisierten Identifizierung von Unternehmen, die angesichts ihres Unfallgeschehens erhöhten Beratungsbedarf haben. Als Träger der gesetzlichen Unfallversicherung haben die Berufsgenossenschaften die Aufgabe, Arbeitsunfälle und Berufskrankheiten zu verhindern. Weil in der Baubranche Unfallzahlen überdurchschnittlich hoch und die Folgen von Unfällen häufig gravierend sind, spielt Präventionsarbeit in diesem Bereich eine besonders wichtige Rolle. BG BAU entwickelte hierfür eine KI-Anwendung, die zeiteffizienter und weniger fehleranfällig Resultate erbringen soll als die bisherige Vorgehensweise mit manuell erstellten Tabellen. Mittels KI können größere Datenmengen als bisher systematisch und schnell analysiert werden, um das Unfallrisiko für betreffende Unternehmen mit Hilfe eines Ampelsystems übersichtlich darzustellen. Durch eine wöchentliche Auswertung sind zudem schneller Trends zu erkennen. Die KI-basierte Identifizierung beratungsbedürftiger Unternehmen soll dazu beitragen, die Kapazitäten der Aufsichtspersonen der BG Bau möglichst zielgerichtet einzusetzen, um in der Folge Unfälle zu vermeiden. Die KI wird durch Rückmeldungen aus dem Einsatz der Aufsichtspersonen weiterentwickelt, um die Treffsicherheit weiter zu verbessern. Bei diesem Projekt handelt es sich um ein durch das Bundesministerium für Arbeit und Soziales (BMAS) gefördertes Leuchtturmprojekt mit Modellcharakter, das auf andere Bereiche ausstrahlen

soll: Aktuell prüft die BG BAU weitere Anwendungsfälle für den Einsatz künstlicher Intelligenz in der eigenen Präventionsarbeit, darüber hinaus ist angedacht, die Ergebnisse auf weitere Träger der gesetzlichen Unfallversicherung und auf darüber hinausreichende Anwendungsfälle im Geschäftsbereich des BMAS zu übertragen. (vgl. Denkfabrik BMAS, 2024a; o.V., 2023)

In der Berufsgenossenschaft Energie Textil Elektro Medienerzeugnisse (BG ETEM), der Trägerin der gesetzlichen Unfallversicherung für die Unternehmen in den Bereichen Energie, Textil, Elektro, Medienerzeugnisse, werden Entscheidungen in Bezug auf die Durchsetzung von Regressansprüchen KI-basiert vorbereitet. In der dortigen Regressabteilung ist KI seit Ende 2019 zur Steuerung der Fallprüfung im Regressmeldeverfahren im Einsatz. Die Berufsgenossenschaft hat im Falle von Unfällen am Arbeitsplatz oder auf dem Arbeitsweg, die durch betriebsfremde Dritte verursacht wurden, zu prüfen, ob eine Regressnahme möglich ist. Bislang sah der Prozess vor, im Falle des Regressverdachts ein standardisiertes Dokument an die zentrale Stelle für Regressverfahren zu übersenden und manuell vorzuprüfen. Unter Anwendung von überwachtem Lernen wurde eine KI anhand von ungefähr einer Million abgeschlossener Regressverfahren trainiert, die Wahrscheinlichkeit für ein erfolgreiches Regressverfahren zu bewerten. Zusätzlich wird mittels Textmining nach kritischen Schlagwörtern gesucht, um durch weitere Informationen die Entscheidungsbasis zu verbreitern. Der Regressabteilung werden die Prognose sowie weitere zur Entscheidung erforderliche Informationen über ein Web-Interface zur Verfügung gestellt. Das System wird laufend überwacht und auf Plausibilität geprüft. (vgl. Engelmann & Puntschuh, 2020; Krause, 2023)

Im Unfallbereich setzt die BG ETEM seit 2021 eine KI-Lösung ein, die bei der Beurteilung von Reha-Fällen unterstützt. Reha-Plus, wie die Anwendung heißt, leistet Mitarbeitern Entscheidungshilfe, wenn es um die Beurteilung geht, ob Versicherte nach schweren Arbeits- oder Wegeunfällen durch das Reha-Management begleitet werden sollen. Gerade bei Grenzfällen kann die KI-basierte Anwendung durch den Abgleich von etwa 320 Merkmalen den Sachbearbeitern wertvolle Entscheidungsunterstützung leisten. (vgl. Krause, 2023) Wie BG BAU hat die BG ETEM zudem eine KI-Anwendung im Einsatz, die bei der Unfallprävention unterstützt, indem Unternehmen mit einer erhöhten Wahrscheinlichkeit für voraussichtlich eintretende Arbeitsunfälle auffindig gemacht werden. Die Zahl der Arbeitsunfälle soll durch die Verbesserung einer zielgerichteten Beratung und Aufsicht reduziert werden. (vgl. Krause, 2023)

Mit dem Projekt KIRA (Künstliche Intelligenz für risikoorientierte Arbeitgeberprüfungen) entwickelt DRV Bund eine Anwendung, die Mitarbeiter des Prüfdienstes KI-basiert bei der Auswahl der Prüfschwerpunkte unterstützen soll. Kernaufgabe des Prüfdienstes ist es sicherzustellen, dass Unternehmen für ihre Mitarbeiter angemessene Sozialabgaben leisten, damit im Fall von Krankheit, Arbeitslosigkeit und im Ruhestand korrekter Versicherungsschutz greift. Die zur Erfüllung dieser Aufgabe erforderlichen Betriebsprüfungen bedeuten ein zeitaufwändiges Verfahren, das aktuell manuell durchgeführt wird. Weil die Kapazitäten nicht ausreichen, um

die Fülle von Unterlagen zur Identifikation von Unregelmäßigkeiten zu sichten, beschränkt man sich auf eine stichprobenhafte Prüfung. Mittels KI soll der Prüfdienst bei der Auswahl der Prüfschwerpunkte unterstützt werden. Dazu durchleuchtet die KI die elektronisch vorliegenden Unterlagen, sucht nach Mustern und liefert zweierlei Arten von Informationen zur Erleichterung der eigentlichen Prüftätigkeit der Mitarbeiter: Zum einen wird ein Score errechnet, der das Ausmaß der gefundenen Auffälligkeiten in einem konkreten Prüffall angibt, und zum anderen zeigt die KI jene Stellen auf, wo die Auffälligkeiten entdeckt wurden. Durch die KI-basierte Voranalyse ergibt sich eine enorme Zeitersparnis, die in der Folge in die detaillierte Prüfung investiert werden kann. Auf diese Weise steigert KIRA die Effektivität des Prüfdienstes und trägt dadurch zu mehr sozialer Gerechtigkeit und dem Schutz des deutschen Sozialversicherungssystems bei. Auch KIRA ist ein Leuchtturmprojekt des Bundesministeriums für Arbeit und Soziales und wird entsprechend vom Ministerium gefördert. Das Projekt befindet sich in der Entwicklungsphase und wird bei erfolgreichem Verlauf im Jahr 2025 im Betriebsprüfdienst der DRV eingesetzt werden. (vgl. Denkfabrik BMAS, 2024b; DRV Bund, o.D.)

Die Österreichische Gesundheitskasse (ÖGK), Trägerin der Pflichtkrankenversicherung in Österreich, nutzt KI zur Beschleunigung der Wahlarztkostenrückerstattung. Als Wahlärzte werden in Österreich niedergelassene Privatärzte ohne Vertrag mit den gesetzlichen Krankenkassen bezeichnet, die ihre Honorierung jedoch an den Krankenkassentarifen orientieren. Krankenversicherte erhalten auch für die Behandlung durch Wahlärzte bis zu 80 Prozent des Tarifs von ihrer gesetzlichen Krankenkasse rückerstattet. Durch Einsatz von KI zur automatisierten Bearbeitung der eingereichten Rechnungen möchte die ÖGK die Wartezeit der Rückerstattung deutlich verkürzen. (vgl. ÖGK, 2024, 6f.)

Estland, das Vorzeigeland der digitalen Transformation, plant eine Vielzahl KI-basierter Anwendungen im Bereich der öffentlichen Verwaltung. So setzt der baltische Staat etwa im Bereich der Arbeitslosenversicherung ein Tool namens OTT ein, um die zeitaufwändige Erstellung von individuellen Strategien zur Rückkehr auf den Arbeitsmarkt zu unterstützen. Die Anwendung basiert auf einem fortschrittlichen Modell maschinellen Lernens und soll Entscheidungen unterstützen, die zu einer besseren Qualität der Vermittlungsleistung führen und die Effizienz der Organisationsprozesse erhöhen. Dazu analysiert das Tool die Profile Arbeitssuchender (z. B. Ausbildung, Berufserfahrung, gesundheitliche Einschränkungen) und berücksichtigt zudem Wirtschafts- und Arbeitsmarktdaten, um die Wahrscheinlichkeit für eine Person vorherzusagen, einen neuen Job zu finden. Zudem wird die Wahrscheinlichkeit berechnet, innerhalb eines Jahres wieder arbeitslos zu werden. OTT präsentiert die diese Wahrscheinlichkeiten beeinflussenden Faktoren, was für die Auswahl geeigneter Services hilfreich sein kann. Die Anwendung zielt darauf, Ressourcen effizient einzusetzen, indem sie aufzeigt, wo die Prioritäten für Vermittlungsbemühungen liegen sollten: Bei Personen mit geringem Risiko wird davon ausgegangen, dass sie selbst einen Job finden, weshalb sie weniger Unterstützung erhalten. Personen, denen OTT ein höheres Risiko für Langzeitarbeitslosigkeit bescheinigt, benötigen hingegen größere Aufmerksamkeit – auf sie richten

sich vorrangig die Unterstützungsleistungen. Insgesamt wird mit dem Einsatz einer KI-basierten Entscheidungshilfe bezweckt, einen besseren Service zu bieten, ohne zusätzliche Kosten zu erzeugen. (vgl. Dreyling III et al., 2023)

In vergleichbarer Weise setzt die flämische Arbeitsverwaltung (VDAB) ein statistisches Profiling-Modell unter Verwendung von Algorithmen maschinellen Lernens ein, um die Wahrscheinlichkeit einzuschätzen, dass jemand langzeitarbeitslos wird. Die zugrundeliegenden Daten werden gesammelt und gespeichert und enthalten detaillierte Informationen über die sozioökonomischen Merkmale der Arbeitssuchenden sowie deren Arbeitshistorie. Außerdem werden Informationen in das Modell einbezogen, die von Sachbearbeitern während früherer oder der aktuellen Arbeitslosigkeit erhoben wurden. Als Indikator für das Verhalten und die Motivation bei der Arbeitssuche werden auch Klickdaten, welche die Aktivitäten der Arbeitssuchenden auf der Webseite der öffentlichen Arbeitsverwaltung nachvollziehen lassen, einschließlich des Anklickens von Stellenangeboten, im Profiling berücksichtigt. Wie auch beim estnischen Modell trifft das System selbst keine Entscheidungen, vielmehr soll den Sachbearbeitern eine Entscheidungsunterstützung gegeben werden, um die knappen Ressourcen im Sinne der Gesamteffizienz zu verwenden. (vgl. Desiere et al., 2019, S. 17)

Auch der österreichische Arbeitsmarktservice (AMS) setzt eine KI-Anwendung ein, um darüber zu entscheiden, wie gut Arbeitssuchende am Arbeitsmarkt vermittelbar sind. In die Bewertung des sogenannten Arbeitsmarktchancen-Assistenzsystem (AMAS) fließen verschiedene personenbezogene Daten wie etwa Alter, Geschlecht, Ausbildung und Arbeitserfahrung ein, um zu einer Priorisierung von Personen nach ihrer Vermittelbarkeit zu gelangen. Der Algorithmus teilt Arbeitssuchende in drei Kategorien mit hohen, mittleren und niedrigen Arbeitsmarktchancen ein, um Fördermaßnahmen zielführender vergeben zu können und insgesamt die Arbeitslosigkeit schneller zu senken. Die Frage, ob diese Vorgehensweise nicht Profiling darstellt und als „automatisierte Einzelfallentscheidung“ nach der Datenschutzgrundverordnung (DSGVO) gilt und demzufolge nur unter sehr engen gesetzlichen Voraussetzungen erfolgen darf, beschäftigt seit Jahren die österreichischen Gerichte. (vgl. Pflügl, 2024)

Der Einsatz von KI zum Zwecke der Aufdeckung von Sozialbetrug ist in der Praxis immer wieder Anlass zu Verletzungen der Privatsphäre und Diskriminierungen. So hat sich etwa Dänemark, ein Land mit einem traditionell stark ausgebauten sozialen Sicherheitsnetz, zur Einrichtung eines Systems zur Aufdeckung von Sozialbetrug entschlossen. Durch die Verknüpfung von Datenbanken, in denen Informationen über Steuern, Wohnungen, Autos, Beziehungen, Arbeitgeber, Reisen und Staatsangehörigkeit vorgehalten werden, entsteht ein weitreichendes Überwachungsnetz. Mit Hilfe eines Modells maschinellen Lernens werden die Daten analysiert, um Vorhersagen zu treffen, welche Personen das System betrügen könnten. Vom Dänischen Institut für Menschenrechte wurde diese Vorgehensweise ebenso kritisiert wie von der Dänischen Datenschutzbehörde. Eine abschließende Beurteilung, inwieweit die Verwendung der Daten zur Betrugserkennung gemäß den europäi-

schen Datenschutzgesetzen verhältnismäßig ist, steht noch aus. (vgl. Geiger, 2023) Auch in den Niederlanden diente ab 2014 ein System namens „System Risk Indication“ (SyRI) zur automatischen Identifizierung potentieller Sozialbetrüger. Auch hier wurden bei verschiedenen Behörden und staatlichen Einrichtungen gespeicherte Daten zusammengeführt und analysiert. Bestimmte Risikoindikatoren sollen erkennen lassen, ob jemand gegen gesetzliche Bestimmungen verstößt. Anfang 2020 wurde wegen Intransparenz und unangemessener Datenerhebung gerichtlich die sofortige Einstellung des Programms angeordnet und daraufhin beendet. Während der Zweck von SyRI nicht in Frage gestellt wurde, zogen Umsetzung und Praxis des Systems heftige Kritik auf sich. (vgl. Nenno, 2021) Nachdem sich herausstellte, dass das System Tausende von Menschen zu Unrecht des Kindergeldbetrugs bezichtigte, sah sich die niederländische Regierung zum Rücktritt gezwungen. Die verantwortliche Behörde forderte die Verdächtigten – oftmals mit geringem Einkommen oder Migrationshintergrund – fälschlicherweise auf, das über Jahre hinweg erhaltene Kindergeld zurückzuzahlen, was Zehntausende von Familien in schwere Verschuldung und Armut trieb sowie die Unterbringung vieler Kinder in Pflegefamilien nach sich zog. (vgl. Rao, 2022)

3.3 VORAUSSETZUNGEN UND ERFOLGSFAKTOREN EINES EFFEKTIVEN EINSATZES VON KI IN ORGANISATIONEN DER SOZIALEN SICHERUNG

Über Erfolg oder Misserfolg von Einführung und nachfolgendem Betrieb von Anwendungen künstlicher Intelligenz entscheidet nicht nur die Technik allein. Die in Sozialversicherungsorganisationen eingesetzten KI-basierten Anwendungen sind letztlich nur so gut wie die Menschen, die mit ihr arbeiten, die System- und Datenlandschaft, in die sie integriert werden, und die Ablauf- und Aufbauorganisation, in die sie eingebettet sind. Das heißt: Hand in Hand mit der Entwicklung, der Erprobung und dem Einsatz der KI-Lösungen müssen die Analyse und Gestaltung mitarbeiterbezogener und organisationaler Aspekte der Arbeit erfolgen.

Dass das Einbringen neuer Technik in Organisationen nicht nur eine technische Herausforderung darstellt, ist Kerngedanke des Gestaltungskonzepts Mensch-Technik-Organisation (MTO), das sich seit den 1950er Jahren in den Arbeitswissenschaften durchgesetzt hat und Unternehmen, Behörden und sonstige Einrichtungen als dynamische und komplexe Systeme betrachtet. Dabei ist die Überzeugung leitend, dass die Teilsysteme Mensch, Technik und Organisation durch die Arbeitsaufgabe verknüpft sind und in ihrer gegenseitigen Abhängigkeit und ihrem Zusammenwirken zu betrachten sind. Für die zufriedenstellende Bewältigung der jeweiligen Arbeitsaufgabe ist es entscheidend, dass die drei Teile des Arbeitssystems gut aufeinander abgestimmt sind. (vgl. Ulich, 2013)

Unter dem Schlagwort Human-in-the-Loop (vgl. etwa Rueckert & Riedl, 2022) wird diskutiert, inwiefern Teams aus Mensch und Maschine zu besseren Ergebnissen führen als dies mit rein automatisierten Entscheidungen möglich wäre. Doch auch aus anderen Gründen wird gefordert, den Menschen mit seinem Urteilsvermögen

in Prozesse zu integrieren: Weil maschinelle Entscheidungen oftmals nicht frei von Vorurteilen sind oder sie ein Verständnis des Kontexts missen lassen, werden KI-basierte Entscheidungen häufig individuellen Situationen nicht gerecht. Human-in-the-Loop verfolgt daher den Ansatz, durch die aktive Einbindung von Menschen einen Unterschied in automatisierten Entscheidungsprozessen zu bewirken. Hierbei stehen Fragen bezüglich der Ausgestaltung einer sinnvollen Interaktion zwischen Mensch und Maschine im Vordergrund.

Gerade im Bereich der Sozialversicherung kommt es nicht nur auf die rechtliche Sicherheit, sondern ebenso auf die Nachvollziehbarkeit von Verfahren an. Ein menschenzentrierter Entwicklungsprozess, der den Menschen von Beginn an im Sinne des Human-in-the-Loop-Ansatzes mitdenkt, erscheint schon allein aus diesem Grund notwendig. Weshalb also neben der konkreten technischen Ausgestaltung des KI-Systems auch Mensch und Organisation eine entscheidende Rolle für ein gutes Gesamtergebnis spielen, wird aus der Betrachtung der drei Faktoren durch die „KI-Brille“ erkennbar:⁹

- **Mensch:** Künstliche Intelligenz ist heute vor allem in der Lage, in ganz bestimmten Bereichen Probleme schnell und effizient zu lösen. KI ist immer noch darauf angewiesen, von Menschen entwickelt und mit passenden Daten trainiert zu werden. Zudem müssen Menschen überhaupt erst die zu lösenden Probleme erkennen und die optimale Art der Zusammenarbeit zwischen Mensch und Maschine gestalten. Insbesondere die Implementierung des KI-Systems erfordert menschliches Zutun, zugleich unternimmt der Mensch – etwa durch die Auswahl und Aufbereitung der Trainingsdaten – wesentliche Weichenstellungen, die über das spätere Leistungsvermögen des Systems bestimmen. Damit KI für eine Organisation Mehrwert schafft, ist es erforderlich, dass die Mitarbeiter ein Verständnis entwickeln, welche Möglichkeiten mit der neuen Technologie realisierbar sind und wo die Grenzen derselben liegen.

Die Einführung neuer Technologien bringt immer auch Zweifel, Ängste und Sorgen der Belegschaft mit sich. Menschen müssen „mitgenommen“ werden, damit die Arbeit mit KI gelingt. Weil gerade künstlicher Intelligenz ein nicht zu ignorierendes Maß an Skepsis voraussetzt, gilt es, Mitarbeiter aufzuklären, Mythen zu entkräften, Vorbehalte auszuräumen und Menschen aktiv und gestalterisch einzubeziehen. Da Einführung und Umgang mit KI-Technologien extensives Knowhow bei Entwicklern und Anwendern voraussetzen, ist darüber hinaus einerseits für ausreichende Aus- und Weiterbildung zu sorgen, andererseits sind permanentes Lernen und ein nachhaltiges Wissensmanagement sicherzustellen. Dabei hat der Kompetenzaufbau nicht nur fachliche Aspekte in den Blick zu neh-

⁹ Die Schwierigkeiten, dieses Konzept auf die Einführung von KI-Systemen anzuwenden, liegen freilich darin, dass die Rolle, die KI innerhalb einer Organisation spielt, sehr unterschiedlich ausgeprägt sowie nur begrenzt sichtbar sein kann. KI-basierte Anwendungen können punktuell in bestimmten Bereichen bzw. für bestimmte Aufgaben in einer Organisation eingesetzt werden oder aber derart umfassend und integriert, dass die Grenzen zwischen den MTO-Faktoren nur noch schwer zu ziehen sind. Zudem ist es möglich, dass KI nicht für den Nutzer erkennbar in Erscheinung tritt bzw. das Systemverhalten nicht nachvollziehbar ist. Wo wird die Maschine tätig, wo beginnt die Arbeit des Menschen?

men: Weil Routineabläufe an lernende Systeme delegiert werden, sind Mitarbeiter gefordert, die vor allem Problemlösungsfähigkeiten und Führungsqualitäten aufweisen.

- **Technik:** Dass das KI-System selbst technisch funktionsfähig sein muss, versteht sich von selbst. Damit nicht nur die Implementierung eines KI-Systems reibungslos vonstatten geht, sondern die nachfolgende Anwendung auch die mit dem Einsatz verknüpften Erwartungen erfüllt, ist es essentiell, die KI fortlaufend zu trainieren. Im Zusammenhang mit den oben dargestellten Beispielen wird immer wieder die Wichtigkeit betont, die Algorithmen zu überwachen und fortlaufend zu optimieren.

Der Technikaspekt betrifft jedoch auch die bestehende IT-Landschaft der betreffenden Organisation, in die KI eingebracht werden soll. Die Nutzbarmachung von künstlicher Intelligenz erfordert mehr als die Installation einer zusätzlichen Software und die Anbindung an eine Datenbank. Vielmehr ist eine umfassende Integration in das bestehende IT-System zu bewerkstelligen, um beispielsweise einen automatischen Zugriff auf benötigte Daten zu erlauben.

Zu den zu berücksichtigenden technologischen Aspekten einer KI-Einführung zählt auch der Aufbau einer soliden Datenbasis und einer geeigneten Dateninfrastruktur. Es sind Strategien und Prozesse erforderlich, um geeignete Daten bereitzustellen und gezielt zu verknüpfen. Ein entscheidender Schritt ist hierbei die Identifizierung von geeigneten Daten und Datenquellen entsprechend den Zwecken der jeweiligen Aufgabe. Zu beachten ist hierbei, dass gegebenenfalls nicht alle benötigten Daten digital vorliegen.

- **Organisation:** Nicht nur in die technische Infrastruktur, sondern auch in die bestehende Prozesslandschaft muss das KI-System integriert werden. Im Zuge der Einführung von KI-Anwendungen ist zugleich die Anpassung der organisatorischen Prozessstrukturen zu prüfen, weil KI zu einer veränderten Art und Weise der Aufgabenerfüllung führen wird. Herkömmliche Prozesse können sich dabei als gänzlich obsolet herausstellen, weil sie automatisiert werden, oder sie können sich verkürzen. Aufgaben und Verantwortlichkeiten in der gesamten Sozialversicherungsorganisation können sich verändern.

Im Zuge der Einführung von KI-Lösungen ist es hilfreich, durch ein aktives Veränderungsmanagement in der Organisationskultur ein Bewusstsein für KI zu schaffen, damit sich Akzeptanz für die neuen Herangehensweisen und Abläufe einstellt sowie eine organisationsweite Zusammenarbeit ermöglicht wird. Der Einsatz von künstlicher Intelligenz ist mehr als das Einbringen eines neuen technischen Arbeitsmittels in die Sozialversicherungsorganisation. Weil KI die Abläufe und Herangehensweisen grundlegend verändert, braucht es starke Führungspersönlichkeiten zur Begleitung des Wandels der Organisationskultur. Diese wird stärker als zuvor gekennzeichnet sein durch die Bereitschaft zu agilem Handeln, kontinuierlicher Anpassung und verantwortlichem Entscheiden.

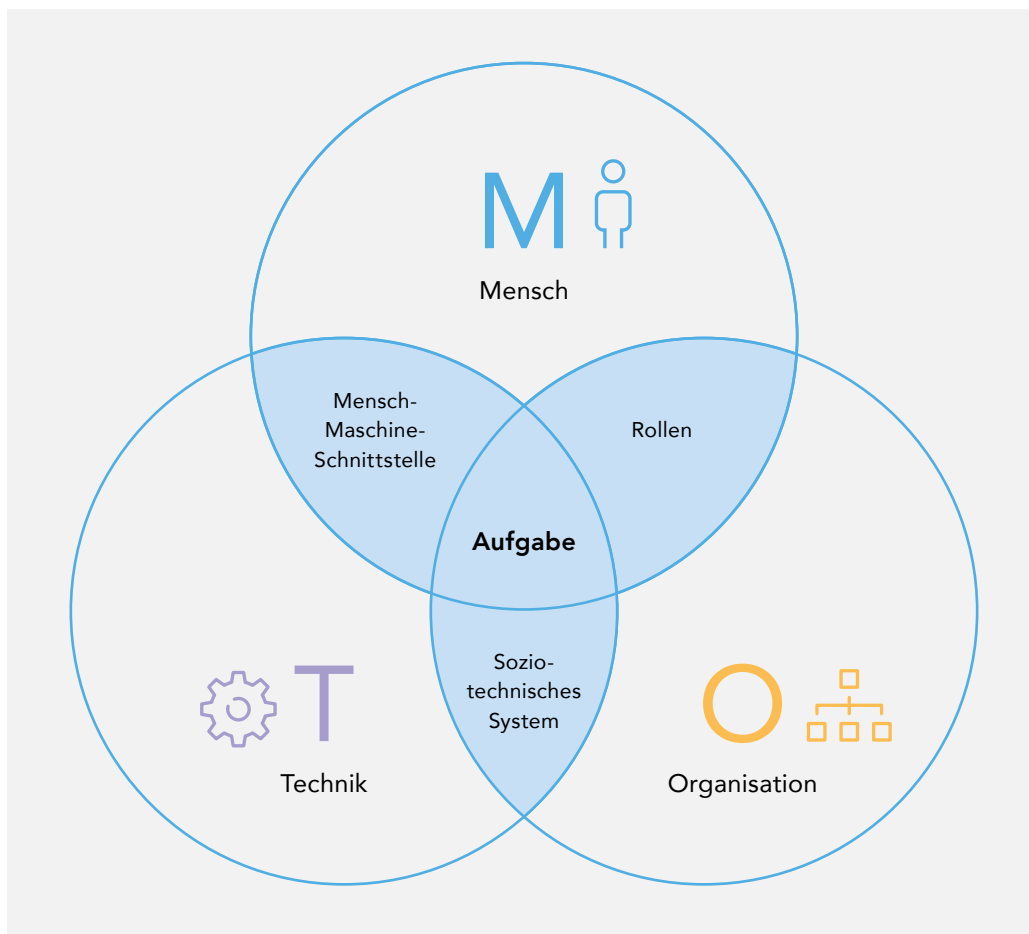


Abbildung 11: MTO-Konzept; Quelle: Eigene Darstellung

Betrachtet man obige Grafik, so fällt ins Auge, dass KI gerade an den Schnittstellen zwischen den drei Faktoren Mensch, Technik und Organisation ganz spezifische Fragen aufwirft. Mensch und Organisation treffen sich aufgrund veränderter Aufgaben und Prozesse auf neue Weise: Menschen füllen neue Rollen aus, weil ihnen gewandelte formelle und informelle Verhaltenserwartungen entgegengebracht werden. Diese neuen Rollen gilt es exakt zu definieren und entsprechend handelnd umzusetzen. KI verspricht vielfältige Effizienz- und Qualitätssteigerungen, verändert zugleich aber die Arbeitsumgebung für die Beschäftigten. Ein performantes Gesamtsystem zu gestalten, setzt die ganzheitliche Sicht auf das soziotechnische System voraus, um natürliche und künstliche Intelligenz mit ihren jeweiligen Stärken und Schwächen aufeinander abzustimmen. Gleichzeitig ist klar, dass KI die Aufgabenteilung zwischen Mensch und Maschine grundlegend verändert. Dabei die Mensch-Maschine-Interaktion „menschengerecht“ auszugestalten, gilt als gesellschaftlicher Konsens. Dies erfordert klare Orientierungsmaßstäbe für die neue Aufgabenverteilung in den betreffenden Organisationen der Sozialversicherung.

Auf lange Sicht spielt für die erfolgreiche Implementierung und Nutzung von KI-Anwendungen in Sozialversicherungsorganisationen darüber hinaus die Bewältigung einer Reihe von spezifischen Herausforderungen eine wesentliche Rolle. Im Rückgriff auf Wirtz et al. (2019, 601ff.) lassen sich vier wesentliche Dimensionen solcher Herausforderungen differenzieren:¹⁰

- **Technologisch-implementierungsbezogene Herausforderungen:** Unter technischen Implementierungsgesichtspunkten geht es darum, die mit der Technologie verknüpften Potenziale auszuschöpfen, wobei Sozialversicherungsorganisationen mit spezifischen Herausforderungen zu kämpfen haben. Allen voran steht die Gewährleistung von **KI-Sicherheit**, womit zum einen gemeint ist, für eine sichere Funktions- und Wirkungsweise von KI-Anwendungen zu sorgen, zum anderen geht es aber auch darum, negative Lerneffekte zu vermeiden sowie Manipulationsmöglichkeiten durch Menschen zu unterbinden. Die Sicherstellung einer adäquaten **System-/Datenqualität sowie -integration** stellt eine weitere wichtige Aufgabe dar, weil KI-Anwendungen nur so gut sein können wie die Daten, mit denen sie trainiert werden. Daher ist dafür Sorge zu tragen, dass keine Störungen und Fehlfunktionen durch qualitativ minderwertige Daten verursacht werden. Das Leistungsvermögen von KI innerhalb von Sozialversicherungsorganisationen wird stark von der Fähigkeit abhängen, unverzerrte, relevante Daten zu sammeln und entsprechend zu nutzen. Nicht zu vernachlässigen ist die Sicherstellung der **finanziellen Machbarkeit** als weitere Herausforderung. Um ein KI-Projekt zu einem erfolgreichen Ende zu führen, bedarf es einer akkuraten Einschätzung der Gesamtkosten, die zum zu erwartenden Nutzen in einem gesunden Verhältnis zu stehen haben. Schließlich spielt die **Verfügbarkeit von KI-Expertise** eine große Rolle bei der erfolgreichen Umsetzung von KI-Projekten. Sozialversicherungsorganisationen benötigen geeignete Strategien, um angesichts einer global hohen Nachfrage nach derartigen Spezialisten und traditionell höheren Gehältern in anderen Arbeitsmarktbereichen Zugriff auf ausreichend Expertise zu haben.
- **Rechtlich-regulatorische Herausforderungen:** Der verantwortungsvolle Betrieb von KI-Anwendungen mit all ihren Konsequenzen setzt die Beachtung einer Reihe von Vorschriften voraus. In dieser Hinsicht agieren Sozialversicherungsorganisationen, die sich mit dem Gedanken des Einsatzes von KI tragen, in einem Umfeld mit erheblicher Unsicherheit, da Hand in Hand mit der technologischen Entwicklung auch die rechtlichen Vorgaben im Fluss sind. Vor diesem Hintergrund besteht eine Herausforderung in der **Governance autonomer intelligenter Systeme:** Damit ist die Schwierigkeit angesprochen, die oftmals intransparenten Entscheidungs- und Handlungsweisen von KI-Systemen als Blackboxes verstehen und kontrollieren zu können. Diese Eigenart von KI zieht die Notwendigkeit nach sich, dass bestimmte Anforderungen hinsichtlich Erklärbarkeit, Transparenz, Fairness und Rechenschaftspflicht erfüllt und Governance-Mechanismen eingeführt werden, die Risiken und potenzielle Fallstricke minimie-

¹⁰ In ihrer Studie nehmen Wirtz et al. (2019) den öffentlichen Sektor in seiner Gesamtheit in den Fokus. Es darf angenommen werden, dass für den Bereich der sozialen Sicherung als Teil des öffentlichen Sektors eben diese Herausforderungen gleichermaßen wirksam sind.

ren. Mit Governance eng verknüpft ist die Frage nach der **Verantwortung und Haftung**, womit gemeint ist, wer für die Entscheidungs- und Handlungsweise einer KI-Anwendung rechtlich verantwortlich ist und im Schadensfall haftet. Im Hinblick auf das Problem der „Verantwortungslücke“ (siehe Kap. 4), die durch Maschinen entstehen kann, deren Verhalten nicht umfassend vorhersehbar und kontrollierbar ist, ist daher zu klären, ob im speziellen Fall das System so gestaltet werden kann¹¹, dass Verantwortlichkeit und Haftbarkeit klar zugewiesen werden können oder aber eine Verantwortungslücke akzeptabel ist. Darüber hinaus sind **Datenschutz und die Sicherung der Privatsphäre** zu beachten. Dies bedingt einerseits einen angemessenen Umgang mit Daten durch die Sozialversicherungsorganisationen selbst und zum anderen sind entsprechende Vorkehrungen betreffend Cybersicherheit zu treffen.

- **Ethische Herausforderungen:** KI-Ethik ist ein in der Öffentlichkeit rege diskutiertes Thema und wirft angesichts der Sensibilität des Einsatzbereichs gerade für Sozialversicherungsorganisationen gewichtige Fragen auf. Hierbei geht es zum einen um Überlegungen im Hinblick darauf, ob die Entwicklung und Nutzung von KI in bestimmten Bereichen aus ethischer Perspektive vertretbar ist und zum anderen stellen sich Fragen, inwiefern ethische Prinzipien in einer Weise in KI-Systeme eingeschrieben werden können, damit diese moralisch handeln. Konkret betreffen die ethischen Herausforderungen etwa die **KI-basierte Regelsetzung für menschliches Verhalten**, d. h. die Auswirkungen von KI-Entscheidungen auf den Menschen. Es besteht stets die Gefahr, dass KI Fehlentscheidungen trifft, weil beispielsweise die Datenbasis verzerrt oder bereits vorurteilsbeladen ist. Es ist aber auch möglich, dass KI-Entscheidungen aus menschlicher Perspektive wie ein Fehlurteil wirken, weil streng rationales Entscheiden, worauf Algorithmen getrimmt sind, schlicht nicht menschlich ist. Menschen wägen zumeist nicht streng rational ab, vielmehr spielen bei Entscheidungen auch andere Faktoren wie etwa Emotionen eine Rolle. Vor diesem Hintergrund stellt sich die Frage, ob es ethisch gerechtfertigt ist, KI ohne normativ-ethische Grundlage entscheiden zu lassen, wenn dadurch menschliche Handlungsmöglichkeiten eingeschränkt werden. Mit diesem Problem eng verknüpft ist der Aspekt der **Kompatibilität zwischen maschinellem und menschlichem Werturteil**. Weil die ethischen Prinzipien, die menschlichen Entscheidungs- und Handlungsweisen zugrundeliegen, nicht klar definierbar sind und folglich den KI-Systemen nicht eingeschrieben werden können, besteht die Gefahr, dass diese im Zuge ihrer Lernprozesse ihr eigenes Wertesystem herausbilden, welches unter Umständen nicht mit dem menschlichen übereinstimmt. Das den KI-Entscheidungen zugrundeliegende Wertesystem mit dem menschlichen Wertesystem kompatibel zu machen, ist folglich eine große Herausforderung. Dass Werturteile keinen klaren Regeln fol-

¹¹ Johnson (2015, S. 714) geht davon aus, dass es stets eine Sache der von Menschen getroffenen Designentscheidungen ist, ob eine Verantwortungslücke entsteht oder nicht. Demnach können autonome Systeme so gestaltet werden, dass sie – oder Teile von ihnen – transparent und kontrollierbar sind. Weil es sich hierbei um menschliche Entscheidungen handelt, entspringen Verantwortungslücken dieser Argumentation zufolge nicht schlichtweg der Komplexität von Maschinen.

gen und nicht berechenbar sind, zeigt sich anhand von **moralischen Dilemmata**. Auch KI-Anwendungen können vor Entscheidungen stehen, die in jedem Fall negative Auswirkungen nach sich ziehen. Wenn schon Menschen in solchen Fällen nicht zu einheitlichen Entscheidungen gelangen, so stellt es eine ungleich größere Schwierigkeit dar, KI-Anwendungen eine ethische Denkfähigkeit zu vermitteln, die sie in die Lage versetzt, Entscheidungen in Situationen zu treffen, die moralisch dilemmatisch sind. In ethischer Hinsicht ist darüber hinaus auch die **KI-basierte Diskriminierung** für Sozialversicherungsorganisationen eine ernstzunehmende Herausforderung. Hierbei gilt es, durch KI-Anwendungen herbeigeführte Diskriminierungen aufzudecken und Präventionsmechanismen einzuführen, um Ungleichbehandlung und Ungerechtigkeit zu verhindern.

- **Soziale Herausforderungen:** Schließlich haben sich Sozialversicherungsorganisationen im Zuge der Einführung von KI mit gesellschaftlichen Veränderungen auseinanderzusetzen, die durch die Technologie bewirkt werden. Weil KI Einfluss auf den Alltag von Menschen nimmt, geht es zum einen um ganz konkrete Konsequenzen der Technologie, zum anderen sind auch Antworten auf diffuse Ängste zu finden, die mit Zukunftsvisionen eines zunehmenden Einflusses von Technik auf den Menschen verknüpft sind. Dass KI einen immensen Einfluss auf den Arbeitsmarkt hat, stellt eine der medial präsentesten Herausforderungen im KI-Diskurs dar. Dabei geht es um die **Substitution und Transformation von Arbeit** durch die KI-basierte Automatisierung. Neben technologischer Arbeitslosigkeit als Folge eines weit verbreiteten KI-Einsatzes ist davon auszugehen, dass sich Jobprofile ändern und Arbeitsrollen stärker überwachender Natur sein werden. Für den Erfolg von KI-Implementierungen ist **soziale Akzeptanz und Vertrauen in KI** entscheidend. Dieser Aspekt ist eng verbunden mit den zuvor beschriebenen Herausforderungen und deren Lösung. Die Herausforderungen ernst zu nehmen, ist daher der erste Schritt, die Kluft zwischen Erwartungen an und Realität der KI-Systeme gar nicht erst entstehen zu lassen. Schließlich ist ein angemessenes Verständnis der **Transformation der Mensch-Maschine- und Maschine-Maschine-Interaktion** zu entwickeln. Eine große Herausforderung besteht vor allem darin, Interaktionen intuitiv zu gestalten und für Nutzer offenzulegen, auf welche Weise die KI eine Interaktion jeweils beeinflusst. Intuitiv zu interagieren setzt voraus, dass die Maschine erkennt, welche Intentionen der jeweilige Interaktionspartner hat. Insgesamt geht es hierbei um nicht weniger, als Maschinen grundlegend beizubringen, wie die (menschliche) Welt funktioniert.

4 CHANCEN UND HERAUSFORDERUNGEN DES KI-EINSATZES IN DER SOZIALEN SICHERUNG

Gerade im Digitalisierungsdiskurs ist immer wieder zu beobachten, dass die Rechtfertigung des Einsatzes neuer Technologien weniger durch ihren Nutzen als vielmehr schlicht durch die Tatsache ihrer Neuheit erfolgt. Dies birgt die Gefahr, dass das Neue aus einem allzu positiven Blickwinkel betrachtet wird und dabei – auf dem kritischen Auge blind bleibend – wichtige Aspekte im Zuge ihrer Ausgestaltung außer Acht gelassen werden. So wenig wie jeder Anwendungsfall von KI begrüßenswert ist, bloß weil es sich um eine Innovation handelt, gilt aber auch das Gegenteil, dass nämlich jede KI-Anwendung von vornherein zu verteufeln ist. Auch dies lässt sich in der gegenwärtigen Debatte rund um KI beobachten: Die Technologie gibt Anlass zu dystopischen Zukunftsbeschreibungen und schürt so Ängste, Zweifel und große Vorbehalte.

In der öffentlichen Debatte rund um künstliche Intelligenz finden sowohl die immensen Chancen als auch die Risiken, die in der Technologie stecken, ihren Platz. Auf der einen Seite werden große Produktivitätssprünge und vollkommen neue Services in Aussicht gestellt. Auf der anderen Seite stehen Bedenken, dass die Technologie Menschen aus ihren Jobs drängt und ein Heer von Arbeitslosen schafft, die Privatsphäre bedroht und Menschen in eine zunehmende Abhängigkeit von automatisierten Prozessen treibt, die für sie intransparent und nicht nachvollziehbar sind. Wer allerdings genau hinsieht, wird schnell erkennen, dass überhöhte Hoffnungen genauso wenig angebracht sind wie es Grund zur übermäßigen Sorge gibt, weil die Wissenschaft noch weit entfernt ist, eine künstliche Superintelligenz zu erschaffen. Die aktuellen Ansätze sind immer noch als „schwache KI“ zu qualifizieren: Als intelligent gelten sie, weil es ihnen gelingt, Aufgaben zu lösen, die normalerweise menschliche Intelligenz erfordern, nicht aber, weil sie menschliche Intelligenz vollumfänglich reproduzieren können. Dennoch: Die technische Entwicklung geht weiter und mit ihr steigt die Versuchung, künstliche Intelligenz als auf einer Stufe mit menschlichem Denken und Handeln stehend zu betrachten. Der vermehrte Einsatz von KI-Anwendungen in immer mehr Lebensbereichen macht vor diesem Hintergrund eine präzise Abwägung der mit der Technologie verknüpften Chancen und Risiken erforderlich. Im Folgenden soll daher untersucht werden, welche Chancen und Risiken sich mit dem Einsatz künstlicher Intelligenz in der sozialen Sicherung verknüpfen.

Auf der positiven Seite ist zu bilanzieren, dass KI-Systeme immense Potenziale eröffnen, die verfügbaren Datenmassen zu nutzen, um administrative Prozesse zu verschlanken und zu beschleunigen, personalisierte Services zu offerieren, Versicherte zu unterstützen und evidenzbasierte Entscheidungen zu treffen. KI-Technologie kann dazu beitragen, Routineaufgaben zu automatisieren, wodurch Zeit und Ressourcen eingespart werden. Durch die Entlastung von Fachkräften von repetitiven, aufwendigen und fehleranfälligen Aufgaben werden Kapazitäten für komplexere Aufgaben sowie individuelle Interaktionen mit Versicherten frei. Dadurch eröffnet sich die Chance, zugleich dem steigenden finanziellen Druck durch Kosteneinsparungen zu begegnen sowie eine höhere Zufriedenheit der Versicherten zu erzielen. Wie einige der in Kap. 3 aufgeführten Beispiele darlegen, erlaubt KI-Einsatz überdies neue Herangehensweisen an die Aufgabenausführung im Rahmen des Leistungserbringungsprozesses und stellt damit die Möglichkeit in Aussicht, optimierte Leistungen zu erbringen. Insgesamt eröffnet KI Raum, um näher an Bürger heranzurücken.

Der sich aus der Implementierung von KI ergebenden Chance, die Aufgaben der sozialen Sicherung in insgesamt optimierter Weise auszuführen, stehen jedoch auf der anderen Seite ernstzunehmende Herausforderungen gegenüber, die aus der Natur der Technologie – auf ihrem heutigen Entwicklungsstand – resultieren. KI unterscheidet sich wesentlich von herkömmlicher, typischerweise regelbasiert programmierter Technik. KI-gestützte Systeme agieren autonom, d. h. sie treffen auf Grundlage von maschinellen Lernvorgängen eigenständig Entscheidungen und passen sich an neue Situationen an. Für die rechtliche Betrachtung ist überdies die Art und Weise der Entscheidungsfindung von Relevanz, basiert diese doch auf Korrelationen und statistischen Wahrscheinlichkeiten. Doch setzt Recht stets Kausalitäten voraus. Zudem ist das Zustandekommen von Ergebnissen ex post kaum nachvollziehbar. (vgl. Beck, 2020, S. 4)

Ein verantwortungsvoller Umgang mit KI setzt daher voraus, Bilanz zu ziehen und zu prüfen, inwiefern die Herausforderungen zu meistern sind bzw. die Möglichkeit besteht, dass Risiken oder konkret absehbare negative Begleiterscheinungen die Aktivseite unterminieren. Zusätzlich ist im Bereich der sozialen Sicherung zu gewährleisten, dass die Gestaltungsprinzipien der Sozialversicherung durch die Implementierung von künstlicher Intelligenz nicht in Frage gestellt werden. Falls Anwendungen künstlicher Intelligenz nicht adäquat gestaltet und überwacht werden, können sie bedeutsame negative Konsequenzen für Individuen sowie die Gesellschaft als Ganzes haben und insgesamt den mit dem Sozialstaatsprinzip verbundenen Auftrag untergraben, für die soziale Gerechtigkeit und die soziale Sicherheit der Bürger zu sorgen. Die im Folgenden aufgezeigten, vielfach in Wechselwirkung stehenden Risiken und ethischen Herausforderungen, die die Kehrseite der positiven Aussichten einer KI-Implementierung darstellen, bringen klar zum Ausdruck, dass Organisationen im Bereich der sozialen Sicherung gefordert sind, potenzielle, nicht intendierte Effekte eines Einsatzes dieser wirkmächtigen Technologie zu antizipieren und entsprechende Präventionsmechanismen vorzusehen. Erschwert wird diese Aufgabe durch den Umstand, dass der Einsatz von KI im Bereich der sozialen Sicherung noch in den Kinderschuhen steckt und Verantwortliche somit nicht auf eine breite Basis von Erfahrungen und Erkenntnisse zurückgreifen können sowie

außerdem spezifische Forschung zur Thematik von KI in Verbindung mit sozialer Sicherheit in sehr überschaubarem Umfang existiert.

Im Folgenden werden einige wichtige mit der Einführung von künstlicher Intelligenz in Sozialversicherungsorganisationen verknüpfte Problembereiche vertieft.

DISKRIMINIERUNG

Computertechnologie eilt der Ruf voraus, bei Ausführung von Aufgaben unvergleichlich präzise und objektiv zu sein. Folglich wurde Computern längste Zeit gerne zugeschrieben, auf der Grundlage von Fakten zu entscheiden – im Gegensatz zu Menschen, deren Entscheidungen, weil auf subjektiven Wahrnehmungen beruhend, zwangsläufig voreingenommen und fehlerhaft sein können. Ebenso sei durch computerbasierte Automatisierung die Fehlerquelle Mensch ausgeschaltet. Der Diskurs rund um Big Data beflügelt diese Auffassung weiter, weil Daten dabei gemeinhin als unverfälschte Abbilder der Realität dargestellt werden. (vgl. etwa Mayer-Schönberger & Cukier, 2013) Big Data, so sehen es boyd & Crawford (201, S. 189), beruhe auf einer „Aura der Wahrheit, der Objektivität und der Genauigkeit“, was zu entsprechenden die Technik begleitenden Narrativen führt.

Die Praxis zeigt allerdings, dass dieses Bild tatsächlich nicht aufrecht zu erhalten ist: In den Medien wird in jüngerer Vergangenheit immer wieder von Fällen berichtet, in denen im Kontext von KI Diskriminierung und Voreingenommenheit durch Algorithmen nachgewiesen wird.¹² Zurückzuführen sind diese Vorkommnisse darauf, dass schlichtweg die Annahme, dass es „Rohdaten“ gäbe, auf den Holzweg führt. (vgl. Gitelman, 2013) „Rohe“, also unverfälschte, neutrale Daten, die man nur einsammeln müsste, um ein Modell der Welt zu erhalten, existieren nicht einfach, vielmehr würden Daten, darauf weist Manovich (2001, S. 224) hin, immer generiert werden. Daten sind ohne Akte der Selektion und Interpretation nicht zu haben – von der Auswahl dessen, was überhaupt gemessen werden soll, bis hin zur Auswertung von Ergebnissen hat man es mit naturgemäß subjektiven Entscheidungen zu tun. Weil der in der Arbeit mit Modellsystemen gut bekannte Grundsatz „Garbage In, Garbage Out“ (GIGO, „Müll rein, Müll raus“) auch für die Entwicklung von KI-Modellen Gültigkeit beanspruchen darf, sind Ergebnisse erwartungsgemäß von den Trainingsdaten abhängig. Sind diese Daten unvollständig, ungenau oder reflektieren historische strukturelle Ungleichbehandlungen, so ist es nicht verwunderlich, dass das maschinelle Lernen und die hervorgebrachten Ergebnisse dadurch beeinflusst werden. Sind bereits die Daten mit einem Bias behaftet, kann dieser durch die KI nicht als solcher erkannt werden, vielmehr wird dieser erlernt. Als Technologie mit der Fähigkeit, selbst zu lernen, können KI-Systeme daher eine

¹² Immer wieder zeigt sich, dass automatisierte Systeme direkt oder indirekt diskriminierende Resultate hervorbringen. Ein bekanntes Beispiel betrifft etwa eine Recruitingsoftware, die Einstellungsprozesse effizienter gestalten sollte. Nachdem sich allerdings herausstellte, dass durch die Empfehlungen der Software Frauen diskriminiert wurden, weil ihre Lebensläufe aussortiert wurden, wurde das Programm eingestellt. (vgl. Dastin, 2018) Noble (2018) zeigt auf, dass die Repräsentation von marginalisierten Gruppen in den Google-Suchergebnissen durchweg von rassistischen und sexistischen Stereotypen geprägt ist. Häufig wird auch kritisiert, dass Gesichtserkennungstechnologie bei nicht-weißen Menschen eine viel höhere Fehlerquote aufweist. (vgl. Bacchini & Lorusso, 2019).

inhärente Befangenheit aufweisen. KI kann helfen, vieles besser und schneller zu erledigen, sie kann aber auch dazu beitragen, Vorurteile zu reproduzieren.

Es versteht sich von selbst, dass der Einsatz von KI, die diskriminierende Ergebnisse hervorbringt und damit Gerechtigkeitsprobleme aufwirft, gerade in Organisationen der sozialen Sicherung besonders kritisch zu beurteilen ist. Neben der Gefährdung der sozialen Gerechtigkeit besteht auch das Risiko von Sicherungslücken – etwa dort, wo Maßnahmen der sozialen Sicherung an eine KI-basierte Entscheidung bzw. Prognose geknüpft ist, können falsch-negative oder falsch-positive Ergebnisse weitreichende Konsequenzen haben. Die Vorgehensweise, im Bereich der Arbeitslosenversicherung von der algorithmischen Bewertung, ob jemand langzeitarbeitslos sein wird, bestimmte Unterstützungsmaßnahmen abhängig zu machen (vgl. Kap. 3.2), ist dementsprechend kritisch zu betrachten.

NACHVOLLZIEHBARKEIT UND RECHTMÄSSIGKEIT

Dass auch KI-basierte Entscheidungen einer gesetzlichen Grundlage bedürfen und den Grundsätzen der Verfahrensgerechtigkeit zu entsprechen haben, klingt zunächst nach einer Binsenweisheit. Doch gibt die Natur von KI-Entscheidungen aus mehrerlei Gründen Anlass zur Annahme, dass die Rückbindung von Entscheidungen an eine Rechtsgrundlage zunehmend schwierig ist.

Zunächst ist festzuhalten, dass es im Falle von Risikobewertungen, wie sie Gegenstand von KI-Anwendungsfällen in der Sozialversicherung sind, immer um die Berechnung von Wahrscheinlichkeiten geht. Klarerweise ist es jedoch problematisch, wenn eine Sozialversicherungsorganisation auf Grundlage von Wahrscheinlichkeiten und nicht auf der Grundlage von Tatsachen entscheidet bzw. handelt, denn das Gesetz knüpft Rechtsfolgen grundsätzlich nur an Gewissheiten. (Henman, 2020, S. 214)

Henman (2020, S. 214) weist ferner darauf hin, dass gewöhnlich menschliche Entscheider in vielen Bereichen der öffentlichen Verwaltung Urteilsvermögen und Ermessensspielraum zu nutzen haben, um eine korrekte Anwendung der Vorschriften in komplexen Situationen sicherzustellen. Mit der Automatisierung von Entscheidungsfindungsprozessen reduziert sich dieser Ermessensspielraum, wodurch fraglich wird, ob in jedem Fall die konkrete Situation von Versicherten angemessen berücksichtigt wird.¹³

Damit eng verknüpft ist die Schwierigkeit der generellen Nachvollziehbarkeit von KI-Entscheidungen. KI führt Analysen in hohem Tempo aus und erkennt Muster, die dem Menschen verborgen bleiben. Doch stellt sich die Frage, ob solche Ergebnisse immer interpretierbar und nachvollziehbar sind. Wie aus einem Input ein Output wird, d.h., die Art und Weise, in der das Ergebnis im Einzelnen zustandekam, bleibt oftmals im Unklaren – sowohl für technische Laien als auch für Experten. Gerade beim Deep Learning ist dieses sogenannte „Black-Box-Problem“ nicht zu ignorieren: Auf welchem Wege sich neuronale Netze bilden, ist

¹³ Der deutsche Gesetzgeber hat Entscheidungen, für die ein Ermessen oder ein Beurteilungsspielraum besteht, explizit in § 35a VwVfG von einer Vollautomatisierung ausgeschlossen.

per definitionem nicht festgelegt und bleibt daher für den Menschen nicht nachvollziehbar. Weil Algorithmen ihre eigenen Regeln generieren, um aus den Inputdaten ein Ergebnis zu generieren, ist der Entscheidungsprozess intransparent. Für das rechtsstaatliche Erfordernis der Rechtsgebundenheit wirft dieser Umstand ernsthafte Probleme auf.

VERANTWORTUNG, TRANSPARENZ UND ERKLÄRBARKEIT

KI-basierte Entscheidungen werfen stets die Verantwortungsfrage auf: Wer oder was trägt die Verantwortung für Entscheidungen und Handlungen bzw. die Haftung für einen eventuellen Fehler? Gerade weil aufgrund der Black-Box-Problematik KI-generierte Ergebnisse nicht immer umfassend vorhersehbar und kontrollierbar sind und infolgedessen im Schadensfall nicht zu klären ist, auf welchen der beteiligten menschlichen Akteure (z. B. Programmierer, Trainer, Betreiber) der Fehler zurückzuführen ist, ist diese Frage nicht zu vernachlässigen. In diesem Zusammenhang wird vom Vorliegen einer „Verantwortungslücke“ (Matthias, 2004) gesprochen.¹⁴

Moderne Rechts- und Verwaltungssysteme zeichnen sich dadurch aus, dass sichergestellt ist, dass menschliche Entscheidungen nachvollziehbar, transparent und anfechtbar sind. Auch im Bereich der sozialen Sicherung sind Verantwortungslücken speziell für jene problematisch, die von den Entscheidungen und Handlungen betroffen sind. Gerade im Bereich der sozialen Sicherung ist der Grundsatz der Fairness im KI-Kontext – von der Europäischen Kommission als Schutz von Personen und Gruppen vor Verzerrung, Diskriminierung und Stigmatisierung definiert – nicht geringzuschätzen. Damit geht in verfahrenstechnischer Hinsicht die Möglichkeit einher, sich gegen KI-basierte Entscheidungen wehren und einen wirksamen Rechtsbehelf einlegen zu können. Dazu müssen die für die Entscheidung verantwortliche Stelle identifizierbar und der Entscheidungsfindungsprozess erklärbar sein. (vgl. Europäische Kommission, 2019, S. 15)

Die Gefahr einer Aushöhlung von Verantwortung und menschlicher Autorschaft kann auch durch psychologische Effekte bedingt sein, vor allem durch den sogenannten Automation Bias. Weil Menschen häufig algorithmisch erzeugten Ergebnissen und automatisierten Entscheidungsprozeduren größeres Vertrauen entgegenbringen als menschlichen Entscheidungen, wird Verantwortung – zumindest

¹⁴ In der Literatur sind durchaus auch Stimmen auszumachen, die das Vorliegen einer Verantwortungslücke bestreiten – und zwar aus verschiedenen Gründen: Santoro et al. (2008) wenden ein, dass das Rechtssystem durchaus Fälle verschuldensunabhängiger Haftung kennt und die fehlende Kontrolle über das Handeln einer Maschine daher kein valides Argument darstellt, das Vorliegen einer Verantwortungslücke zu bejahen. Nagenborg et al. (2008) machen geltend, dass es eine professionelle Verantwortung gibt, aufgrund derer Entwickler von KI-Systemen haften. Ein weiterer Debattenbeitrag besteht darin, die Möglichkeit der Verantwortung durch die autonomen künstlichen Agenten selbst ins Feld zu führen. Matthias (2008, 235ff.) regt an, Maschinen als Träger von Rechten und Pflichten zu begreifen, was eine Verantwortlichkeit bestimmter lernfähiger, autonom agierender Formen von KI nach sich zöge. Auch Hellström (2013) hält es mit den fortschreitenden Lernfähigkeiten von Maschinen für angemessen, diesen selbst Verantwortung zuzuschreiben. Vor diesem Hintergrund ist auch der Vorstoß des Europäischen Parlaments zu begreifen, die Einführung der rechtlichen Kategorie einer „elektronischen Person“ zu fordern, wodurch die Möglichkeit eröffnet wird, eventuelle Schäden der Maschine selbst zuzurechnen. (vgl. Europäisches Parlament, 2017) Die von der Bundesregierung einberufene Datenethikkommission (DEK) spricht sich hingegen entschieden gegen den Gedanken aus, Systemen hoher Autonomie künftig Rechtspersönlichkeit zuzuerkennen mit dem Ziel, solche Systeme selbst haften zu lassen. (vgl. Datenethikkommission, 2019, S. 219).

unbewusst – an „Quasi-Akteure“ delegiert. Dieser Effekt kann dazu führen, dass selbst bei Systemen, deren Rolle strikt auf die bloße Entscheidungsunterstützung begrenzt ist, das KI-System allmählich in die Rolle des eigentlichen Entscheiders gerät. (vgl. Deutscher Ethikrat, 2023, S. 185)

PRIVATHEIT UND ÜBERWACHUNG

Mit der Herstellung einer adäquaten Datengrundlage ist eine weitere Herausforderung für den Einsatz von KI im Bereich der sozialen Sicherung gegeben: Insbesondere Modelle nach dem Machine-Learning-Ansatz sind auf eine große Menge von (personenbezogenen) Daten angewiesen. Diese müssen in hinreichender Qualität und Quantität vorliegen sowie in Bezug auf die betreffende Problemstellung relevant sein. Die große Menge an sensiblen Daten, mit denen im Bereich der sozialen Sicherung hantiert wird, verleiht entsprechenden Vorkehrungen im Hinblick auf Daten- und Privatsphäreschutz eine vordringliche Bedeutung. Aus datenschutzrechtlicher Sicht ergibt sich angesichts der großen Datenmassen, die von den KI-Systemen verarbeitet werden, insbesondere die Frage nach Einhaltung von Zweckbindungs- und Erforderlichkeitsprinzip: Werden die Daten nur für den Zweck verarbeitet, für den sie erhoben wurden und werden sie bloß verarbeitet, soweit die Verarbeitung für den jeweiligen Zweck erforderlich ist?

Die Erfassung von Daten und die damit einhergehende Möglichkeit der Prognoseerstellung und Mustererkennung kann Personen nicht nur, wie oben beschrieben, vulnerabel gegenüber möglichen Benachteiligungen machen, sondern überdies die Privatsphäre und Autonomie von Personen beeinträchtigen. Selbst in Fällen, in denen Auswertungen keine Rückschlüsse auf individuelle Personen zulassen, kann sich bereits die bloße Sorge vor der Möglichkeit der personenbezogenen Auswertung negativ auf Autonomie und Entfaltungsfreiheit von Menschen auswirken. In diesem Kontext beschreibt der sogenannte Chilling-Effekt (vgl. Penney, 2017) Einschränkungen für Personen, die sich aus der Sorge ergeben, dass ihr Verhalten beobachtet und analysiert wird. Der Einsatz von KI kann also einerseits negative Auswirkungen auf Freiheit und Autonomie durch die tatsächliche Überwachung von Personen haben und andererseits können auch Effekte auftreten, die als Folge der Anpassung an die Sorge vor Überwachungsmaßnahmen entstehen. (vgl. Deutscher Ethikrat, 2023, 357f.)

LEGITIMATION

Zu alledem sollte man sich einer weiteren Konsequenz der durch den KI-Einsatz immer weiter getriebenen Automatisierung im Sozialsicherungssystem bewusst sein: die Ablösung dessen, was Bovens & Zouridis (2002) eine „street-level bureaucracy“ nennen, durch eine, wie sie sagen, „system-level bureaucracy“. Die Bürokratie des rechtsstaatlich verfassten Wohlfahrtsstaats, so lässt sich diese Beobachtung zusammenfassen, verändert durch Einbringung von digitaler Technik ganz grundlegend ihr Gesicht und wirft ernstzunehmende Legitimationsprobleme auf.

Betrachtet man die Geschichte der Digitalisierung und der damit verknüpften Automatisierung im Bereich der öffentlichen Verwaltung, so lässt sich feststellen, dass die Automatisierung zunehmend systemisch erfolgt: Es werden voneinander getrennte automatisierte Systeme verknüpft, sodass stärker integrierte Prozesse

entstehen, nicht selten werden dabei sogar organisatorische Grenzen überschritten. Zugleich erstreckt sich die Automatisierung auf stets weitere Bereiche, um immer mehr solche Aufgaben zu umfassen, die traditionell der Sphäre des menschlichen Urteilsvermögens überlassen wurden. Eine das gesamte System erfassende Automatisierung setzt eine weitgehende Standardisierung und Modularisierung von Organisationsprozessen voraus. Es geht darum, Gesetze in Computercode umzusetzen¹⁵, was faktisch dazu führt, dass Code zum de-facto-Gesetz wird. (vgl. Lessig, 1999) Vorangetrieben werden solche Entwicklungen zudem durch ein Data Warehousing in großem Maßstab sowie die Verknüpfung von Datenbeständen. Auf diese Weise werden sozialstaatliche Prozesse akkurater, nahtloser und vernetzter sowie leichter zugänglich, etwa über Webseiten oder Smartphone-Apps. (vgl. Henman, 2022, S. 269)

Bovens & Zouridis (2002) beschreiben die traditionelle Bürokratie als eine solche, in der Staatsbedienstete Vorschriften und Verwaltungsroutinen auf konkrete Fälle anwenden. In der täglichen bürokratischen Praxis wurden die Vorschriften lebendig, weil es letzten Endes immer die Angestellten des öffentlichen Dienstes waren, die Entscheidungen trafen, d. h. über Anträge entschieden, die Auszahlung einer Zuwendung bewilligten, bestimmte Bedingungen an eine Genehmigung knüpften, die Arbeitsfähigkeit Arbeitsuchender beurteilten und dergleichen mehr. In der IT-basierten Bürokratie finden persönliche Interaktionen zwischen Bürger und Staatsbediensteten nicht mehr statt. Informationstechnologie übernimmt die tragende Rolle in der Arbeit der Organisationen – nicht nur im Hinblick auf die Erhebung und Sammlung von Daten, wie in den Anfangstagen der Automatisierung, sondern in der Ausführung und Steuerung des gesamten Leistungserbringungsprozesses. Den entscheidenden Unterschied zwischen den beiden bürokratischen Modellen sehen Bovens & Zouridis (2002) darin, dass sich der einst bei den die einzelnen Fälle bearbeitenden Angestellten des öffentlichen Diensts liegende Ermessensspielraum nun Systemdesignern zukommt. War Gesetzesvollzug einst die Anwendung von Vorschriften auf Einzelfälle, so geht es in einem IT-basierten Umfeld um die Gestaltung von Ausführungssystemen. Dazu müssen Gesetze in Code übersetzt werden, wodurch Systemgestaltung zu Politikgestaltung wird. Der Ermessensspielraum bezieht sich nicht mehr auf die Anwendung genereller Regeln auf individuelle Fälle, sondern auf die Gestaltung der Regeln selbst. In seiner eindringlichen Warnung vor der Digitalisierung von Sozialsystemen weist der frühere UN-Sonderberichterstatter Alston (2019) unter anderem auf das Problem hin, dass es problematisch sei, wenn der private Sektor eine führende Rolle bei der Gestaltung, dem Aufbau und sogar dem Betrieb wesentlicher Teile des digitalen Wohlfahrtsstaates übernehme.

Vor diesem Hintergrund weist Henman (2022, S. 269) darauf hin, dass mit diesem Wandel erhebliche Nachteile verbunden sind: Eine „digital by default“-Strategie schließt solche Personen von Leistungen der sozialen Sicherung aus, die keinen Zugang zu digitaler Technologie haben oder denen die Fähigkeit zu deren Nutzung fehlt, sofern nicht andere Zugangswege bereitgestellt werden. Zudem führt

¹⁵ Zu den vielfältigen Problemen im Zusammenhang mit der Übertragung der Rechtsanwendung auf Computerprogramme siehe ausführlich Effer-Uhe (2023).

die Automatisierung sozialstaatlicher Prozesse dazu, dass die Verantwortung und der Aufwand für die Bereitstellung immer größerer Mengen digitaler Informationen zunehmen. Zugleich schränkt Automatisierung die Ermessensfreiheit ein, da Informationstechnologie komplexe menschliche Realitäten und Umstände in Code zu pressen versucht, woraus Entscheidungen resultieren, die diesen Umständen dann oftmals nicht gerecht werden. Ferner gibt Henman (2022, S. 269) zu bedenken, dass fehlerhaft arbeitende Systeme systematisch Fehler begehen, was weitreichende Schäden verursacht.

UNKNOWN UNKNOWNNS UND KATEGORIENFEHLER

Selbst wenn es möglich sein sollte, sämtlichen vorgenannten Problemen Herr zu werden, so ist niemals ausgeschlossen, dass KI immer noch auf eine Weise agieren könnte, die vom Menschen nicht vorhersehbar ist. Im Zuge der Entwicklung von KI-Anwendungen wird stets stark auf deren Vorteile und die zu lösenden Probleme fokussiert, während eine Folgenabschätzung unterbleibt – nicht zuletzt, weil negative Auswirkungen nicht bekannt sind oder sich erst auf lange Sicht zeigen. Es ist es ja schon rein logisch gar nicht möglich, die „unknown unknowns“, die unbekanntes Unbekanntes, also jene Sachverhalte, von denen wir gar nicht wissen, dass wir sie nicht kennen, vorherzusehen. Modelle können blinde Flecken haben, die benutzten Trainingsdaten sind unter Umständen nicht für den Kontext geeignet, innerhalb derer die KI genutzt wird, oder aber die Fehler werden von den Entwicklern nicht verstanden. Es besteht immer die Möglichkeit, dass KI vollkommen falsche Ergebnisse liefert. Zwar lässt sich das Risiko derartiger Vorfälle durch kontinuierliches Training der Modelle mit neuen Daten reduzieren, doch selbst für ein verbessertes System wird es immer Informationen geben, die außerhalb des „Horizonts“ der KI liegen.

Wenngleich die immensen Fortschritte auf dem Gebiet der künstlichen Intelligenz unbestritten sind, wovon auch der in den letzten Jahren zunehmende praktische Einsatz der Technologie zeugt, besteht ebenso Konsens dahingehend, dass die Forschung keineswegs am Ziel ist, weil KI insofern verbesserungsfähig ist, als sie die menschliche Welt nicht wirklich „versteht“. KI-Systeme mögen heute Schach und Go spielen, Autos lenken sowie sonstige Aufgaben ausführen und Probleme lösen können, doch sind wir weit davon entfernt, über eine starke KI zu verfügen. Wenn es nämlich darum geht, eine „Artificial General Intelligence“ (AGI) zu entwickeln, also eine Künstliche allgemeine Intelligenz, die imstande ist, Wissen eines bestimmten Bereichs auf einen anderen Bereich zu übertragen und dadurch nicht bloß bestimmte einzelne Aufgaben zu lösen, sind die technischen Fortschritte recht überschaubar. Angesichts der Beobachtung, dass Entwicklungsbemühungen nach wie vor stark von der auf die Arbeiten von Turing (1937) und Church (1936) zurückgehenden Überzeugung geprägt sind, dass eine rechnerische Simulation des menschlichen Verstandes möglich sei, lässt sich aus gutem Grund mit Bishop (2021) fragen, ob wir nicht einem Kategorienfehler aufsitzen und uns bewusst machen sollten, dass Maschinen rein gar nichts „verstehen“.

So stehen die vermeldeten Fortschritte der KI-Entwicklung in starkem Kontrast zu Fehlern, die KI-Systemen immer wieder unterlaufen und diese eher dumm als intelligent aussehen lassen: Bilderkennung scheitert etwa an verrauschten Bildern oder bereits an einem einzigen veränderten Pixel; KI versagt vollends, die Nuancen der Alltagssprache zu deuten. Solcherart Fehler können nicht darüber hinwegtäuschen, dass stets eine unüberbrückbare Kluft zwischen der technischen Problemlösungsfähigkeit von Maschinen und der allgemeinen Problemlösungsfähigkeit von Menschen bleiben wird: Maschinen „verstehen“ menschliche Sprache ebensowenig wie sie Bilder auf dieselbe Weise erkennen, wie Menschen dies tun. Der Mensch kann all dies auf Grundlage eines tiefgreifenden Verständnisses der Struktur der menschlichen Lebenswelt. Bereits John Searle (1980) hat mit seinem „Chinese Room“-Gedankenexperiment gezeigt, dass sich Verstand, Verstehen und Bewusstsein nicht durch Computer erreichen lassen, die Programme ausführen – unabhängig davon, wie intelligent oder menschenähnlich das Programm die Maschine erscheinen lässt. Damit hat Searle nicht nur die These starker KI und den sogenannten Turing-Test (vgl. Turing, 1950) als Vorgehen zur Feststellung, ob eine Maschine ein dem Menschen gleichwertiges Denkvermögen habe, widerlegt. Zugleich hat er gezeigt, dass ein Computer ein Programm ausführen und regelbasiert Zeichenreihen verändern kann, ohne die Bedeutung der Zeichen zu verstehen und gefolgert, dass Syntax für Semantik nicht hinreicht.

Auch wenn natürlich bestimmte Aufgaben durch ein technisches System erfolgreich ohne tiefes menschliches Verständnis der Zusammenhänge der Welt ausgeführt werden können, so sollte man sich doch stets die Kluft im Verstehen zwischen Mensch und Maschine vergegenwärtigen. Wenn wir nämlich Maschinen komplexere Probleme zur Ausführung überlassen, ihnen Entscheidungs- und Handlungsfähigkeit zugestehen und mit ihnen interagieren, kann wirkliches menschliches Verstehen wichtig sein, das sich nicht durch Rechengänge imitieren lässt.

Es gibt keine einfachen Antworten auf all diese Fragen. Doch gibt es seit einigen Jahren eine rege Debatte rund um die Ethik der KI, bei der eines feststeht: Es braucht eine Kontrolle der KI-basierten autonom agierenden Systeme. Das Problem lässt sich grundsätzlich von zwei Seiten angehen: Zum einen sind technische Innovationen gefragt, die die Probleme der KI-Systeme beseitigen, zum anderen können Regulierungsmaßnahmen dazu beitragen, den Einsatz von KI sicherer zu gestalten, um sozial gerechte Ergebnisse zu generieren. (vgl. Henman, 2020, 216f.) Auf der technologischen Seite sind zunächst die Bestrebungen hin zu einer erklärbaren künstlichen Intelligenz (Explainable Artificial Intelligence, XAI) (vgl. etwa Kraus et al., 2021; Samek et al., 2019) zu nennen. Mit Hilfe von erklärbarer KI soll etwa das Black-Box-Problem adressiert werden, indem die unklaren Entscheidungswege der KI ausgeleuchtet werden und nachvollziehbar gemacht wird, wie die KI zu ihren Ergebnissen gelangt. Zu diesem Zweck liefern erklärbare KI-Systeme eine für Menschen interpretierbare Beschreibung der Ergebnisfindung. Durch die Herstellung von Erklärbarkeit sollen ein verantwortungsvollerer Umgang mit der KI sowie mehr Akzeptanz und Vertrauen erreicht werden, indem sichergestellt wird, dass das System erwartungsgemäß funktioniert und regulatorischen Anforderungen gerecht wird und Betroffenen einer Entscheidung die Möglichkeit verschafft

wird, das Ergebnis anzufechten. Abgesehen von der technischen Schwierigkeit, die Entscheidungsfindung maschineller Lernprozesse zu erklären, sind mit erklärbarer KI eine Reihe von Tradeoffs bzw. Herausforderungen verknüpft: Zum einen ist zu gewährleisten, dass es nicht zu Privatsphäneverletzungen kommt und zum anderen benötigt wiederum auch erklärbare KI Kontrollmechanismen. Insbesondere geht es darum, zu verhindern, dass mittels mangelhafter Erklärungen bloß der Anschein von Regelkonformität erweckt werden soll.

Die Ursache für Bias im KI-Kontext kann in den verschiedensten Phasen der KI-Entwicklung liegen – im Grunde genommen kann das Problem in der Datenbasis oder dem Algorithmus seinen Ursprung haben. Parallel dazu gibt es eine Vielzahl von Methoden, die an den einzelnen Phasen ansetzen und zum Ziel haben, Bias zu minimieren. Dementsprechend lassen sich die folgenden Methoden unterscheiden: (vgl. Siddique et al., 2024)

- Pre-processing setzt an der Datenbasis an und umfasst die entsprechende Aufbereitung der Daten, bevor diese dem Algorithmus gefüttert werden. Dadurch soll der Einfluss von verzerrten Daten auf den maschinellen Lernvorgang möglichst gering gehalten werden.
- In-processing beruht auf der Sicherstellung, dass das KI-Modell während des Trainings keine Verzerrungen entstehen lässt. Dabei wird der Algorithmus während des Trainingsvorgangs beobachtet, um Bias zu entdecken und gegebenenfalls zu korrigieren.
- Post-processing umfasst die Anpassung des KI-generierten Outputs, um zu gerechteren Ergebnissen zu gelangen.

Henman (2020, S. 216) weist allerdings richtigerweise darauf hin, dass die Frage, ob Bias vorliegt oder nicht keine triviale Ja/Nein-Frage darstellt. Menschen unterschiedlich zu behandeln, muss nicht in jedem Fall als Diskriminierung, sondern kann durchaus auch als angemessene Form der Personalisierung zu interpretieren sein. So wird man beispielsweise in der Gesundheitsversorgung bestimmte Initiativen, die sich gezielt an Frauen statt an Männer oder an alte statt an junge Menschen richten, nicht als diskriminierend auffassen, sondern darin eine angemessene Differenzierung sehen.

Grundsätzlich ist es wünschenswert, ethische Prinzipien so früh wie möglich in den KI-Entwicklungsprozess einfließen zu lassen, anstatt nur die Ergebnisse einer ethischen Bewertung zu unterziehen. Morley et al. (2020) plädieren dafür, im KI-Diskurs stärker darauf zu fokussieren, wie ethischen Gesichtspunkten in der Entwicklung von KI-Systemen besser Rechnung getragen werden kann. In diesem Zusammenhang zieht Henman (2020, S. 216) eine Parallele zum „Privacy by Design“-Konzept (vgl. Cavoukian, 2012), das fordert, Technik von Grund auf so zu gestalten, dass der Schutz der Privatsphäre garantiert ist. Analog lässt sich für die Einhaltung ethischer Prinzipien durch KI-Anwendungen fordern, proaktiv und präventiv ethische Gesichtspunkte in der Technikgestaltung mitzudenken, ein ethisch einwandfrei-

es Verhalten als Default zu setzen sowie ethische Prinzipien in das Design und die Architektur der KI-Systeme einzubetten. In ihren ETHIK-LEITLINIEN FÜR EINE VERTRAUENSWÜRDIGE KI (Europäische Kommission, 2019) weist auch die Hochrangige Expertengruppe für künstliche Intelligenz auf die konzeptuelle Integration von Werten in das System hin, um eine Brücke zwischen den abstrakten Prinzipien, an die das KI-System gebunden ist, und den spezifischen Entscheidungen bei der Umsetzung zu schlagen.

Abgesehen von technologischen Verbesserungen sind auch Governance-Maßnahmen im Gespräch, die den Einsatz von KI sicherer machen sollen. Die Akzeptanz von KI und damit die Chance, KI im Bereich der sozialen Sicherung zum gesellschaftlichen Wohl anzuwenden, wird wesentlich davon abhängen, wie effektiv es gelingt, unethische und sonst schädliche Auswirkungen des Technologieeinsatzes regulatorisch in den Griff zu bekommen. Zugleich besteht die Herausforderung darin, nicht durch eine Überregulierung zu große Hindernisse für die Nutzung und Weiterentwicklung von KI zu errichten. Vor diesem Hintergrund sind die rechtlichen Rahmenbedingungen im Hinblick auf die genannten Schwächen von KI zu überarbeiten.¹⁶ Insgesamt ist zu beobachten, dass in den zurückliegenden Jahren weltweit die Anzahl der erlassenen Gesetze mit KI-Bezug stark angestiegen ist (vgl. Maslej et al., 2023, 267ff.). Neben Regulierung kommen des Weiteren Verhaltenskodices in Betracht, um die Vertrauenswürdigkeit von KI zu stärken. Auch Maßnahmen der Standardisierung, wie etwa Normen für Konzeption, Betrieb etc., deren Einhaltung durch Gütesiegel bestätigt wird, können KI sicherer machen. Eine Zertifizierung könnte einer breiten Öffentlichkeit gegenüber bestätigen, dass ein KI-System transparent, rechenschaftspflichtig und fair ist. (vgl. Europäische Kommission, 2019, 27f.)

Vor diesem Hintergrund kommt auch solchen Bemühungen große Bedeutung zu, die die Entwicklung von praktischen Instrumenten und Methoden zur Bewertung von KI betreffen. Während weitgehend Einigkeit über die als Mindestanforderungen für KI-Anwendungen entscheidenden Prinzipien und Werte besteht, von Transparenz und Sicherheit über Datenschutz bis zu Fairness und Nichtdiskriminierung, bleibt unklar, wie die Umsetzung der ethischen Prinzipien sichergestellt werden kann. (vgl. Henman, 2020, S. 217) In diesem Zusammenhang ist etwa die Initiative des Verbands der Elektrotechnik, Elektronik, Informationstechnik (VDE) und der Bertelsmann Stiftung hervorzuheben, die ein Rahmenwerk entwickelten (vgl. Hallensleben & Hustedt, 2020), das durch Kombination von drei Werkzeugen die Messbarkeit und konkrete Umsetzbarkeit allgemeiner ethischer Prinzipien erlauben soll:

- KI-Ethik-Label: Eine dem Energieeffizienzlabel auf Elektrogeräten nachempfundene Markierung soll einen schnellen Überblick über ethisch relevante Merkmale eines KI-Systems (z. B. Verantwortlichkeit, Gerechtigkeit, Transparenz) bieten. Solcherart kompakte Informationen sollen die Kontrolle sowie Transparenz und Vergleichbarkeit von Produkten erleichtern.

- WKIO-Modell: Das sogenannte WKIO-Modell (im Englischen VCIO: Value, Criteria, Indicators, Observables) dient der Konkretisierung und Messbarmachung der im KI-Ethik-Label aufgeführten Werte, indem es Werten zu bestimmende (variable) Kriterien zuordnet, deren Bewertung anhand von Indikatoren und Observablen (messbar) erfolgt und so den Grad der jeweiligen Werterfüllung anzeigt.
- Risikomatrix: Da die ethische Beurteilung eines KI-Systems vom konkreten Anwendungsfall abhängig ist, soll anhand der Risiko-Matrix eine Klassifizierung des Anwendungskontexts möglich sein. Beurteilt werden hierbei einerseits die Intensität des potenziellen Schadens sowie andererseits die Abhängigkeit von der Entscheidung. Hierbei wird die Intensität des potenziellen Schadens als besonders hoch angenommen, wenn Auswirkungen auf Grundrechte, Gleichheit oder soziale Gerechtigkeit möglich sind, eine große Anzahl von Menschen betroffen sind sowie mit Effekten auf die Gesellschaft als Ganzes zu rechnen ist. In Zusammenhang mit dem Aspekt der sozialen Gerechtigkeit nennen die Autoren explizit Renten- und Krankenversicherung als Beispiele – nicht überraschend handelt es sich beim Anwendungskontext soziale Sicherung um einen solchen, der die Existenz eines höheren Risikos nahelegt. Die zweite Dimension, Abhängigkeit von der Entscheidung, wird dann als kritischer und folglich stärker regulierungsbedürftig gesehen, wenn keine menschliche Überwachung der KI erfolgt, keine Möglichkeit besteht, sich KI-basierten Entscheidungen, ohne gravierende Nachteile zu gewärtigen, entziehen zu können sowie kein Weg offensteht, KI-Entscheidungen zu korrigieren oder anzufechten.

Ähnliche Initiativen und Vorschläge gibt es auch außerhalb Deutschlands: So hat beispielsweise die britische Regierung 2019 einen Leitfaden für den Einsatz künstlicher Intelligenz im öffentlichen Sektor herausgegeben. (vgl. Government Digital Service & Office for Artificial Intelligence, 2019) Das Ada Lovelace Institute hat gemeinsam mit DataKind UK einen Überblick über Instrumente für die Bewertung von algorithmischen Systemen veröffentlicht. Vorgeschlagen werden Audits zur Bewertung von Bias sowie zur Einhaltung von regulatorischen Anforderungen sowie Folgenabschätzungen algorithmischer Systeme. (vgl. Ada Lovelace Institute & DataKind UK, 2020) Vom Centre for Government Excellence der Johns Hopkins University in Baltimore stammt ein Instrumentarium, das sich insbesondere an den öffentlichen Bereich wendet und dabei unterstützen soll, die Auswirkungen des KI-Einsatzes zu verstehen, die potenziellen Risiken klar zu formulieren und Möglichkeiten zu deren Abschwächung zu erkunden. (<https://ethicstoolkit.ai/>) Auch die Expertengruppe der Europäischen Kommission hat ihren abstrakten Anforderungskatalog konkretisiert und eine Checkliste erarbeitet, anhand derer überprüft werden kann, inwieweit ein KI-System der Vertrauenswürdigkeitanforderung entspricht. (vgl. High-Level Expert Group on Artificial Intelligence, 2020)

5 LEITLINIEN FÜR DEN EINSATZ VON KI IN DEN SOZIALEN SICHERUNGSSYSTEMEN

Es besteht kaum ein Zweifel daran, dass künstliche Intelligenz viele unserer Lebensbereiche grundlegend verändern wird – auch im Bereich der sozialen Sicherung sind die bereits produktiven oder zurzeit erprobten und erforschten Einsatzzwecke ein Vorgeschmack auf gravierende Wandelprozesse. Weniger zweifelsfrei steht fest, welche Auswirkungen die neue Technologie genau haben wird: Werden deren Chancen oder Gefahren überwiegen? (vgl. Kap. 4) Für eine dem gesellschaftlichen Wohl zuträgliche Antwort auf diese Frage ist die Schaffung eines geeigneten Rahmens für den Einsatz von KI-Technologie unerlässlich.

Wie so häufig im Digitalisierungsbereich hinken Regulierungsbemühungen auch in Fragen der künstlichen Intelligenz dem technologischen Fortschritt stets hinterher. Doch befördern die enormen ethischen, sozialen, politischen und rechtlichen Herausforderungen, die mit dem Einzug von KI in unsere Gesellschaft im Allgemeinen und in den Bereich der sozialen Sicherung im Besonderen aufkommen, Fragen der Regulierung in jüngster Zeit auf die Agenden vieler nationaler Regierungen, internationaler Regierungsorganisationen und anderer Institutionen, die u. a. in Form von Diskussions- und Positionspapieren zu dieser Problematik Stellung beziehen. Gemeinsamer Nenner ist dabei das unter dem Schlagwort „Vertrauenswürdige KI“ (Trusted AI) zusammenzufassende Verständnis, dass für die Akzeptanz der Technologie sowie deren rechtskonforme, sichere Anwendung die Vertrauenswürdigkeit von KI maßgeblich ist.

Mit ihrer erstmals 2018 vorgelegten und 2020 im Lichte neuer Entwicklungen fortgeschriebenen KI-Strategie zielt die Bundesregierung auf zweierlei: Zum einen geht es um die Wahrung individueller Freiheitsrechte, Autonomie, Persönlichkeitsrechte und die Entscheidungsfreiheit des Einzelnen. Zum anderen steht die Stärkung Deutschlands als Industriestandort im Fokus. Die KI-Strategie für Deutschland gibt die wesentlichen Rahmenbedingungen vor, um die Chancen der neuen Technologie im globalen Wettbewerb zum gesellschaftlichen Wohl zu nutzen.

Auf europäischer Ebene findet seit einigen Jahren eine intensive Auseinandersetzung mit der Nutzung künstlicher Intelligenz statt. Hierbei steht für die europäischen Regulierungsbehörden in erster Linie die Gewährleistung des Schutzes der Bürgerrechte, der Sicherheit und einer ethisch vertretbaren Nutzung von KI im Zentrum.

Die Europäische Datenschutz-Grundverordnung (DSGVO) bildet den datenschutzrechtlichen Rahmen innerhalb der Europäischen Union, der auch für KI-Systeme gilt. Die Verordnung sieht mit dem „Recht auf Auskunft“ ein bedeutendes Betroffenenrecht vor: Demnach haben betroffene Personen nicht nur ein Recht darauf zu erfahren, ob eine bestimmte Stelle betreffende personenbezogene Daten verarbeitet, sondern ebenso haben sie umfassend über die Art und Weise der Datenverarbeitung ins Bild gesetzt zu werden. Sofern eine KI zur automatisierten Entscheidungsfindung eingesetzt wird, sind gemäß Art. 15 Abs. 1 lit. h DSGVO „aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung für die betroffene Person“ bereitzustellen.

Mit dem AI Act haben darüber hinaus 2024 die EU-Mitgliedstaaten das weltweit erste umfassende gesetzliche Regelwerk zur Regulierung von KI verabschiedet. Entsprechend der Gratwanderung zwischen einem angemessenen Grundrechtsschutz einerseits und der Erhaltung von Freiraum für Innovationen andererseits verfolgt der AI Act einen risikobasierten Ansatz zur Schaffung eines einheitlichen, verbindlichen Rechtsrahmens für einen vertrauenswürdigen Einsatz von KI in der EU: Je höher das Risiko einer Anwendung eingeschätzt wird, desto strenger sind die Vorgaben. Dementsprechend werden vier Risikostufen unterschieden:

- **Inakzeptables Risiko:** Die Bereitstellung und Nutzung von Systemen, die ein unannehmbar hohes Risiko für die Schutzgüter des KI-Gesetzes bergen, werden künftig verboten sein. Dies betrifft beispielsweise KI-Systeme, die soziales Verhalten bewerten („Social Scoring“) oder solche, die Bürgerrechte einschränken, Handlungsmöglichkeiten beeinflussen oder den freien Willen beschneiden.
- **Hohes Risiko:** Stellt ein KI-System für die geschützten Rechte und Rechtsgüter ein zwar hohes, jedoch akzeptables Risiko dar, unterliegen Anbieter solcher Systeme der Verpflichtung, weitreichende Anforderungen zu erfüllen – sie haben etwa ein Risikomanagementsystem umzusetzen und für die Überwachung durch menschliches Personal zu sorgen. Ebenso bestehen besondere Aufzeichnungs- und Transparenzanforderungen. Beispiele für solche hochriskanten Systeme sind solche zur biometrischen Identifizierung und Kategorisierung von natürlichen Personen oder zur Auswahl von Bewerbern und Leistungsüberwachung in Arbeitsverhältnissen. Auch Systeme, die für die Verwaltung und den Betrieb kritischer Infrastrukturen eingesetzt werden, fallen in diese Risikoklasse.
- **Begrenztes Risiko:** In die dritte Kategorie fallen etwa KI-Systeme zur Interaktion mit Menschen. Dazu zählen beispielsweise Chatbots oder Anwendungen zur Emotionserkennung. Anbieter solcher Systeme unterliegen Transparenzpflichten, die in erster Linie die Offenlegung verlangen, dass es Betroffene mit einer KI und nicht mit einem anderen Menschen zu tun haben.
- **Niedriges Risiko:** Alle Systeme, die nicht einer der ersten drei Stufen zugeordnet werden können, fallen in die letzte Kategorie. Für Systeme, wie etwa KI-basierte Spamfilter oder Systeme im Bereich Predictive Maintenance, deren Risiko als gering eingeschätzt wird, sieht der AI-Act keine besonderen Anforderungen vor.

Bereits im April 2019 legte eine von der EU-Kommission eingesetzte Hochrangige Expertengruppe Leitlinien für eine vertrauenswürdige KI vor. (vgl. Europäische Kommission, 2019) Eine solche zeichnet sich definitionsgemäß dadurch aus, dass die KI rechtmäßig, ethisch und robust sein soll, d. h. mit jeglichem Einsatz von KI muss sichergestellt sein, dass geltendes Recht und ethische Grundsätze und Werte eingehalten werden sowie kein Schaden in technischer und sozialer Hinsicht verursacht wird. Aus dieser Definition werden vier ethische Grundsätze abgeleitet, die beachtet werden müssen, um die Entwicklung, Einführung und Nutzung von KI-Systemen auf vertrauenswürdige Art und Weise zu gewährleisten:

- **Achtung der menschlichen Autonomie:** Menschliche Interaktion mit KI-Systemen darf nicht die Selbstbestimmung über die eigene Person sowie die Teilhabe am demokratischen Prozess einschränken. Die Aufgabenteilung zwischen Mensch und Maschine sollte nach menschenzentrierten Entwicklungsgrundsätzen erfolgen und sinnvolle Spielräume für menschliche Entscheidungen lassen.
- **Schadensverhütung:** Negative Auswirkungen durch KI-Systeme auf den Menschen sind zu verhindern. Daher ist sicherzustellen, dass KI-Systeme sicher und geschützt operieren sowie dass sich keine ungleiche Macht- oder Informationsverteilung ergibt.
- **Fairness:** Eine faire Entwicklung, Einführung und Nutzung von KI-Systemen impliziert, dass die Verteilung von Vorteilen und Kosten gerecht erfolgt und Menschen vor unfairer Verzerrung, Diskriminierung und Stigmatisierung geschützt werden. In verfahrenstechnischer Hinsicht setzt Fairness voraus, sich gegen Entscheidungen der KI-Systeme und der sie betreibenden Personen wehren und einen wirksamen Rechtsbehelf einlegen zu können.
- **Erklärbarkeit:** Prozesse müssen transparent und Entscheidungen den betroffenen Personen erklärbar sein. Ebenso sind die Fähigkeiten und der Zweck von KI-Systemen offen zu kommunizieren.

Auf Basis dieser vier Grundsätze listet die Europäische Kommission (2019, 17ff.) sieben Anforderungen auf, die es während des gesamten Lebenszyklus eines KI-Systems kontinuierlich zu bewerten und zu berücksichtigen gilt:

- **Vorrang menschlichen Handelns und menschliche Aufsicht:** KI-Systeme sollen die menschliche Autonomie und Entscheidungsfindung unterstützen sowie die Grundrechte fördern. Damit die menschliche Autonomie nicht untergraben wird, ist für angemessene Aufsicht zu sorgen. Hierbei sind menschliche Eingriffsmöglichkeiten („Human-in-the-Loop“), Lenkungs- und Kontrollmöglichkeiten („Human-on-the-Loop“) und die Möglichkeit, den Gesamtbetrieb des KI-Systems zu beaufsichtigen sowie die Fähigkeit, zu entscheiden, wann und wie das System in einer bestimmten Situation eingesetzt werden soll („Human-in-Command“) zu gewährleisten.

- **Technische Robustheit und Sicherheit:** KI-Systeme müssen widerstandsfähig und sicher sein. Für Problemfälle müssen die Systeme über einen Notfallplan verfügen; außerdem müssen sie Sachverhalte richtig beurteilen, Ergebnisse müssen reproduzierbar und zuverlässig sein. Nur so kann sichergestellt werden, dass auch unbeabsichtigte Schäden verhindert oder zumindest minimiert werden können.
- **Schutz der Privatsphäre und Datenqualitätsmanagement:** In allen Phasen des Lebenszyklus eines KI-Systems müssen diese den Schutz der Privatsphäre und den Datenschutz gewährleisten. Da die Leistungsfähigkeit von KI-Systemen entscheidend von den eingehenden Daten abhängt, ist deren Qualität und Integrität sicherzustellen. Darüber hinaus ist der Datenzugriff, d. h. wer unter welchen Umständen auf Daten zugreifen kann, zu regeln.
- **Transparenz:** Eng mit dem Grundsatz der Erklärbarkeit ist die Anforderung der Transparenz verknüpft, die sich auf die Daten, das System und die KI-Geschäftsmodelle bezieht. Transparenz ist zum einen durch Rückverfolgbarkeitsmechanismen herzustellen, indem Datensätze und Prozesse, die zu einer Entscheidung geführt haben, dokumentiert werden. Zum zweiten sollten KI-Systeme und ihre Entscheidungen in einer Weise erläutert werden, die den betroffenen Interessengruppen angemessen ist. Zum dritten müssen KI-Systeme als solche erkennbar sein und dürfen gegenüber Nutzern nicht als Menschen auftreten.
- **Vielfalt, Nichtdiskriminierung und Fairness:** Eine vertrauenswürdige KI setzt Inklusion und Vielfalt während des gesamten Lebenszyklus des KI-Systems voraus. Unfaire Verzerrungen sind zu vermeiden sowie ein gleichberechtigter Zugang zu gewährleisten. Zur Umsetzung dieser Grundsätze empfiehlt sich die frühzeitige Beteiligung und Konsultation von Betroffenen bereits in der Entwicklungsphase von KI-Systemen.
- **Gesellschaftliches und ökologisches Wohlergehen:** Die Grundsätze der Fairness und Schadensverhütung legen nahe, die Nachhaltigkeit und ökologische Verantwortung der KI-Systeme zu fördern. Ebenso sind andere Lebewesen und soziale Auswirkungen zu berücksichtigen. Darüber hinaus sind nicht nur Auswirkungen auf den Einzelnen, sondern immer auch solche auf Institutionen, Demokratie und die Gesellschaft als Ganzes zu bedenken.
- **Rechenschaftspflicht:** Schließlich sind Vorkehrungen zu treffen, die die Verantwortung und Rechenschaftspflicht für KI-Systeme sicherstellen. Die Ermöglichung von Nachprüfbarkeit spielt hierbei eine Schlüsselrolle. Ebenso muss gewährleistet sein, dass über Ergebnisse berichtet und auf deren Folgen reagiert werden kann. Weil die oben aufgezählten Anforderungen zu wechselseitigen Spannungen führen können, sind Kompromisse unabdingbar. Ethisch vertretbare Kompromisse sind auf rationale und methodische Weise herbeizuführen. Für den Fall, dass es trotz alledem zu ungerechten und nachteiligen Folgen kommt, muss ein adäquater Rechtsschutz sichergestellt sein.

Die gegenwärtige Debatte rund um Leitlinien für eine dem gesellschaftlichen Wohl dienende KI wird außerdem von Empfehlungen und Prinzipien bereichert, wie sie beispielsweise das Gutachten der Datenethikkommission, der Abschlussbericht der Enquete-Kommission zu KI des Deutschen Bundestages, die „Hambacher Erklärung zur Künstlichen Intelligenz“ der Datenschutzkonferenz oder die Empfehlungen des Council on Artificial Intelligence der OECD beinhalten. Innerhalb dieses Diskursrahmens hat sich das „Netzwerk KI in der Arbeits- und Sozialverwaltung“¹⁷ unter Federführung des Bundesministeriums für Arbeit und Soziales (BMAS) auf gemeinsame Werte und Prinzipien für den Einsatz von KI im Bereich der Arbeits- und Sozialverwaltung verständigt. Entstanden sind selbstverpflichtende Leitlinien für den KI-Einsatz in der behördlichen Praxis, die darauf zielen, die Einführung von KI-Anwendungen zu beschleunigen und deren Qualität und Sicherheit angesichts der besonderen Verantwortung der Arbeits- und Sozialverwaltung sicherzustellen. Die Liste der grundlegenden Rechte, Werte und Prinzipien für den Einsatz von KI in der Arbeits- und Sozialverwaltung ist eng an die oben dargestellten Kernanforderungen der EU-Leitlinie für eine vertrauenswürdige KI angelehnt: (vgl. Netzwerk KI in der Arbeits- und Sozialverwaltung, 2022)

- Menschenzentrierung & Gemeinwohl
- Fairness & Nichtdiskriminierung
- Erklärbarkeit & Transparenz
- Privatsphäre & Persönlichkeitsschutz
- Sicherheit & Robustheit
- Intervenierbarkeit & Verantwortung
- Ökologische Nachhaltigkeit & Ressourcenschonung

Auch die „Deutsche Sozialversicherung Arbeitsgemeinschaft Europa e. V.“ leistet einen Beitrag zum KI-Diskurs und stellt in ihrer Replik auf das KI-Weißbuch der Europäischen Kommission (2020) fest, dass im Bereich der sozialen Sicherheit ganz besondere Voraussetzungen für den Einsatz von KI zu erfüllen sind. Zunächst wird die Forderung erhoben, auszuschließen, dass Entscheidungen über individuelle Sozialleistungsansprüche der Versicherten entscheidend von KI abhängig gemacht werden. Konkretisiert wird diese Forderung dahingehend, dass KI vor allem „nicht durch einen automatischen Bescheid anstelle eines Menschen hoheitlich handeln“ (Deutsche Sozialversicherung, 2020, S. 6) sollte. Klargestellt wird ferner, dass beim KI-Einsatz Menschen stets die Letztverantwortung für Entscheidungen zu tragen haben, weshalb KI folgerichtig in die Schranken der Entscheidungshilfe verwiesen

¹⁷ Mitgliedsorganisationen im Netzwerk KI sind die folgenden Behörden aus der Arbeits- und Sozialverwaltung: Berufsgenossenschaft der Bauwirtschaft (BG BAU), Berufsgenossenschaft Energie Textil Elektro Medienerzeugnisse (BG ETEM), Berufsgenossenschaft für Gesundheitsdienst und Wohlfahrtspflege (BGW), Berufsgenossenschaft Handel und Warenlogistik (BGHW), Berufsgenossenschaft Holz und Metall (BGHM), Berufsgenossenschaft Nahrungsmittel und Gastgewerbe (BGN), Berufsgenossenschaft Rohstoffe und chemische Industrie (BG RCi), Berufsgenossenschaft Verkehrswirtschaft Post-Logistik Telekommunikation (BG Verkehr), Bundesagentur für Arbeit (BA), Bundesamt für Soziale Sicherung (BAS), Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA), Bundesministerium für Arbeit und Soziales (BMAS), Deutsche Gesetzliche Unfallversicherung e. V. (DGUV), Deutsche Rentenversicherung Bund (DRV Bund), Deutsche Rentenversicherung Knappschaft-Bahn-See (DRV KBS), Künstlersozialkasse (KSK), Sozialversicherung für Landwirtschaft, Forsten und Gartenbau (SVLFG) (im Rahmen des Netzwerks auch für das Zusatzversorgungswerk für Arbeitnehmer in der Land- und Forstwirtschaft – ZLF VVaG – (ZLF) und die Zusatzversorgungskasse für Arbeitnehmer in der Land- und Forstwirtschaft (ZLA) handelnd), Unfallversicherung Bund und Bahn (UVB), Verwaltungs-Berufsgenossenschaft (VBG).

wird – „nur vorbereitend und unterstützend“ (Deutsche Sozialversicherung, 2020, S. 6) solle sie eingesetzt werden. Zudem wird darauf hingewiesen, dass ein KI-Einsatz die Unterscheidung zwischen unerwünschter Diskriminierung auf der einen und gewollter Differenzierung auf der anderen Seite zu meistern habe, wobei es oftmals schwierig ist, die Grenze zu ziehen, weil letztlich Kontext und gesellschaftlicher Konsens eine Rolle spielen.

Neben den an dieser Stelle genannten Institutionen äußerten sich eine weite Bandbreite weiterer Organisationen zur Frage der zu beachtenden Prinzipien beim KI-Einsatz. Fjeld et al. (2020) untersuchten den Inhalt von 36 solchen Publikationen, die Leitlinien für die ethische, in Einklang mit dem Recht stehende und gesellschaftlich nützliche KI-Nutzung umfassen und entdeckten dabei überraschende Gemeinsamkeiten, die darauf hindeuten, dass die Debatte rund um die normative Ausrichtung der KI-Nutzung bereits weitgehend zu einem Konsens gelangt ist. In die Analyse gingen Dokumente ein, die sich hinsichtlich ihrer Zielgruppe, ihres Umfangs und ihrer Detaillierungsgrades unterscheiden. Sie stammten aus unterschiedlichen Weltregionen (Lateinamerika, Ost- und Südasien, Naher Osten, Nordamerika und Europa) und weisen folglich kulturelle Unterschiede auf. Zudem wurden die Prinzipien von unterschiedlichen Akteuren verfasst: Regierungen und zwischenstaatliche Organisationen, Unternehmen, Berufsverbände, Interessengruppen und Multi-Stakeholder-Initiativen. Obwohl sich also die untersuchten Dokumente stark unterschieden, zeigte die Analyse, dass weitgehende Einigkeit bezüglich der folgenden acht Kernthemen besteht:

- **Datenschutz:** KI-Systeme haben die Privatsphäre des Einzelnen zu respektieren sowie den betroffenen Personen die Kontrolle über ihre Daten sowie mit diesen getroffenen Entscheidungen zu überlassen.
- **Rechenschaftspflicht:** Geeignete Mechanismen haben sicherzustellen, dass die Verantwortung für die Auswirkungen von KI-Systemen angemessen verteilt ist und dass adäquate Rechtsmittel zur Verfügung stehen.
- **Sicherheit und Schutz:** KI-Systeme müssen die Anforderungen erfüllen, die beabsichtigte Leistung zu erbringen sowie nicht von Unbefugten beeinträchtigt zu werden.
- **Transparenz und Erklärbarkeit:** KI-Systeme müssen auf eine Weise konzipiert und implementiert werden, dass eine Kontrolle möglich ist, etwa durch Übersetzung ihrer Operationen in verständliche Ergebnisse und die Bereitstellung von Informationen darüber, wo, wann und wie sie eingesetzt werden.
- **Fairness und Nicht-Diskriminierung:** Der Einsatz von KI-Systemen muss möglichst fair sein und die Integration fördern.
- **Menschliche Kontrolle:** Der Einsatz von KI hat dergestalt zu erfolgen, dass wichtige Entscheidungen weiterhin von Menschen überprüft werden müssen.

- **Professionelle Verantwortung:** Hiermit wird die wichtige Rolle von Personen anerkannt, die an der Entwicklung und dem Einsatz von KI-Systemen beteiligt sind, und an ihre Professionalität und Integrität appelliert um sicherzustellen, dass die entsprechenden Interessengruppen konsultiert und die langfristigen Auswirkungen geplant werden.
- **Förderung menschlicher Werte:** Es ist zu gewährleisten, dass die Ziele, denen KI gewidmet ist, sowie die Mittel, mit denen sie umgesetzt wird, mit menschlichen Grundwerten übereinstimmen und generell das Wohlergehen der Menschheit fördern.

Auffällig ist, darauf sollte an dieser Stelle hingewiesen werden, die weithin unhinterfragte Akzeptanz von Vertrauenswürdigkeit als primäres Beschreibungsmerkmal einer KI, wie dies vorrangig von der EU und im Anschluss daran auch von der deutschen Bundesregierung gefördert wird. Kritische Stimmen sind in dieser Debatte schwer ausfindig zu machen. Gegen die Bedeutung von Vertrauenswürdigkeit an sich lässt sich ja auch nur schwerlich argumentieren. Es macht sich jedoch bezahlt, genauer hinzusehen und zu fragen, ob Vertrauenswürdigkeit im vorliegenden Zusammenhang tatsächlich die passende Kategorie ist.

Zum einen darf nicht übersehen werden, dass mit der Markierung einer Technologie als „vertrauenswürdig“ nicht selten der strategische Zweck verfolgt wird, deren Anwendung zu fördern, um die ökonomischen Potentiale dieser Technologie auszuschöpfen. Der Philosoph Thomas Metzinger (2023), selbst Mitglied der High-Level Expert Group on Artificial Intelligence, die die EU-Ethikrichtlinie ausgearbeitet hat, erhebt gar den Vorwurf, mit „Trustworthy AI“ sei „ein von der Industrie erdachtes Marketing-Narrativ“ erschaffen worden. Der deutschen Bundesregierung bzw. der Europäischen Union dient dieses Attribut nicht zuletzt dazu, ein charakteristisches Alleinstellungsmerkmal im internationalen Wettbewerb zu etablieren. (vgl. Hirsch-Kreinsen & Krokowski, 2023) Vertrauenswürdigkeit soll zum Qualitätsmerkmal einer KI „made in Europe“ bzw. „made in Germany“ stilisiert und eine „vertrauenswürdige KI“ ganz bewusst als Gegenentwurf zu der auf Kommerz bzw. Kontrolle zielenden US-amerikanischen und chinesischen Konzeptualisierung von KI propagiert werden. (vgl. Beckert, 2021, S. 17)

Zum anderen stellt sich natürlich auch die Frage, inwieweit „Vertrauenswürdigkeit“ überhaupt ein geeignetes Attribut für ein technisches System sein kann. Zunächst ist es alles andere als selbstverständlich, das Vertrauenswürdigkeitskonzept auf Maschinen anzuwenden, weil keineswegs feststeht, dass ein KI-System überhaupt imstande ist, ein wirklicher Gegenstand des Vertrauens zu sein. Metzinger (2023) hält den Gedanken einer „vertrauenswürdigen KI“ gar für „begriffliche[n] Unsinn. Maschinen sind nicht vertrauenswürdig, nur Menschen können vertrauenswürdig sein“. Versteht man Vertrauen als ein soziales Phänomen, das diskursiv konstruiert und kontextabhängig ist und kontinuierlich als notwendige Basis sozialer Interaktion ausgehandelt wird, ist es tatsächlich nur schwer vorstellbar, dass Maschinen imstande sein sollen, unser Vertrauen zu verdienen. Normalerweise schreiben wir

Vertrauenswürdigkeit menschlichen Akteuren zu, von denen wir überzeugt sind, dass sie über bestimmte menschliche Eigenschaften wie Redlichkeit, Ehrlichkeit und Glaubwürdigkeit verfügen. Eine KI könnte demzufolge dann zuverlässig sein, jedoch nicht vertrauenswürdig. Allenfalls das sozioökonomische System aus Mensch und Maschine könnte die Eigenschaft aufweisen, vertrauenswürdig zu sein. (vgl. Rieder et al., 2021)

Laux et al. (2024) sehen das im AI-Act zur Anwendung kommende enge Begriffsverständnis von „Vertrauenswürdigkeit“ im Sinne von „Akzeptanz“ der Risiken von KI kritisch. Insbesondere weil letztere durch Konformitätsbewertungen von Technikexperten beurteilt wird, verwischt die Verquickung von Vertrauenswürdigkeit und Risikoakzeptanz wichtige Unterschiede. Zudem habe man sich vor Augen zu führen, dass Vertrauenswürdigkeit ein auf lange Sicht angelegtes Konzept ist, das sich aus Lernprozessen aus Erfahrungen in sozialen Beziehungen speist und einen iterativen Prozess von Kontrollen, Kommunikation und Rechenschaftspflicht erfordert, um die Vertrauenswürdigkeit zu etablieren und dauerhaft aufrechtzuerhalten. Vor diesem Hintergrund ist die Herangehensweise der EU-Kommission zu simpel, einen klaren Trennstrich zwischen akzeptablen und inakzeptablen Risiken und folglich vertrauenswürdigem und nicht vertrauenswürdigem KI-Systemen zu ziehen.

Krüger & Wilson (2023) zeigen auf, wie stark das Vertrauenswürdigkeitskonzept in einem – selbstverständlich auch durch Machtrelationen geprägten – Diskurs strapaziert wird und dabei Vertrauen instrumentalisiert und blindes Vertrauen in KI vorausgesetzt wird. Weil die Förderung der Technologie ausgemachte Sache ist und Regierungen insistieren, den Aufbau von Vertrauen in KI voranzutreiben, gehe es stärker darum, Misstrauen in der Bevölkerung als Fortschrittshindernis abzubauen als um die Frage, wie sich sichere und zuverlässige Systeme schaffen lassen, die innerhalb der Grenzen agieren, die ihnen gezogen wurden.

Auch die deutsche Bundesregierung definiert die KI-Nutzung im öffentlichen Bereich als vorrangig – nicht zuletzt, um die Verbreitung der Technologie zu fördern. Somit bildet auch der Einsatz künstlicher Intelligenz im Sozialversicherungswesen die Speerspitze der Bemühungen, den Einsatz der Technologie voranzutreiben. Vor dem Hintergrund, dass in diesem Bereich mit sensiblen Daten umgegangen wird und Fehler gravierende Auswirkungen auf Individuen sowie die Gesellschaft haben können, gilt es, die wahre Stoßrichtung der Vertrauenswürdigkeitsanforderungen ernst zu nehmen, d. h. sicherzustellen, dass KI-Einsatz stets rechtmäßig, ethisch unbedenklich und technisch robust erfolgt.

Führt man sich Bedeutung und Tragweite der sozialen Sicherung für Individuen sowie das Gemeinwesen vor Augen, müssen bei der Nutzung von KI alle oben genannten Aspekte ernsthaft bedacht werden. Die bereits bestehende Regulierung und die entworfenen Leitlinien, die einen verantwortungsvollen KI-Einsatz regeln sollen, weisen in die richtige Richtung, indem sie einen Rahmen für KI-Maßnahmen abstecken. Weitergehenderer Schutz ist allerdings dort erforderlich, wo es um die kontinuierliche Kontrolle der in der Sozialverwaltung eingesetzten KI-Anwendungen geht. Erforderlich sind – im Unterschied zu „traditionellem“ Verwaltungshan-

deln – nicht nur Möglichkeiten der Anfechtung von Einzelentscheidungen, vielmehr müssen aufgrund der Eigenart algorithmischer Entscheidungen Wege installiert werden, gegen strukturell problematische Entscheidungen vorzugehen. Dies kann etwa durch regelmäßige Reviews, die Einrichtung von Kontrollinstanzen oder dergleichen geschehen.

Die Formulierung leitender Prinzipien kann selbstverständlich immer nur der erste Schritt zu KI-Systemen sein, die verantwortungsvoll im Bereich der sozialen Sicherung einzusetzen sind. Keine geringere Bedeutung kommt der konkreten Umsetzung von KI-Systemen zu. Auch weil sich Vertrauen und Vertrauenswürdigkeit nicht auf Kommando erschaffen lassen, ist es essentiell, den Fokus stärker auf die Voraussetzungen für Vertrauen zu richten: die Gestaltung und Umsetzung der KI-Systeme. Zum einen fällt es schwer, die formulierten Prinzipien einer vertrauenswürdigen KI zu konkretisieren, d. h. die ethische Debatte handhabbar zu machen. Zum anderen, weil sich KI-Systeme von „normalen“ Softwareprojekten unterscheiden, bedarf es zur Umsetzung der Vertrauenswürdigkeit ausmachenden Prinzipien in die Praxis eines besonderen Augenmerks auf verschiedenen Ebenen: Auf der Ebene der Softwareentwicklung sind neben Verifikations- und Validierungstests spezielle Tests zur Erkennung von Bias vorzusehen. Ebenso ist den Anforderungen einer Erklärbaren KI gerecht zu werden, indem Wege vorgesehen werden, auf denen Betroffenen Erklärungen für KI-Entscheidungen gegeben werden. Auf Organisationsebene ist von Bedeutung, dass die Führungsebene sich zu den Grundsätzen einer vertrauenswürdigen KI bekennt und eine entsprechende Kultur fördert. Zudem ist für eine geeignete Personalausstattung sowie Aus- und Weiterbildung zu sorgen. Organisationsweit sind Strukturen der Aufsicht und Kontrolle zu installieren. Es sollte sichergestellt werden, dass kontinuierliche Fehleranalysen stattfinden sowie im Rahmen von internen Reviews die Leistungsfähigkeit und eventuelle Optimierungspotenziale des KI-Systems analysiert werden. Darüber hinaus ist die Einhaltung von gesetzlichen und sonstigen Normen zu gewährleisten. (vgl. Beckert, 2021, S. 20; Shneiderman, 2020)

6 FAZIT UND AUSBLICK

Zweifelsohne sind mit dem Einsatz digitaler Technologien und insbesondere künstlicher Intelligenz im Bereich der sozialen Sicherung große Potenziale verknüpft, die administrativen Prozesse effizienter zu gestalten sowie Bürgern effektiveren Sozialschutz zu gewähren. Von der Automatisierung von Verwaltungsprozessen profitieren die Sozialversicherungsorganisationen und Bürger gleichermaßen: Einerseits gelingt es durch den KI-Einsatz, Mitarbeiter bei alltäglichen Routinearbeiten zu entlasten und andererseits kommen Bürgern kürzere Bearbeitungszeiten und Entscheidungen auf gleichmäßig hohem Qualitätsniveau zugute. Insgesamt können zudem neue Leistungsangebote umgesetzt werden.

Doch ist die Einbringung der neuen, vielversprechenden Technologie kein Selbstläufer. Inwiefern KI ihren Verheißungen gerecht wird, hängt in hohem Maße von der Ausgestaltung des Systems ab. Diese wiederum ist maßgeblich dadurch bestimmt, welches Bewusstsein für die mit dem Einsatz von künstlicher Intelligenz in einem so sensiblen Bereich wie sozialer Sicherung verknüpften Risiken jene Verantwortlichen an den Tag legen, zu deren Aufgaben das Design, die Programmierung und der Betrieb des Systems zählen. Ende 2019 hat der von den Vereinten Nationen ernannte Sonderberichterstatter für extreme Armut und Menschenrechte, Philip Alston, in einem Bericht aufs Schärfste verurteilt, wie Regierungen Sozialsysteme privatisieren und automatisieren. Wie seine Warnungen vor Augen führen, ist es ein schmaler Grat zwischen einem technologisch effektivierten Sozialstaat auf der einen Seite und einer „digitalen Wohlfahrtsdystopie“ (Alston, 2019) auf der anderen Seite. Die Fallhöhe für einen demokratischen Wohlfahrtsstaat wie Deutschland ist hierbei enorm.

Es kann daher gar nicht beharrlich genug dafür plädiert werden, die Einführung von KI unter bewusster Abwägung von Chancen und Risiken zu gestalten und nicht an politischen Willensbildungsprozessen vorbei den großen Technologieunternehmen zu überlassen. Es darf nicht vergessen werden, dass KI insofern immer schon eine „politische Technologie“ ist, als ihre Entwicklung in der Vergangenheit von Regierungen gefördert wurde und sie eng mit den auf Informationsverarbeitung basierenden Aspekten Macht und Kontrolle verknüpft ist. (vgl. Zajko, 2023, S. 248) Wird eine solche Debatte über KI in der sozialen Sicherung geführt, so ist in diesem Rahmen darauf zu achten, den Fokus richtig zu setzen: So wichtig und richtig es ist, die vordergründigen Vor- und Nachteile der Technologie – von den Effizienzgewinnen und Kosteneinsparungen über die Möglichkeit der Betrugserkennung bis hin zu nie gekannten Personalisierungschancen – zu beleuchten, so darf dabei nicht aus dem Blick geraten, worum es beim Einsatz jeglicher Technik in der sozialen Sicherung letztlich geht: Wie es durch Technologie gelingen kann, das begrenzt verfügbare Budget für Sozialleistungen so einzusetzen, dass das gesellschaftliche Wohl gesteigert wird. Erinnern wir uns an die eingangs aufgeworfene Luhmannsche Auffassung von Gesellschaft als „Konglomerat aus Entscheidungen“. Ob man die damit notwendig verknüpfte Bürokratie hinnehmen müsse, fragt Luhmann und schlägt als Lösung vor, nach den Wirkungen zu fragen, die in der Umwelt eines Sys-

tems eigentlich erzielt wurden. Mit Blick auf den Einsatz künstlicher Intelligenz in Organisationen der sozialen Sicherung kann der Kreis, innerhalb dessen man nach Wirkungen sucht, nicht weit genug gezogen werden.

Die Betrachtung des Einsatzes von KI in der Sozialverwaltung lässt vielfältige Tendenzen erkennen, die in ihrer Wirkung über den engen Kreis der eigentlichen Aufgabendurchführung hinausreichen. So führt die Einführung von KI in die Sozialverwaltung zu einer gänzlich neuen Stufe der Prozessautomatisierung. Während digitale Technologien zunächst nur auf die automatisierte Ausführung von klar definierten Routinetätigkeiten fokussierten (z. B. Berechnung von Beiträgen und Ansprüchen, Anfertigung von Statistiken) weitete sich die KI-basierte Automatisierung auf neue Einsatzgebiete aus: Nicht-routinemäßige Aufgaben und Entscheidungen, die einst als Domäne des Menschen betrachtet wurden, werden zunehmend von KI übernommen. Mit maschinellem Lernen, so die Logik autonom agierender Systeme, lässt sich menschliches Urteilsvermögen ersetzen. Damit verändert sich nicht nur die Art der Aufgabenausführung innerhalb der Sozialversicherungsorganisationen, sondern werden die Aufgaben selbst neu definiert. Man denke beispielsweise nur an die neuen Möglichkeiten des Prognostizierens und Kategorisierens oder Betrugsfälle auf gänzlich neuartige, unvergleichlich wirkungsvollere Weise abzuwehren. Insgesamt ist zu erwarten, dass die Einbringung von KI in die Sozialverwaltung der digitalen Vernetzung sowie neuen Rationalitäten im Bereich der sozialen Sicherung Vorschub leisten wird, um im Sinne eines Whole-of-Government-Ansatzes die verstärkte Zusammenarbeit zwischen Verwaltungseinheiten zu ermöglichen und dadurch Wirksamkeit und Kohärenz von staatlichen Maßnahmen zu erhöhen.

Überdies legt künstliche Intelligenz den Grundstein zu vollkommen neuen Herangehensweisen in der Sozialverwaltung. Die Umsetzung der sozialen Sicherungsmaßnahmen kann etwa stärker differenziert und personalisiert erfolgen. Anstelle von universellen Einheitslösungen macht KI-basierte Automatisierung ein zielgerichteteres Eingehen auf individuelle Belange sowie die Gestaltung nuancierter Maßnahmen möglich, um der menschlichen Vielfalt besser gerecht zu werden. Dabei kommen Entscheidungen, die einst durch menschliche Ermessensspielräume charakterisiert waren, heute durch kodifizierte Algorithmen zustande. Hand in Hand mit der Algorithmizität geht eine Form von Konditionalität, bei der der Anspruch auf bestimmte Leistungen von Umständen oder Verhaltensweisen abhängig gemacht wird, die in digitalen Datenbanken erfasst sind. Im Endergebnis erhöhen sowohl Differenzierung als auch Konditionalität die Komplexität des Systems der sozialen Sicherung, was sich einschränkend auf den Zugang zu Sozialleistungen sowie auf die Möglichkeit der Zurechenbarkeit von Entscheidungen und damit Verantwortung auswirkt. Dabei ist zu erwarten, dass im Big-Data-Zeitalter die Ausgestaltung der sozialen Sicherung immer stärker datenbasiert vonstatten gehen wird. Ausgereifte Techniken der Datenanalyse sowie computerbasierte Mikrosimulationsmodelle zur Abschätzung der Folgen politischer Maßnahmen werden einen stärker zielgerichteten Zuschnitt von Maßnahmen unterstützen. Die Tragweite der durch KI eröffneten Möglichkeiten verleitet gar dazu, einen Paradigmenwechsel anzukündigen: Könnte es nicht sein, dass in einem Zukunftsszenario der sozialen

Sicherung diese stärker vom Vorsorgegedanken geprägt sein wird? Mit KI sind die technischen Möglichkeiten gegeben, um bei der Gewährung von Sozialleistungen durch rechtzeitiges Erkennen von Risiken stärker auf Prävention und frühzeitiges Eingreifen zu setzen anstatt nachträglich zu „reparieren“. Ist ein solches Vorgehen aber auch gesellschaftlich gewollt?

Bei aller Zukunftsmusik darf nicht vergessen werden, dass das hartnäckige Narrativ, wonach datenbasierte, algorithmische Entscheidungen stets genau, richtig und objektiv sind, mit ein Grund ist, weshalb eine kritische Würdigung des Einsatzes KI-basierter Anwendungen schwer fällt. Weil falsche Entscheidungen im Bereich des Sozialschutzes allerdings besonders empfindliche Konsequenzen für Einzelne, aber auch die Gesellschaft als Ganzes haben können, ist die Zurückhaltung mit Kritik an dieser Stelle besonders schwerwiegend, führt sie doch zu einer Aufweichung von Verantwortung. Zusätzlich zu den ohnehin der KI-Technologie inhärenten Transparenzproblemen könnte der KI-fundierte Sozialstaat auf eine Situation zulaufen, die sich in Abwandlung von Michael Lipskys (1984) Schlagwort der „bürokratischen Entmündigung“ als „algorithmische Entmündigung“ bezeichnen ließe. Kannte der Wohlfahrtsstaat Lipsky (1984) zufolge immer schon Mechanismen, durch fragwürdige oder unterlassene Entscheidungen Ansprüche zuzubilligen, ohne jene Verantwortlichkeiten daran zu binden, die überhaupt erst den Anknüpfungspunkt für eine umfassende Diskussion kritischer Verteilungsfragen bilden, so stellt sich dieses Problem in der KI-basierten Sozialverwaltung mit nie gekannter Brisanz dar. Den mittels künstlicher Intelligenz getroffenen Entscheidungen fehlt in der Regel gleich auf drei Ebenen Nachvollziehbarkeit und Transparenz: Weder verfügt der Betroffene über das Wissen um die Funktionsweise des algorithmischen Systems, noch ist das verarbeitete Datenmaterial bekannt, noch ist klar, ob in spezifischen Fällen ein KI-System zum Einsatz kommt oder nicht.

Wird das Sozialstaatsprinzip ernst genommen, müssen soziale Gerechtigkeit und Rechtsstaatlichkeit gleichermaßen verfolgt werden. Aus diesem Grund ist das Wissen darum, wann und wie es zu KI-basierten Entscheidungen kommt, von so enormer Bedeutung. Gerade im Bereich der sozialen Sicherung ist es wichtig, eine breit angelegte Debatte über die Rahmenbedingungen solch weitreichender Technologieeinführungen zu führen. Naturgemäß muss dies eine interdisziplinäre Debatte sein: Auch wenn KI-Anwendungen immer nur so gut sein können, wie ihre technische Umsetzung dies erlaubt, so ist der Einsatz künstlicher Intelligenz kein rein technisches Problem. Um die Auswirkungen dieser wirkmächtigen Technologie wirklich zu verstehen und in die gesellschaftlich gewünschten Bahnen zu lenken, braucht es eine umfassendere Perspektive: Zum einen ist ins Auge zu fassen, auf welche Weise Algorithmen Anspruchsberechtigungen bestimmen oder den Zugang zu Sozialleistungen einschränken. Es darf nicht darüber hinweggesehen werden, dass Artefakten Aufforderungscharakter zukommt und unsere dingliche Umwelt als Präfiguration aufzufassen ist, bestimmte Handlungen herauszufordern und andere zu hemmen.

Damit zusammenhängend ist die Einsicht wichtig, dass KI-Anwendungen kein weltanschaulich neutrales Werkzeug sind und folglich nicht als solches behandelt werden dürfen. Weil vom Menschen geschaffen, ist ihnen – bewusst oder unbewusst – Denkweise und Weltvorstellung ihrer Schöpfer eingebettet. Ebenso ist zu berücksichtigen, dass mit der Art der gesammelten Daten jeweils stets ein ganz bestimmter Ausschnitt der Welt ins Blickfeld gerückt wird und mittels Datenbankmanagement, Data Mining oder maschinellem Lernen eine Strukturierung dieser Daten stattfindet. Die gewählten Ausschnitte und Strukturierungen folgen stets bestimmten Zwecken und resultieren in spezifischen Ergebnissen. Somit kommt Algorithmen soziale Macht zu (vgl. Beer, 2016), sie haben Auswirkungen in der sozialen Welt. Damit verknüpft sich im Hinblick auf die Anwendung von algorithmischer Entscheidungsfindung im Bereich der Sozialverwaltung in ganz besonderer Weise das Erfordernis, die zugrunde gelegten Zielsetzungen und Annahmen der vorgenommenen Strukturierungen zu identifizieren und klar zu benennen. Ein kritischer Blick auf die KI-Nutzung hat sich nicht zuletzt auf herrschende Machtverhältnisse zu richten. Im Anschluss an Foucaults Gouvernementalitätsbegriff fasst Rouvroy (2020) die Idee einer „algorithmischen Gouvernementalität“ als die Gestaltung der sozialen Welt auf, die auf der algorithmischen Verarbeitung großer Datensätze beruht und nicht auf politischen Entscheidungen. Anders gesagt: Was Ausfluss eines politischen Prozesses sein sollte, wird durch Algorithmen berechnet.

Es besteht kaum ein Zweifel, dass künstliche Intelligenz das Potenzial hat, soziale Sicherung disruptiv zu verändern. In welchen Fällen und auf welche Weise sollen die Möglichkeiten der Technologie tatsächlich genutzt werden? In welchen Fällen sollten auf die neuen Möglichkeiten besser verzichtet werden? Wie werden im konkreten Kontext deren Gefahren und Risiken bewertet? Wo sind die Grenzen der Nutzung künstlicher Intelligenz – abgesehen von der technischen Machbarkeit – zu ziehen? Auf dieses Konvolut an Fragen gilt es in einem politischen Aushandlungsprozess Antworten zu finden. Spielkamp (2019) ist in seiner Argumentation uneingeschränkt beizupflichten, dass es vorrangig eine politische Frage ist, ob Bürger von der KI-Nutzung im öffentlichen Sektor profitieren. Darüber hinaus erscheint es empfehlenswert, Fragen der Nutzung künstlicher Intelligenz in der sozialen Sicherung im Zusammenhang einer umfassenderen Politik der sozialen Gerechtigkeit zu erörtern, anstatt primär auf ökonomische Ziele zu fokussieren. Gerade bei einem für das gesellschaftliche Wohl derart zentralen Bereich wie der sozialen Sicherung gilt es, nach gesellschaftlich wünschenswerten Lösungen zu streben, anstatt das Feld einem Technikeinsatz zu überlassen, der schlicht durch die Auswirkungen des Umstands charakterisiert ist, dass die Entwicklung künstlicher Intelligenz stets derer Regulierung vorausseilt.

7 INTERVIEWS: PERSPEKTIVEN AUF DIE KI-NUTZUNG IN DER SOZIALVERSICHERUNG

7.1 ELLEN HELLMANN (BG BAU) Referentin Digitalisierung, Stabsabteilung Digitalisierung und Unternehmensentwicklung, BG BAU – Berufsgenossenschaft der Bauwirtschaft

Die Berufsgenossenschaft der Bauwirtschaft (BG BAU) setzt künstliche Intelligenz (KI) ein, um den Beratungsbedarf ihrer Mitgliedsunternehmen vorherzusagen. Worin bestand Ihre Motivation, KI für diese Aufgabe zu nutzen?

EH: Die Motivation zu diesem Projekt war tatsächlich vielschichtig. Das Projekt geht ursprünglich zurück auf die Mitarbeit der BG BAU in dem vom Bundesministerium für Arbeit und Soziales (BMAS) ins Leben gerufenen Netzwerk „KI in der Arbeits- und Sozialverwaltung“ Ende 2022. In diesem Rahmen haben wir KI-Leitlinien entwickelt, die den verantwortungsvollen Einsatz von KI in der Arbeits- und Sozialverwaltung regeln sollen. In dieses Werk floss zwar neben wissenschaftlicher auch praktische Expertise ein, die Leitlinien waren zu diesem Zeitpunkt jedoch noch nie wirklich in der Praxis angewandt worden. Das BMAS suchte vor diesem Hintergrund Träger, die die Leitlinien anhand eines echten Projekts erproben. Dafür wurde eine finanzielle Unterstützung in Aussicht gestellt. Da wir ohnehin KI einsetzen wollten, gab dies den konkreten Anstoß zum Projekt. Wir überlegten damals bereits, ob KI uns bei unseren im SGB VII geregelten hoheitlichen Aufgaben unterstützen kann.

Die BG BAU war also auf der Suche nach konkreten Anwendungsfällen für den Einsatz von KI?

EH: Unsere wichtigste Aufgabe ist die Prävention, also die Verhütung von Arbeitsunfällen und Berufskrankheiten. In diesem Zusammenhang überlegten wir, welches Vorhaben sich eignen könnte, KI einzusetzen und zu erproben. Es war uns klar, dass wir nicht einfach nur einen Anwendungsfall finden wollten, nur um KI einzusetzen, also als Selbstzweck. Vielmehr ging es uns darum, einen Use Case zu finden, mit dem wir auch unsere Ziele verfolgen können. Also haben wir geschaut, in welchen Bereichen aktuell unsere größten Herausforderungen liegen, und sind so auf unsere Aufsichtspersonen gekommen. Diese sind auf den Baustellen oder anderen baunahen Arbeitsstätten unterwegs und stehen unseren Versicherten- und Mitgliedsunternehmen im Hinblick auf Arbeits- und Gesundheitsschutz hilfreich beratend zur Seite. Hierbei stellte sich die Frage: Wie schaffen wir es eigentlich, dass wir unsere Aufsichtspersonen zu denjenigen Unternehmen schicken, die unsere Beratung am dringendsten benötigen?

Auf welche Weise erhofften Sie sich bei dieser Fragestellung Unterstützung durch KI zu erhalten?

EH: Unseren ca. 500 Aufsichtspersonen stehen deutschlandweit über 600.000 Mitgliedsunternehmen und mehr als drei Millionen Versicherte gegenüber. Für das Jahr 2022 zeigen unsere Zahlen, dass intensive Beratungsgespräche, bei denen die Aufsichtsperson mit dem Unternehmer den aktuellen Arbeits- und Gesundheitsschutz komplett durchsprechen und Verbesserungsvorschläge machen kann, nur bei zehn Prozent unserer Mitgliedsunternehmen stattfinden. Aufgrund unserer begrenzten Kapazitäten kamen wir zu der Fragestellung, wie KI dabei unterstützen kann, auf Basis unserer Daten den Aufsichtspersonen eine Empfehlung zu geben, welche Unternehmen sie optimalerweise aufsuchen und beraten sollten, um Unfälle zu verhindern.

Es ist in der Präventionsarbeit grundsätzlich schwierig, den Erfolg im Nachhinein zu beurteilen. Können Sie bereits irgendwelche Aussagen dahingehend treffen, ob sich der KI-Einsatz lohnt?

EH: Ja, richtig, auch vor Einführung der KI war die Messung, wann Präventionsarbeit überhaupt wirksam ist, schon immer eine Herausforderung. Ob ein Unfall nicht passiert ist, weil eine Aufsichtsperson vor Ort war oder ob er auch sonst nicht passiert wäre, lässt sich natürlich nicht feststellen. Was wir aber machen können, ist, die Trefferquoten zu vergleichen, d. h. wir vergleichen die Empfehlung der KI mit dem, was die Aufsichtsperson dann vorfindet. Anhand von Checklisten wird für jedes Unternehmen ein Score-Wert ermittelt und dieser kann mit dem durch die KI berechneten Score-Wert verglichen werden. Das KI-System wurde Ende 2023 eingeführt und wir haben diesen Vergleich bereits in der ersten Testphase durchgeführt, wobei sich gezeigt hat, dass sich die Erfolgsquote fast verdoppelt hat: Im Jahr 2022, also ohne KI, haben Aufsichtspersonen in 35 Prozent der Fälle jene Unternehmen aufgesucht, die dann auch tatsächlich einen roten Score hatten, also Mängel im Arbeitsschutz. Wenn man sich heute die Empfehlungen der KI anguckt, kommt man auf 58 Prozent. Richtig interessant wird es aber, wenn man die beiden Intelligenzen kombiniert und die Erfolgsquote für den Fall betrachtet, dass ein Mensch die KI-Empfehlung kontrolliert – dann kommen wir auf eine Trefferquote von 64 Prozent. Schon aus diesem Grund ist uns wichtig, dass letztendlich immer die Aufsichtsperson entscheidet, die KI empfiehlt nur. Ich hatte ja die KI-Leitlinien erwähnt und dabei gilt als ein Credo, dass die letzte Entscheidung beim Menschen bleibt. Hierbei ist auch der Aspekt von Bedeutung, dass dadurch die Akzeptanz von KI gesteigert wird – bei denen, die entscheiden sollen, und bei denen, über die entschieden wird.

Gibt es Pläne, die Anwendung auch anderweitig einzusetzen?

EH: Ja, das konkrete Modell ist auch für andere Berufsgenossenschaften sehr interessant, aber auch für die Landesaufsichten, für die ebenfalls Aufsichtspersonen tätig sind. Wir haben darüber hinaus auch Anfragen aus der Wirtschaft erhalten. Das System kann all jene unterstützen, die selbst Bauherren sind und auf Baustellen mit Sicherheitsfachkräften – ähnlich unseren Aufsichtspersonen – agieren. Es gibt sehr viele, die genau vor derselben Herausforderung stehen, nämlich dort zu sein, wo man am dringendsten gebraucht wird. Aus Datenschutzgründen gibt es Einschränkungen, wie umfangreich sich das System zur Verfügung stellen lässt. Aber auch innerhalb der BG BAU sind wir dabei, andere KI-Modelle für weitere Anwendungsfälle umzusetzen. Wir haben ja nicht nur das KI-System entwickelt, sondern auch eine KI-Plattform aufgebaut. Auf diese Weise wollen wir das Thema skalieren.

Welche Anwendungsfälle schweben Ihnen hier etwa vor?

EH: Aktuell setzen wir zusammen mit anderen Berufsgenossenschaften eine generative Besichtigungsassistenz um. Das bedeutet, dass Aufsichtspersonen Unterstützung erhalten, wenn sie bei einer Besichtigung, etwa auf einer Baustelle, Mängel entdecken – also etwa: Helme werden nicht getragen, die Absturzsicherung fehlt und dergleichen. In solchen Fällen ist es immer so, dass ein Bericht verfasst und an die betreffenden Unternehmen geschickt wird. Zur Beschleunigung dieses Vorgangs soll eine KI die von der Aufsichtsperson angesprochenen Mängel mittels Spracherkennung in Text umwandeln und aus dem bestehenden ca. 1.800 Mängel umfassenden Mängelkatalog die mit der höchsten Wahrscheinlichkeit zutreffenden Mängel vorschlagen. Die Aufsichtsperson muss dann nur noch den passenden Mangel anhaken. Auf diese Weise wird nicht nur Zeit gespart, sondern auch die Qualität der Besichtigungsberichte verbessert.

Was sind aus Ihrer Sicht die größten Hürden bei der Durchführung Ihrer KI-Projekte?

EH: Als erstes kommt mir bei dieser Frage das Thema Datenschutz in den Sinn. Die gesetzlichen Rahmenbedingungen sind in Bezug auf KI zumeist nicht explizit, es muss viel Subsumtionsarbeit geleistet werden. Daher besteht große Rechtsunsicherheit, die viel Abstimmung erfordert, was sehr zeitaufwendig ist. Das ist eine echte Herausforderung, weil nicht jedes KI-Projekt gleich ist und jeweils andere Daten, andere Anwendergruppen etc. betroffen sind. Das Gleiche gilt für das Thema Cloud. Wir machen alles On-Premise, aber das System wäre viel skalierbarer und kostengünstiger, wenn man Cloud-Produkte einsetzen könnte. Eine dritte Herausforderung besteht darin, dass viele Beschäftigte noch keinen Kontakt mit dem Thema KI hatten, damit nicht viel anfangen können oder auch Sorgen und Ängste haben – es fehlt eine Art Grundbildung. Alle mitzunehmen und aufzuklären, ist sehr zeitintensiv.

Sie sprachen bereits kurz die Akzeptanz von KI an. Besteht hier nicht die Gefahr, mit zunehmender Gewöhnung an die Technologie in eine Zwickmühle zu geraten, wenn einerseits die Sorgen und Ängste abnehmen, andererseits damit aber auch zu befürchten ist, dass Menschen Maschinen zunehmend blindlings vertrauen?

EH: Grundsätzlich besteht natürlich die Gefahr, dass Menschen gar nicht mehr selbst nachdenken und entscheiden, sondern einfach der KI folgen. Wir machen den Beschäftigten der BG BAU klar, dass wir nicht davon ausgehen, dass die Maschine Recht hat und immer gilt, was sie sagt. Im Gegenteil: wir betonen, dass die Maschine nur empfiehlt und jede Aufsichtsperson die Entscheidungspflicht hat. Dadurch suggerieren wir gerade auch, dass die Maschine nicht immer Recht hat, sondern wir eher darauf vertrauen, wie der Mensch entscheidet.

7.2 JENS KLOSTERMANN UND MANUELA MAKIOLA (DRV BUND)

JENS KLOSTERMANN

Bereichsleitender vom Bereich Verfahrensmanagement
der Zentralen Zulagenstelle für Altersvermögen,
Deutsche Rentenversicherung Bund (DRV Bund)

MANUELA MAKIOLA

Bereichsleitende vom Bereich Öffentlichkeitsarbeit
der Zentralen Zulagenstelle für Altersvermögen (ZfA),
Deutsche Rentenversicherung Bund (DRV Bund)

Es ist kaum zu übersehen, dass künstliche Intelligenz (KI) weltweit und in Deutschland zurzeit ein sehr angesagtes Thema ist. Es ist medial überaus präsent, Unternehmen suchen nach Anwendungsfällen und auch vom öffentlichen Sektor wird die Thematik stark vorangetrieben. Welche Überlegungen gibt es innerhalb der Deutschen Rentenversicherung Bund im Hinblick auf die Nutzung von KI?

JK: Künstliche Intelligenz stellt uns vor die Herausforderung einer Gratwanderung: Auf der einen Seite erkennen wir, dass KI einen bestimmten Nutzen entfalten kann. Auf der anderen Seite gibt es – insbesondere für uns als Deutsche Rentenversicherung Bund – eine ganze Reihe von Restriktionen beim Einsatz dieser Technologie. Es gibt einerseits technische Abwägungen, andererseits sind rechtliche Anforderungen, man denke nur an das Datenschutzrecht, zu beachten.

Auf der von der Zentralen Zulagenstelle für Altersvorsorgevermögen (ZfA) betreuten Internetseite zur staatlich geförderten, privaten Altersvorsorge (Riester-Förderung) steht ein Chatbot zur Beantwortung von Fragen rund um die Riester-Förderung zur Verfügung. Wie funktioniert dieser?

MM: Der Chatbot basiert auf den FAQs, die man früher auf der Internetseite fand. Diese haben wir entsprechend überarbeitet und erweitern sie auch kontinuierlich: Wenn wir feststellen, dass bestimmte Themen regelmäßig gefragt werden und es keine diesbezüglichen Antworten gibt, dann füllen wir diese Lücke schnellstmöglich. Insofern steht hier immer der Nutzende mit seinem Bedarf im Mittelpunkt. Daher achten wir auch auf eine möglichst einfache Sprache. Auch in dieser Hinsicht stehen wir als Deutsche Rentenversicherung Bund vor besonderen Herausforderungen: Nutzenden wollen wir verständliche und möglichst nicht zu lange Informationen zur Verfügung stellen. Zugleich muss die Information rechtlich korrekt sein. In manchen Fällen stellt es sich daher auch als vorteilhaft heraus, Informationen wegzulassen und den Nutzenden lieber eine Nachfrage stellen zu lassen, die dann konkret beantwortet wird. Außerdem kontrollieren wir laufend, ob der Chatbot erwartungsgemäß antwortet und nehmen erforderlichenfalls Anpassungen vor. Nutzende haben auch die Möglichkeit, Feedback zu geben: Zum einen kann die Unter-

stützung durch den Chatbot mittels Skala – hilft mir bzw. hilft mir nicht – bewertet werden und zum anderen können Kommentare hinterlassen werden. All dies führt dazu, dass wir den Chatbot ständig für den Nutzenden verbessern können.

Bei Chatbots denkt man heute schnell an KI. War es in Ihrem Fall eine bewusste Entscheidung, den Chatbot ohne KI zu realisieren?

JK: Wenn man sich an die Steuererklärung mittels früher Versionen vom ELSTER-Formular zurückerinnert: Dort gab es neben den Formularen eine Spalte mit Erklärungen. So etwas ist für Anwender sehr nützlich und in technischer Hinsicht gar nicht aufwändig umzusetzen. Wenn die Unterstützung dann noch etwas zielorientierter ausfiele, wären kontextorientierte Angebote in bestimmten Anwendungsfällen sogar hilfreicher als ein abstrakter Chatbot. Zudem darf man nicht vergessen, dass es bei KI-Lösungen immer unzählige rechtliche Rahmenbedingungen zu berücksichtigen gibt. Vor diesem Hintergrund ist unsere Herangehensweise, verschiedene Szenarien zu betrachten und uns die Frage zu stellen, was man möglicherweise mit KI tun kann oder gegebenenfalls unterlässt – all dies ist im Moment noch sehr offen. Wir sind gerade dabei, mit einem Large Language Model zu experimentieren. Derzeit warten wir auf Ergebnisse und werden dann entscheiden, inwiefern dies zu unseren Kundenanforderungen passt.

Man muss sich auch vor Augen halten, dass technische Weichenstellungen weitreichende Konsequenzen haben. Setze ich KI in der Cloud oder On-Premises ein? Welche Implikationen hat das für den Datenschutz? Wird KI in der Cloud eingesetzt, dann ist ein Zugriff auf persönliche Daten ausgeschlossen. Das bedeutet aber, dass KI bestimmte Fragen nicht beantworten kann: Habe ich Wartezeiten für Renten erfüllt? Habe ich das Mindestkapital erreicht, um eine Entnahme vornehmen zu können? Perspektivisch wären solche Informationsangebote bei anderer technischer Verwirklichung aber möglicherweise denkbar.

Auch wenn wir den Chatbot ohne KI realisierten, so befassen wir uns mit dieser Technologie aber durchaus in anderen Anwendungsbereichen.

In welchen Bereichen schwebt Ihnen vor, mit KI zu arbeiten?

JK: Grundsätzlich gehen wir stark experimentell vor: Aus unseren Erprobungen leiten wir Nutzen und Hypothesen ab, um fokussiert auf das Thema Altersvorsorge zum Schluss zu kommen, ja, an dieser Stelle lohnt sich der Einsatz von KI. Es muss dann eben nicht der Chatbot sein, es könnte auch der Telefonservice sein – und sei es nur, indem die KI das Anliegen der anrufenden Person erkennt und den Anruf an die richtige Stelle leitet.

Was konkrete Anwendungen betrifft, so ist ja in Bezug auf Informationsbereitstellung der Chatbot als Dienstleistung für Kunden nur die eine Seite. Wir haben aber natürlich auch Informationen für Mitarbeitende. Wenn man an das Sozialgesetzbuch und dazu noch die Kommentare denkt, sind das immense zu bewältigende Informationsmassen.

MM: In Bezug auf die Riester-Rente ist auch noch Steuerrecht relevant. All diese Rechtsgebiete müssen die Kolleg:innen kennen, um Anfragen korrekt zu beantworten. Hier können die Mitarbeitenden mit künstlicher Intelligenz unterstützt werden.

JK: Dabei macht sich die Arbeit mit dem Chatbot bezahlt, weil wir damit gute Erfahrungen sammeln können. Wir können nachvollziehen, woher welche Antworten kommen, und erkennen, weshalb eine Antwort falsch oder nicht optimal war. Natürlich ist der Einsatz von KI immer auch eine Frage der Kosten. Insgesamt fragen wir immer danach, wo die Nutzung von KI sinnvoll ist und wo nicht und welche Qualität und Verlässlichkeit die durch KI erzielten Ergebnisse haben.

Sie sprechen damit wahrscheinlich die sogenannte Blackbox-Problematik an...

JK: Das ist tatsächlich ein kritischer Punkt – möglicherweise weniger bei einem Chatbot, aber bei automatisierten Verwaltungsentscheidungen ist die Herleitung schon wesentlich, wie die Entscheidung zustande kam. Stellt man ChatGPT eine Frage, dann erhält man eine Antwort; stelle ich morgen dieselbe Frage, erhalte ich möglicherweise eine andere Antwort und niemand kann sagen, worauf diese zurückzuführen ist. Zwar gibt es technische Möglichkeiten, Erklärbarkeit herzustellen, aber dadurch wird die Umsetzung eines Systems insgesamt auch nicht einfacher.

Wird der Chatbot von Nutzern angenommen? Wie ist die Akzeptanz?

MM: Aktuell erhalten wir positives Feedback zum Chatbot. Es gab auch mal Überlegungen, ihn menschlicher zu gestalten. Wir haben uns aber dafür entschieden, eine klare Abgrenzung zur Kommunikation mit Menschen zu machen: Der Chatbot erscheint daher nur in Form des Logos und hat auch keinen Namen. Auch im Vergleich zu anderen Chatbots bleibt dieser eher unauffällig, springt nicht sofort ins Auge, hält sich dezent zurück, ist unten rechts am Bildschirm platziert und geht nicht automatisch auf, sondern muss aktiv angeklickt werden.

JK: Wir sind noch nicht ganz zufrieden mit der Anzahl der Nutzenden. Das hat sicherlich einen wirtschaftlichen Aspekt: Natürlich wird man langfristig fragen müssen, ob Mittel gut eingesetzt werden, wenn der Chatbot Fragen in geringerem Umfang beantwortet als Redaktionsmitarbeitende für dessen Betrieb nötig sind. Im Moment geht es in erster Linie um die Erprobung der Technologie. Im Hinblick auf einen Proof of Concept (POC), um die technische Machbarkeit zu demonstrieren und einen Proof of Value (POV) zur Darstellung des Nutzens ist die Anwendung auf jeden Fall eine gute Sache. Dass dieser einfache FAQ-Chatbot Nutzen bringt, können wir klar sagen. Es geht uns immer darum, was macht wo Sinn. Im Bereich der privaten Altersvorsorge wollen wir diesen Dienst für Riester auf jeden Fall aufrechterhalten, weil wir ihn für zielführend halten und zufriedene Kunden haben. Wir bekommen gutes Feedback und wollen weiter daraus lernen.

Eine Herausforderung besteht sicherlich darin, den Chatbot bekannt zu machen und Vorbehalte abzubauen. Man kennt es ja aus eigener Erfahrung, dass Chatbots nicht immer ein Erfolg sind. Ich finde es etwa eine gute Idee, Beta-Versionen zu

veröffentlichen und Nutzer ausprobieren zu lassen. Aber auch hier könnten wir als Deutsche Rentenversicherung Bund schnell an rechtliche Grenzen stoßen – Stichwort: Verlässlichkeit von Aussagen.

Häufig ist ja zu beobachten, dass sich Organisationen beim Einsatz neuer Technologien zuallererst am technisch Machbaren orientieren. Worin besteht die Hauptmotivation für die Auseinandersetzung mit KI innerhalb der Deutschen Rentenversicherung Bund? Ist es dieser Reiz des Neuen und das Mitmachen am KI-Wettlauf? Sind es vorwiegend ökonomische Gründe? Welche Faktoren spielen eine Rolle?

JK: Wir als Deutsche Rentenversicherung Bund nehmen das Denken vom Kunden her sehr ernst. Wir blicken uns in der Welt um und überlegen, was benötigt wird und welche Entwicklungen Kunden dann auch annehmen; unsere Agenda ist es nicht, Dinge zu entwickeln, weil wir sie selbst toll finden. Aber natürlich spielen auch wirtschaftliche Faktoren eine Rolle, denn idealerweise werden durch KI-Anwendungen andere unserer Services entlastet. Wer einem Chatbot eine Frage stellt und mit der Antwort zufrieden ist, wird diese Frage nicht mehr der Hotline stellen.

MM: Wichtig ist auch: Die Frage kann vom Chatbot 24/7 beantwortet werden und wenn der Kunde nicht das Gefühl hat, mit einer Maschine zu kommunizieren: umso besser!

JK: Wenn man überlegt, wie viele Mitarbeitende noch mit Betrieb und Weiterentwicklung des Chatbots beschäftigt sind, bin ich nicht sicher, ob wir bereits im gleichen Umfang Mitarbeitende in den Servicestellen entlasten. Das ist aber auch gar nicht unsere vorrangige Intention. Es geht immer auch darum, private Altersvorsorge sozialpolitisch zu fördern. Ein Business Case muss daher immer auch den Nutzen für die Gesellschaft insgesamt miteinbeziehen. Was wir hier tun, muss letzten Endes für jede Einzelne und jeden Einzelnen Nutzen entfalten.

Welche Rolle spielt also, würden Sie abschließend sagen, KI für die Deutsche Rentenversicherung Bund?

JK: Wir haben KI fest im Blick, aber wir verlassen uns keineswegs auf sie. Die Auseinandersetzung mit dem Chatbot und mit KI hat bei uns auch intern zu einem Wandel geführt: dadurch, dass wir aktiv Feedback einholen, beschäftigen wir uns viel intensiver mit den Bedürfnissen der Nutzenden. Früher waren FAQs viel statischer. Wir erhielten keinerlei Feedback, ob diese für gut befunden wurden oder ob sie praktisch anwendbar waren. In meinen Augen ist es ein wesentlicher Punkt, dass wir im Laufe dieses ganzen Prozesses interaktiver geworden sind.

Schließlich sollte man noch festhalten, dass der Umgang mit Technologie jeweils spezifische Kompetenzen erfordert, die sich auch aus der Erfahrung heraus entwickeln. Nicht ohne Grund lernen wir als Kinder zunächst Tretroller fahren, dann Fahrradfahren, dann irgendwann fahren wir Auto und erst dann werden einige von uns Rennfahrer. Vor diesem Hintergrund versuchen wir, diese Entwicklungen, die

wir in der ZfA im Kleinen haben, nach und nach in die gesamte DRV Bund auszurollen. Aufgrund der viel größeren Mitarbeiterschaft ergeben sich dann auch andere Business Cases. Der wichtige Punkt ist aber: Wir gehen hier bei der ZfA das eine oder andere Experiment ein, übernehmen auch eine gewisse Vorreiterrolle, aber als Teil der DRV Bund geht es uns letzten Endes um integrative Lösungen.

7.3 DR. HENRIKE STEPHANI (FRAUNHOFER ITWM) Wissenschaftliche Referentin Institutsstrategie und Stellvertretende Abteilungsleiterin „Bildverarbeitung“, Fraunhofer-Institut für Techno- und Wirtschafts- mathematik ITWM

Im Projekt „PflegeForensik“ hat das Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM in Kooperation mit der Generalstaatsanwaltschaft Dresden und dem Kommissariat für Wirtschaftskriminalität der Polizeidirektion Leipzig eine Softwarelösung entwickelt, die mit künstlicher Intelligenz (KI) dabei unterstützt, Abrechnungsbetrug in der Pflege aufzudecken. Worin bestand Ihre Motivation, KI für diesen Anwendungsfall einzusetzen, welches Problem versuchen Sie mit der KI zu lösen?

HS: Gerade im Bereich der ambulanten Pflege hat man es in Deutschland mit einer besonderen Situation zu tun, die dazu führt, dass Abrechnungen sehr wenig kontrolliert werden. Zum Teil erfolgen Abrechnungen handschriftlich. Es gibt wenig Vorgaben in diesem Bereich, was nicht zuletzt daran liegt, dass die Politik in Zeiten des Pflegenotstands den Pflegediensten die Arbeit nicht noch schwerer machen möchte. Als Kehrseite ergibt sich allerdings ein gewisser Wildwuchs. Weder bei den Krankenkassen noch bei der Polizei kann daher alles systematisch geprüft werden, weil dies viel zu aufwendig wäre. Die Polizei prüft ohnehin nur, wenn es einen Anfangsverdacht gibt, wenn sie also entweder von ehemaligen Pflegekräften, von Patienten oder der Krankenkasse auf Auffälligkeiten in größerem Umfang aufmerksam gemacht wird. Weil die Prüfung mit komplexer, manueller Papierarbeit verbunden und daher sehr zeitaufwendig ist, beschränkt man sich zumeist auf einen extrem kleinen Tatzeitraum – kleiner Tatzeitraum bedeutet aber immer auch kleiner Betrugsumfang. Und an diesem Punkt setzen wir an, um zu helfen.

Wie geht die KI vor, um Betrugsfälle zu erkennen?

HS: Die KI übernimmt zwei Aufgaben: Zum einen unterstützt sie bei der Digitalisierung der Papierdokumente. Dabei werden mittels Bildverarbeitung den Dokumenten die entsprechenden Informationen entnommen. Zum anderen werden diese Daten dann abgeglichen. Beispielsweise wird die Liste mit den Qualifikationen der Pflegekräfte den erbrachten Leistungen gegenübergestellt oder es wird geprüft, ob jemand an einem Tag, an dem er angeblich gearbeitet hat, nicht laut Dienstplan im Urlaub war. Auch wird geprüft, ob jemand an einem Tag nicht viel zu viele Leistungen erbracht hat.

Stichwort Pflegenotstand: Ist es vor dem Hintergrund der extrem angespannten Personalsituation nicht eine Gratwanderung, welche Auffälligkeiten als Betrug markiert werden und welche Unregelmäßigkeiten schlichtweg notwendig sind, um die Versorgung aufrecht zu erhalten?

HS: Wir haben einen Pflegenotstand und zu wenige qualifizierte Pflegekräfte. Insofern sind wir in einer schwierigen Situation, weil natürlich überhaupt kein Interesse besteht, gut und zuverlässig tätige Pflegedienste zu bestrafen und die Arbeit zu erschweren. Aber natürlich muss sichergestellt werden, das Geld im System zu halten.

Es geht bei dem Projekt überhaupt nicht darum, vor Ort nachzuprüfen, ob Pflegekräfte das Richtige machen. Wir haben kein Interesse daran, Fälle, die aus einer Mangelsituation heraus entstehen – z. B. keine Verfügbarkeit eines Vormundes des zu Pflegenden – oder extreme Zeitknappheit zu bestrafen. Grundsätzlich besteht ein großes Grundvertrauen in die Pflegekräfte, dass vor Ort gute Arbeit geleistet wird. Uns geht es um Fälle systematischen Betrugs. Dabei nehmen wir den gesamten Pflegedienst, nicht die einzelne Pflegekraft ins Visier. Wenn zum Beispiel Pflegekräfte minderer Qualifikation eingesetzt werden, kann das auch problematisch für Patienten sein – etwa in Fällen, in denen sich um Intensivpflegepatienten keine speziell geschulte Intensivpflegekraft kümmert. Es gibt eine ganze Reihe solcher Betrugssachverhalte, bis hin dazu, dass Patienten gar nicht angefahren und versorgt werden. Letztendlich sind es ja auch nicht die einzelnen Pflegekräfte, die am Betrug verdienen, sondern es ist der Pflegedienst, der in betrügerischer Weise verdient, wenn beispielsweise zu gering qualifizierte Arbeitskräfte geschickt werden.

Lässt sich feststellen, wie wirksam die KI-Anwendung bei der Betrugserkennung im Vergleich zur manuellen Prüftätigkeit ist?

HS: Im Rahmen der Erstellung des Demonstrators haben wir festgestellt, dass ca. fünf Monate Arbeit einer Polizeifachkraft auf zwei Monate reduziert werden kann. Das ist eine ganz erhebliche Reduktion. Aber davon abgesehen kann der Umfang der Analyse ausgeweitet werden. Das Schöne an KI-Systemen ist ja, dass das anfängliche Aufsetzen des Systems einen sehr großen Aufwand bedeutet, aber wenn das System erst einmal läuft, kann es beliebig mit Daten erweitert werden. Das sind ganz klare Vorteile der KI im Vergleich zur menschlichen Arbeitskraft, die immer ein Skalierungsproblem hat.

Dann scheint Ihr System ein guter Weg zu sein, das Pflegesystem durch KI insgesamt zu verbessern. Wird es bereits eingesetzt?

HS: In dem Projekt haben wir zunächst nur einen Demonstrator entwickelt, d. h. wir haben die Machbarkeit eines entsprechenden Systems gezeigt. Wir sind gerade dabei, uns um eine Anschlussfinanzierung zu bemühen, damit die Software dann wirklich von Ermittelnden bei der Polizei oder bei der Krankenkasse genutzt werden kann. Wir kooperieren mit der Polizei und setzen dabei die Software im Rahmen von Gutachten ein.

Der Weg zum wirklichen Einsatz ist schwierig: Zurzeit wird nicht viel gefördert und das Fördersystem in Deutschland ist prinzipiell so aufgebaut, dass in erster Linie auf Grundlagen fokussiert wird. Aber an dem Punkt, an dem man zum Schluss kommt, dass etwas machbar ist, bleiben die Projekte dann häufig stecken, weil die Entwicklungen zu dem Zeitpunkt für Verwerter noch zu unattraktiv und riskant sind. Was die Krankenkassen betrifft muss man sagen, dass diese zwar großes Interesse an dem Thema haben, aber die Investition intern nicht einfach zu rechtfertigen ist: Oft ist es ja so, dass sich Investitionen dadurch begründen lassen, dass etwas billiger wird, was vorher teurer war. Da mir aber scheint, dass Abrechnungsprüfungen überhaupt nur sehr stichprobenartig durchgeführt werden, geht diese Rechnung nicht auf. Schließlich sind auch die Kapazitäten für die Entwicklung begrenzt – so ist etwa bei der Polizei, mit der wir zusammenarbeiten, deren Bereich des Gesundheitswirtschaftsbetrugs derzeit stark durch die Betrugsfälle rund um die Corona-Teststationen in Beschlag genommen.

Gibt es ansonsten irgendwelche Probleme und Hürden, die Sie bei der Entwicklung des KI-Systems in Ihrem Projekt „PflegeForensik“ ausbremsen?

HS: Der Datenschutz ist in Deutschland eine große Herausforderung. Die Vorgaben sind natürlich nachvollziehbar und wir agieren daher sehr datenschutzkonform. Weniger nachvollziehbar sind Probleme, die im Zusammenhang mit der Komplexität des deutschen Gesundheitssystems als föderales Multikrankenkassensystem auftreten. Zum einen führt das zum ganz praktischen Problem, dass die Krankenkassen, weil sie nur über die Daten ihrer eigenen Patienten verfügen, die KI-Systeme gar nicht so einfach nutzen könnten. Zum anderen führt der Föderalismus neben dem Erfordernis vieler Abstimmungen dazu, dass es immer ein Bundesland als Vorreiter braucht, das die Dinge angeht und finanziert, die aber dann doch für alle gleichermaßen sinnvoll sind.

Generell wird zurzeit ja viel mit KI experimentiert. Geht man in dem sensiblen Bereich der Sozialversicherung etwas vorsichtiger an die Sache heran bzw. gibt es irgendwelche Besonderheiten in dem Bereich zu beachten?

HS: Wir haben in dem Projekt mit der Staatsanwaltschaft und der Polizei zusammengearbeitet und dabei wurde gefordert, dass jeder Schritt einer Prüfung einem Richter plausibel zu machen ist. Im Zweifel wird jedes Mal nachgefragt, weshalb an der betreffenden Stelle welche Entscheidung getroffen wurde. Das ist ein riesiger Unterschied zu anderen KI-Projekten. Wir bereiten daher alles so auf, dass es nachprüfbar ist. Das ist auch der Grund, weshalb wir die KI zunächst einmal einsetzen für die Digitalisierung, d. h. um mittels Bildverarbeitung aus den Dokumenten das Richtige zu erkennen. Aber auch hier versuchen wir, alles so weit wie möglich nachvollziehbar und durch Menschen nachprüfbar auszugestalten. Für die Betrugserkennung selbst nutzen wir regelbasierte Systeme. Das bedeutet, dass zunächst definiert wird, was einen Betrugsfall ausmacht und in der Folge maschinell Abgleiche durchgeführt werden.

Gewissermaßen helfen Sie also zunächst einmal der Digitalisierung auf die Sprünge, die eigentlich Voraussetzung ist, um das volle Potenzial der KI auszuschöpfen...

HS: Um mittels Mustererkennung diffusere Auffälligkeiten zu erkennen, müsste die Digitalisierung in Deutschland weiter sein. Dafür haben wir die Daten überhaupt nicht digital vorliegen. Daher müssen wir erst einmal ganz viel handfeste Arbeit machen, um zur Magie der KI zu kommen. Aber insgesamt sollte man vorsichtig sein, welche Erwartungen an die KI geknüpft werden: In einem anderen Projekt arbeiten wir gerade daran, Auffälligkeiten in Gesundheitsdaten aufzuspüren. Oft passiert es, dass das KI-System interessante Auffälligkeiten findet und wenn wir die Ergebnisse den Leuten zeigen, die die Daten kennen, liefern sie ganz profane Erklärungen für die Auffälligkeiten, wie zum Beispiel, dass das Internet am betreffenden Tag ausgefallen sei oder ein neuer Mitarbeiter die Daten dreimal eingegeben habe. Die KI weiß natürlich in diesen Fällen nicht, dass es sich um ganz triviale, erklärbare Abweichungen handelt.

Welche Rolle sollte Ihrer Meinung nach KI in einer Zukunftsvision der Sozialversicherung spielen?

HS: Ich denke, KI kann stark bei der Digitalisierung helfen sowie bei Assistenzsystemen und überall dort, wo es um formelles Abarbeiten geht. Dokumentation wäre hier ein Beispiel, die viel Arbeitskraft und Energie von hochqualifizierten Menschen kostet. Im Rahmen der Leistungserbringung kann ich mir vorstellen, KI zum Sortieren von Medikamenten zu nutzen. Es ist wichtig, KI dort einzusetzen, wo Freiräume für die menschliche Interaktion und das menschliche Urteilsvermögen geschaffen werden. Eine Arbeitsteilung, in der die KI Empfehlungen abgibt und am Ende immer die Einschätzung durch den Menschen steht, erscheint mir sinnvoll. KI sollte einen Hintergrund bilden, der das System einfacher macht – weniger Formalitäten, mehr Freiheit für den Menschen, das zu tun, was Menschen gut können.

8 QUELLENVERZEICHNIS

- Ada Lovelace Institute & DataKind UK (2020) *Examining the Black Box: Tools for assessing algorithmic systems* [Online], London.
- Alston, P. (2019) *Report of the Special Rapporteur on extreme poverty and human rights* [Online]. Verfügbar unter <https://documents.un.org/doc/undoc/gen/n19/312/13/pdf/n1931213.pdf> (Abgerufen am 20 August 2024).
- Bacchini, F. & Lorusso, L. (2019) "Race, again: how face recognition technology reinforces racial discrimination", *Journal of Information, Communication and Ethics in Society*, Vol. 17, No. 3, S. 321–335.
- Badidi, E. (2023) "Edge AI for Early Detection of Chronic Diseases and the Spread of Infectious Diseases: Opportunities, Challenges, and Future Directions", *Future Internet*, Vol. 15, No. 11, S. 370.
- Barca, V. & Hebbar, M. (2021) "Delivering social transfers", in Schüring, E. & Loewe, M. (Hg.) *Handbook on Social Protection Systems*, Cheltenham, UK, Edward Elgar Publishing, S. 426–436.
- Beck, S. (2020) „Künstliche Intelligenz – ethische und rechtliche Herausforderungen“, in Mainzer, K. (Hg.) *Philosophisches Handbuch Künstliche Intelligenz*, Wiesbaden, Springer VS, S. 1–28.
- Beckert, B. (2021) „Vertrauenswürdige künstliche Intelligenz: Ausgewählte Praxisprojekte und Gründe für das Umsetzungsdefizit“, *TaTuP*, Vol. 30, No. 3, S. 17–22.
- Beer, D. (2016) "The social power of algorithms", *Information, Communication & Society*, Vol. 20, No. 1, S. 1–13.
- Bertsimas, D., Pawlowski, C. & Zhuo, Y. D. (2018) "From Predictive Methods to Missing Data Imputation: An Optimization Approach", *The Journal of Machine Learning Research*, Vol. 18, No. 1, S. 7133–7171.
- Birkeland, S., Bogh, S. B. & Morsø, L. (2024) "From systematic complaint analysis to quality improvement in healthcare", *BMJ Open Quality*, Vol. 13.
- Bishop, J. M. (2021) "Artificial Intelligence Is Stupid and Causal Reasoning Will Not Fix It", *Frontiers in Psychology*, Vol. 11, No. 513474.
- Bitkom Research (2023) *Künstliche Intelligenz – Wo steht die deutsche Wirtschaft?* [Online]. Verfügbar unter <https://www.bitkom.org/sites/main/files/2023-09/bitkom-charts-ki-im-unternehmen.pdf> (Abgerufen am 27 Juni 2024).
- Bovens, M. & Zouridis, S. (2002) "From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control", *Public Administration Review*, Vol. 62, No. 2, S. 174–184.

- boyd, d. & Crawford, K. (2011) „Big Data als kulturelles, technologisches und wissenschaftliches Phänomen. Sechs Provokationen“, in Geiselberger, H. & Moorstedt, T. (Hg.) *Big Data: Das neue Versprechen der Allwissenheit*, Berlin, Suhrkamp, S. 187–218.
- Braun, C. (2024) „Klicken und Verticken“, *DIE ZEIT*, No. 27, S. 23.
- Brussig, M. (2017) „Methoden und Daten zur Erforschung spezieller Organisationen: Organisationen der Sozialversicherung“, in Liebig, S., Matiaske, W. & Rosenbohm, S. (Hg.) *Handbuch Empirische Organisationsforschung*, Wiesbaden, Springer Gabler, S. 589–608.
- Büchel, J., Bakalis, D., Scheufen, M. & Schmitz, E. (2024) *Digitalisierung der Wirtschaft in Deutschland: Digitalisierungsindex 2023* [Online]. Verfügbar unter <https://www.de.digital/DIGITAL/Navigation/DE/Lagebild/Digitalisierungsindex/digitalisierungsindex.html> (Abgerufen am 2 Juli 2024).
- Büchel, J. & Engels, B. (2022) „Viele Unternehmen sind nicht bereit für die Datenwirtschaft“, *IW-Kurzbericht*, No. 96, S. 1–2.
- Buchholz, K. (2023) *Threads Shoots Past One Million User Mark at Lightning Speed* [Online], Statista. Verfügbar unter <https://www.statista.com/chart/29174/time-to-one-million-users/> (Abgerufen am 26 Juni 2024).
- Bughin, J. & Hazan, E. (2017) *The new spring of artificial intelligence: A few early economies* [Online]. Verfügbar unter <https://www.mckinsey.com/mgi/overview/in-the-news/the-new-spring-of-artificial-intelligence-a-few-early-economics> (Abgerufen am 2 Juli 2024).
- Bundesagentur für Arbeit (o.D.) <https://www.arbeitsagentur.de/vor-ort/it-systemhaus/themen/advanced-analytics> [Online]. Verfügbar unter <https://www.arbeitsagentur.de/vor-ort/it-systemhaus/themen/advanced-analytics> (Abgerufen am 29 Juli 2024).
- Bundesagentur für Arbeit (2024) *Deutscher Digitaltag: BA setzt Künstliche Intelligenz (KI), maschinelles Lernen und Automatisierung ein* [Online]. Verfügbar unter <https://www.arbeitsagentur.de/presse/2024-26-deutscher-digitaltag-ba-setzt-kuenstliche-intelligenz-maschinelles-lernen-und-automatisierung-ein> (Abgerufen am 29 Juli 2024).
- Cavoukian, A. (2012) “Privacy by Design: Origins, Meaning, and Prospects for Assuring Privacy and Trust in the Information Era”, in Yee, G. O. (Hg.) *Privacy Protection Measures and Technologies in Business Organizations: Aspects and Standards*, Hershey, PA, IGI Global, S. 170–208.
- Church, A. (1936) “An Unsolvable Problem of Elementary Number Theory”, *American Journal of Mathematics*, Vol. 58, No. 2, S. 345–363.
- Cockburn, I. M., Henderson, R. & Stern, S. (2019) “The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis”, in Agrawal, A., Gans, J. & Goldfarb, A. (Hg.) *The Economics of Artificial Intelligence: An Agenda*, Chicago, IL, University of Chicago Press, S. 115–148.

Damioli, G., van Roy, V. & Vertesy, D. (2021) "The impact of artificial intelligence on labor productivity", *Eurasian Business Review*, No. 11, S. 1–25.

Dastin, J. (2018) *Insight - Amazon scraps secret AI recruiting tool that showed bias against women* [Online], Reuters. Verfügbar unter <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/> (Abgerufen am 20 August 2024).

Datenethikkommission (2019) *Gutachten der Datenethikkommission der Bundesregierung* [Online], Berlin.

Deloitte (2024) *Now decides next: Insights from the leading edge of generative AI adoption in Germany: Deloitte's State of Generative AI in the Enterprise Quarter one report - German Cut* [Online]. Verfügbar unter <https://www2.deloitte.com/content/dam/Deloitte/de/Documents/Innovation/us-state-of-gen-ai-report-q2.pdf> (Abgerufen am 27 Juni 2024).

Denkfabrik BMAS (2024a) *Verwaltung modern und digital: BG BAU setzt eine neue KI-Anwendung zur Vermeidung von Arbeitsunfällen in der Bauwirtschaft ein* [Online]. Verfügbar unter <https://www.denkfabrik-bmas.de/projekte/ki-in-der-verwaltung/verwaltung-modern-und-digital-bg-bau-setzt-eine-neue-ki-anwendung-zur-vermeidung-von-arbeitsunfaellen-in-der-bauwirtschaft-ein> (Abgerufen am 25 Juli 2024).

Denkfabrik BMAS (2024b) *Verwaltungsinnovation für Beitragsgerechtigkeit, soziale Sicherheit und einen zukunftsfähigen Sozialstaat: Mit KI Beschäftigte unterstützen und die Sozialversicherung schützen* [Online]. Verfügbar unter <https://www.denkfabrik-bmas.de/projekte/ki-in-der-verwaltung/verwaltungsinnovation-fuer-beitragsgerechtigkeit-soziale-sicherheit-und-einen-zukunftsfahigen-sozialstaat-mit-ki-beschaefigte-unterstuetzen-und-die-sozialversicherung-schuetzen> (Abgerufen am 29 Juli 2024).

Desiere, S., Langenbucher, K. & Struyven, L. (2019) "Statistical profiling in public employment services: An international comparison: OECD Social, Employment and Migration Working Papers (No. 224) [Online], Paris, OECD Publishing.

Deutsche Sozialversicherung (2020) *Die öffentliche Konsultation der EU-Kommission zum Weißbuch "Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen": Stellungnahme der Deutschen Sozialversicherung vom 11. Juni 2020* [Online]. Verfügbar unter <https://dsv-europa.de/lib/2020-06-11-DSV-Position-Weissbuch-Endfassung.pdf> (Abgerufen am 14 August 2024).

Deutscher Ethikrat (2023) *Mensch und Maschine: Herausforderungen durch Künstliche Intelligenz* (Stellungnahme) [Online], Berlin, Deutscher Ethikrat.

Dreyling III, R. M., Tammet, T. & Pappel, I. (2023) „Digital Transformation Insights from an AI Solution in Search of a Problem“, in Dang, T. K., Küng, J. & Chung, T. M. (Hg.) *Future Data and Security Engineering: Big Data, Security and Privacy, Smart City and Industry 4.0 Applications*, Singapur, Springer, S. 341–351.

- DRV Bund (o.D.) *Künstliche Intelligenz entlastet Mitarbeitende und schützt das Sozialversicherungssystem* [Online]. Verfügbar unter <https://www.deutscherentenversicherung.de/Bund/DE/Ueber-uns/Digitalstrategie/KIRA.html> (Abgerufen am 29 Juli 2024).
- Effer-Uhe, D. (2023) „Überlegungen zur Automatisierbarkeit der Rechtsanwendung“, *JuristenZeitung*, Vol. 78, No. 19, S. 833–842.
- Engelmann, J. & Puntschuh, M. (2020) *KI im Behördeneinsatz: Erfahrungen und Empfehlungen* [Online], Kompetenzzentrum Öffentliche IT. Verfügbar unter <https://www.oeffentliche-it.de/documents/10181/14412/KI+im+Beh%C3%B6rdeneinsatz+-+Erfahrungen+und+Empfehlungen> (Abgerufen am 26 Juli 2024).
- Engels, B. (2023) „Künstliche Intelligenz in der deutschen Wirtschaft: Ohne Digitalisierung und Daten geht nichts“, *Wirtschaftsdienst*, Vol. 103, No. 8, S. 525–529.
- Europäische Kommission (2018) *Communication Artificial Intelligence for Europe* [Online]. Verfügbar unter <https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe> (Abgerufen am 1 Juli 2024).
- Europäische Kommission (2019) *Ethik-Leitlinien für eine vertrauenswürdige KI* [Online]. Verfügbar unter <https://data.europa.eu/doi/10.2759/22710> (Abgerufen am 7 August 2024).
- Europäische Kommission (2020) *Weißbuch: Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen* [Online]. Verfügbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0065> (Abgerufen am 1 Juli 2024).
- Europäische Kommission (2021) *Coordinated Plan on Artificial Intelligence 2021 Review* [Online]. Verfügbar unter <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review> (Abgerufen am 1 Juli 2024).
- Europäisches Parlament (2017) *Bericht mit Empfehlungen an die Kommission zu zivilrechtlichen Regelungen im Bereich Robotik* [Online]. Verfügbar unter https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_DE.html (Abgerufen am 8 August 2024).
- Europäisches Parlament (2023) *Was ist künstliche Intelligenz und wie wird sie genutzt?* [Online]. Verfügbar unter <https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt> (Abgerufen am 28 Juni 2024).
- Europäisches Parlament (2024) *Verordnung über künstliche Intelligenz* [Online]. Verfügbar unter https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_DE.pdf (Abgerufen am 1 Juli 2024).

- Farbmacher, H., Löw, L. & Spindler, M. (2022) "An explainable attention network for fraud detection in claims management", *Journal of Econometrics*, Vol. 228, No. 2, S. 244–258.
- Feike, M., Bienzeisler, B. & Neuhüttler, J. (2024) *Künstliche Intelligenz aus Sicht von Unternehmen: Status quo und Potenziale in der Region Heilbronn-Franken* [Online], Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO.
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. & Srikuma, M. (2020) *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI* [Online], Berkman Klein Center for Internet & Society.
- Ford, J., Jain, V., Wadhvani, K. & Gupta, D. G. (2023) "AI advertising: An overview and guidelines", *Journal of Business Research*, Vol. 166, S. 1–15.
- Forster, J. & Entrup, B. (2017) *A Cognitive Computing Approach for Classification of Complaints in the Insurance Industry* [Online], *IOP Conference Series: Materials Science and Engineering*, Vol. 261, No. 1.
- Gantchev, V. (2019) "Data protection in the age of welfare conditionality: Respect for basic rights or a race to the bottom?", *European Journal of Social Security*, Vol. 21, No. 1, S. 3–22.
- Gao, B., Wang, Y., Xie, H. & Hu, Y. (2023) "Artificial Intelligence in Advertising: Advancements, Challenges, and Ethical Considerations in Targeting, Personalization, Content Creation, and Ad Optimization", *SAGE Open*, S. 1–20.
- Gardner, H. (1983) *Frames of Mind: The Theory of Multiple Intelligences*, New York, NY, Basic Books.
- Geiger, G. (2023) "How Denmark's Welfare State Became a Surveillance Nightmare", *Wired* [Online]. Verfügbar unter <https://www.wired.com/story/algorithms-welfare-state-politics/> (Abgerufen am 20 August 2024).
- Geißler, J. (2017) *Die Digitalisierung von Geschäftsprozessen in der gesetzlichen Kranken- und Unfallversicherung als Herausforderung für das Management: ifgs Schriftenreihe der FOM*, No. 9 [Online], Essen, MA Akademie Verlags- und Druck-Gesellschaft mbH.
- Gitelman, L (Hg.) (2013) *"Raw Data" Is an Oxymoron*, Cambridge, MA, The MIT Press.
- GKV-Spitzenverband (2024) *Anreize für die Krankenhausabrechnung – einfach korrekt abrechnen!: Argumentationspapier des GKV-Spitzenverbandes zur Krankenhausabrechnungsprüfung* [Online]. Verfügbar unter https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/krankenhaeuser/abrechnung/abrechnungspruefung/20240426_Argumentationspapier_KH-Abrechnung_bf.pdf (Abgerufen am 11 Juli 2024).

- Government Digital Service & Office for Artificial Intelligence (2019) *A guide to using artificial intelligence in the public sector* [Online]. Verfügbar unter <https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector> (Abgerufen am 8 August 2024).
- Hallensleben, S. & Hustedt, C. (2020) *From Principles to Practice: An interdisciplinary framework to operationalise AI ethics* [Online], AI Ethics Impact Group, VDE, Bertelsmann Stiftung.
- Hassan, A. F., Barakat, S. & Rezk, A. (2022) "Towards a deep learning-based outlier detection approach in the context of streaming data", *Journal of Big Data*, Vol. 9, No. 120, S. 1–16.
- Hausmann, G. & Lämmel, U. (2021) „Künstliche Intelligenz in der automatisierten Dokumentenverarbeitung am Beispiel von Krankenversicherungen“, in Barton, T. & Müller, C. (Hg.) *Künstliche Intelligenz in der Anwendung: Rechtliche Aspekte, Anwendungspotenziale und Einsatzszenarien*, Wiesbaden, Springer Vieweg, S. 177–193.
- Hehner, S., Körs, B., Martin, M., Uhrmann-Klingen, U. & Waldron, J. (2017) *Artificial intelligence in health insurance: Smart claims management with self-learning software* [Online], McKinsey & Company. Verfügbar unter <https://www.mckinsey.com/industries/healthcare/our-insights/artificial-intelligence-in-health-insurance-smart-claims-management-with-self-learning-software> (Abgerufen am 11 Juli 2024).
- Hellström, T. (2013) "On the moral responsibility of military robots", *Ethics and Information Technology*, Vol. 15, S. 99–107.
- Henman, P. (2020) "Improving public services using artificial intelligence: possibilities, pitfalls, governance", *Asia Pacific Journal of Public Administration*, No. 42, S. 209–221.
- Henman, P. (2022) "Digital technologies and artificial intelligence: A computer science perspective", in Adler, M. (Hg.) *A Research Agenda for Social Welfare Law, Policy and Practice*, Cheltenham, UK, Edward Elgar Publishing, S. 265–282.
- High-Level Expert Group on Artificial Intelligence (2020) *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment* [Online]. Verfügbar unter <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment> (Abgerufen am 15 August 2024).
- Hirsch-Kreinsen, H. & Krokowski, T. (2023) "Trustworthy AI: AI made in Germany and Europe?", *AI & Society*.
- Hofmann, M. (2021) *Erfahrungen mit KI in der Bundesagentur für Arbeit: Impulsvortrag* [Online]. Verfügbar unter <https://www.denkfabrik-bmas.de/projekte/ki-in-der-verwaltung/impulsvortrag-von-martina-hofmann> (Abgerufen am 29 Juli 2024).

- Hohnstein, A. (2021) „KI für die Intensivstation: Weniger lebensbedrohliche Komplikationen durch digitales Frühwarnsystem“, *Klinik Management aktuell*, Vol. 26, No. 11, S. 38–40.
- Humm, B. G., Buxmann, P. & Schmidt, J. C. (2022) „Grundlagen und Anwendungen von KI“, in Gethmann, C. F., Buxmann, P., Distelrath, J., Humm, B. G., Lingner, S., Nitsch, V., Schmidt, J. C. & Spiecker gen. Döhmann, I. (Hg.) *Künstliche Intelligenz in der Forschung: Neue Möglichkeiten und Herausforderungen für die Wissenschaft*, Berlin, Springer, S. 13–42.
- Ismailbekovich, B. D. (2024) “Implementing chatbots for consumer complaint response”, *Proceedings of International Conference on Modern Science and Scientific Studies*, Vol. 3, No. 2, S. 440–445.
- IVSS (2020) *Künstliche Intelligenz in der sozialen Sicherheit: Hintergründe und Erfahrungen* [Online]. Verfügbar unter <https://www.issa.int/de/analysis/artificial-intelligence-social-security-background-and-experiences> (Abgerufen am 24 Juli 2024).
- IVSS (2022a) *Aufdeckung von Betrug im Gesundheitswesen durch neue Technologien* [Online]. Verfügbar unter <https://www.issa.int/de/analysis/detecting-fraud-health-care-through-emerging-technologies> (Abgerufen am 25 Juli 2024).
- IVSS (2022b) *Künstliche Intelligenz in Institutionen der sozialen Sicherheit: Der Fall intelligenter Chatbots* [Online]. Verfügbar unter <https://www.issa.int/de/analysis/artificial-intelligence-social-security-institutions-case-intelligent-chatbots> (Abgerufen am 24 Juli 2024).
- Johnson, D. G. (2015) “Technology with No Human Responsibility?”, *Journal of Business Ethics*, Vol. 127, S. 707–715.
- Johnson, J. & Khoshgoftaar, T. M. (2023) “Data-Centric AI for Healthcare Fraud Detection”, *SN Computer Science*, Vol. 4, No. 389.
- Kasper, T. (2020) *Wie der Sozialstaat digital wurde: Die Computerisierung der Rentenversicherung im geteilten Deutschland*, Göttingen, Wallstein.
- Kraus, T., Ganschow, L., Eisenträger, M. & Wischmann, S. (2021) *Erklärbare KI: Anforderungen, Anwendungsfälle und Lösungen* (Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi) im Rahmen der Begleitforschung zum Technologieprogramm „Künstliche Intelligenz als Treiber für volkswirtschaftlich relevante Ökosysteme“ (KI-Innovationswettbewerb)) [Online], Berlin.
- Krause, M. (2023) „Fortschritt mit Verantwortung“, *etem*, No. 2, S. 18.
- Krüger, S. & Wilson, C. (2023) “The problem with trust: on the discursive commodification of trust in AI”, *AI & Society*, Vol. 38, S. 1753–1761.
- La Salle, D. (2021) “Delivering social insurance”, in Schüring, E. & Loewe, M. (Hg.) *Handbook on Social Protection Systems*, Cheltenham, UK, Edward Elgar Publishing, S. 437–447.

- Langquillon, C. & Schacht, S. (2023) *Knowledge Science - Grundlagen: Methoden der Künstlichen Intelligenz für die Wissensextraktion aus Texten*, Wiesbaden, Springer Vieweg.
- Laux, J., Wachter, S. & Mittelstadt, B. (2024) "Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk", *Regulation & Governance*, Vol. 18, No. 1, S. 3–32.
- Lessig, L. (1999) *Code and Other Laws of Cyberspace*, New York, NY, Basic Books.
- Lindert, K, Karippacheril, TG, Caillava, IR & Chávez, KN (Hg.) (2020) *Sourcebook on the Foundations of Social Protection Delivery Systems*, Washington, DC, The World Bank.
- Lipsky, M. (1984) "Bureaucratic Disentitlement in Social Welfare Programs", *Social Service Review*, Vol. 58, No. 1, S. 3–27.
- Luhmann, N. (2022) „Bürokratie im Wohlfahrtsstaat“, in Tacke, V. & Lukas, E. (Hg.) *Schriften zur Organisation 5: Vorträge, Lexikonartikel, Rezensionen*, Wiesbaden, Springer VS, S. 247–252.
- Manovich, L. (2001) *The Language of New Media*, Cambridge, MA, The MIT Press.
- Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J. & Perrault, R. (2023) *Artificial Intelligence Index Report 2023* [Online], Stanford, CA, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University.
- Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, Juan Carlos, Shoham, Y., Wald, R. & Clark, J. (2024) *The AI Index 2024 Annual Report* [Online], AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Verfügbar unter https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf (Abgerufen am 3 Juli 2024).
- Matthias, A. (2004) "The responsibility gap: Ascribing responsibility for the actions of learning automata", *Ethics and Information Technology*, Vol. 6, S. 175–183.
- Matthias, A. (2008) *Automaten als Träger von Rechten*, Berlin, Logos Verlag.
- Mayer-Schönberger, V. & Cukier, K. (2013) *Big Data: A Revolution that Will Transform how We Live, Work, and Think*, Boston, Houghton Mifflin Harcourt.
- McKinsey (2024) *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value* [Online]. Verfügbar unter <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai> (Abgerufen am 27 Juni 2024).
- Metzinger, T. (2023) *Ethik-Waschmaschinen made in Europe* [Online], Tagesspiegel. Verfügbar unter <https://background.tagesspiegel.de/digitalisierung-und-ki/briefing/ethik-waschmaschinen-made-in-europe> (Abgerufen am 14 August 2024).

- Moravec, H. (1988) *Mind Children: The Future of Robot and Human Intelligence*, Cambridge, MA, Harvard University Press.
- Morley, J., Floridi, L., Kinsey, L. & Elhalal, A. (2020) "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices", *Science and Engineering Ethics*, Vol. 26, S. 2141–2168.
- Nagenborg, M., Capurro, R., Weber, J. & Pingel, C. (2008) "Ethical regulations on robotics in Europe", *AI & Society*, Vol. 22, S. 349–366.
- Nair, K. & Gupta, R. (2021) "Application of AI technology in modern digital marketing environment", *World Journal of Entrepreneurship, Management and Sustainable Development*, Vol. 17, No. 3, S. 318–328.
- Nass, E. & Schneider, M. (2022) „Maschinen mit Moral für eine gute Pflege der Zukunft?“, in Pfannstiel, M. A. (Hg.) *Künstliche Intelligenz im Gesundheitswesen: Entwicklungen, Beispiele und Perspektiven*, Wiesbaden, Springer Gabler, S. 311–323.
- Nenno, S. (2021) „Siris böse Schwester. Wenn der niederländische öffentliche Dienst Ihre Daten stiehlt“, *Digital Society Blog* [Online]. Verfügbar unter <https://www.hiig.de/siris-boese-schwester-wenn-der-niederlaendische-oeffentliche-dienst-ihre-daten-stiehlt/> (Abgerufen am 20 August 2024).
- Netzwerk KI in der Arbeits- und Sozialverwaltung (2022) *Selbstverpflichtende Leitlinien für den KI-Einsatz in der behördlichen Praxis der Arbeits- und Sozialverwaltung* [Online]. Verfügbar unter https://www.denkfabrik-bmas.de/fileadmin/Downloads/Publikationen/Selbstverpflichtende_Leitlinien_fuer_den_KI-Einsatz_in_der_behoerdlichen_Praxis_der_Arbeits-_und_Sozialverwaltung.pdf (Abgerufen am 13 August 2024).
- Noble, S. U. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, NY, NYU Press.
- o.V. (2023) „BG BAU mit KI-Projekt auf dem Deutschen Baugewerbetag“, *BauPortal*, Vol. 135, No. 4, S. 16.
- ÖGK (2024) *ÖGK Jahresbericht: Aktuelles im Jahr 2023* [Online]. Verfügbar unter <https://www.gesundheitskasse.at/cdscontent/load?contentid=10008.789038&version=1720602478> (Abgerufen am 30 Juli 2024).
- Patel, K., Mistry, C., Mehta, D., Thakker, U., Tanwar, S., Gupta, R. & Kumar, N. (2022) "A survey on artificial intelligence techniques for chronic diseases: open issues and challenges", No. 55, S. 3747–3800.
- Penney, J. W. (2017) "Internet surveillance, regulation, and chilling effects online: a comparative case study", *Internet Policy Review*, Vol. 6, No. 2.
- Peters, M. (2024) „KI in der ambulanten Pflege: Fantasie oder Unterstützung?“, *Pflege Zeitschrift*, Vol. 77, No. 4, S. 52–55.

- Pflügl, J. (2024) *Entscheidet die KI über Jobs? Höchstgericht lässt AMS-Algorithmus erneut prüfen* [Online]. Verfügbar unter <https://www.derstandard.at/story/3000000206845/entscheidet-die-ki-ueber-jobs-hoechstgericht-laesst-ams-algorithmus-erneut-pruefen> (Abgerufen am 30 August 2024).
- Rammer, C. (2022) *Kompetenzen und Kooperationen zu Künstlicher Intelligenz: Ergebnisse einer Befragung von KI-aktiven Unternehmen in Deutschland* (Studie im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz) [Online].
- Rammer, C., Fernández, G. P. & Czarnitzki, D. (2022) "Artificial intelligence and industrial innovation: Evidence from German firm-level data", *Research Policy*, Vol. 51, No. 7, S. 1–15.
- Rao, R. (2022) *The Dutch Tax Authority Was Felled by AI — What Comes Next?: European regulation hopes to rein in ill-behaving algorithms* [Online], IEEE Spectrum. Verfügbar unter <https://spectrum.ieee.org/artificial-intelligence-in-government> (Abgerufen am 20 August 2024).
- Rashid, J., Batool, S., Kim, J., Wasif Nisar, M., Hussain, A., Juneja, S. & Kushwaha, R. (2022) "An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction", *Frontiers in Public Health*, Vol. 10.
- Rezaeipourfarsangi, S. & Milios, E. E. (2023) "AI-powered Resume-Job matching: A document ranking approach using deep neural networks", *DocEng '23: Proceedings of the ACM Symposium on Document Engineering 2023*.
- Rich, E., Knight, K. & Nair, S. B. (2009) *Artificial Intelligence*, Neu-Delhi, Tata McGraw-Hill.
- Rieder, G., Simon, J. & Wong, P.-H. (2021) "Mapping the Stony Road toward Trustworthy AI: Expectations, Problems, Conundrums", in Pelillo, M. & Scantamburlo, t. (Hg.) *Machines We Trust: Perspectives on Dependable AI*, Cambridge, MA, The MIT Press, S. 27–40.
- Rojas-Galeano, S., Posada, J. & Ordoñez, E. (2022) "A Bibliometric Perspective on AI Research for Job-Résumé Matching", *The Scientific World Journal*, No. 1.
- Rouvroy, A. (2020) "Algorithmic Governmentality and the Death of Politics", *Green European Journal* [Online]. Verfügbar unter <https://www.greeneuropeanjournal.eu/algorithmic-governmentality-and-the-death-of-politics/> (Abgerufen am 20 August 2024).
- Rueckert, M. & Riedl, M. (2022) „Human-in-the-Loop: Wie Mensch und KI Aufgaben besser lösen“, *Digitale Welt*, Vol. 6, No. 4, S. 36–39.
- Samek, W, Montavon, G, Vedaldi, A, Hansen, LK & Müller, K-R (Hg.) (2019) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, Cham, Springer.

- Santoro, M., Marino, D. & Tamburrini, G. (2008) "Learning robots interacting with humans: from epistemic risk to responsibility", *AI & Society*, Vol. 22, No. 3, S. 301–314.
- Schlude, A., Schwind, M., Mendel, U., Stürz, R. A., Harles, D. & Fischer, M. (2023) *Verbreitung und Akzeptanz generativer KI in Deutschland und an deutschen Arbeitsplätzen* [Online], bidt. Verfügbar unter <https://www.bidt.digital/publikation/verbreitung-und-akzeptanz-generativer-ki-in-deutschland-und-an-deutschen-arbeitsplaetzen/> (Abgerufen am 27 Juni 2024).
- Schweighöfer, S. C. & Pfannstiel, M. A. (2022) „Künstliche Intelligenz im Entwicklungsprozess von Medikamenten in der Pharmaindustrie“, in Pfannstiel, M. A. (Hg.) *Künstliche Intelligenz im Gesundheitswesen: Entwicklungen, Beispiele und Perspektiven*, Wiesbaden, Springer Gabler, S. 139–151.
- Searle, J. (1980) "Minds, brains, and programs", *The Behavioral And Brain Sciences*, Vol. 3, No. 3, S. 417–457.
- Shneiderman, B. (2020) "Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems", *ACM Transactions on Interactive Intelligent Systems (TiiS)*, Vol. 10, No. 4, S. 1–31.
- Siddique, S., Haque, M. A., George, R., Gupta, K. D., Gupta, D. & Faruk, M. J. H. (2024) "Survey on Machine Learning Biases and Mitigation Techniques", *Digital*, Vol. 4, No. 1, S. 1–68.
- Sikder, M. N. K. & Batarseh, F. A. (2023) "Outlier detection using AI: a survey", in Batarseh, F. A. & Freeman, L. J. (Hg.) *AI Assurance: Towards Trustworthy, Explainable, Safe, and Ethical AI*, London, Academic Press, S. 231–291.
- Spielkamp, M. (2019) *Mind The Algorithm* [Online]. Verfügbar unter <https://algorithmwatch.org/en/mind-the-algorithm/> (Abgerufen am 30 August 2024).
- Spindler, M. & Kögel, H. (2020) *Erkennung von Versicherungsbetrug mit künstlicher Intelligenz* [Online], Berlin. Verfügbar unter https://www.bitkom.org/sites/default/files/2020-08/200817_sof9_versicherungsbetrug.pdf (Abgerufen am 11 Juli 2024).
- Statista (2024a) *Generative artificial intelligence (AI) market size worldwide from 2020 to 2030* [Online]. Verfügbar unter <https://www.statista.com/forecasts/1449838/generative-ai-market-size-worldwide> (Abgerufen am 27 Juni 2024).
- Statista (2024b) *Global market size of the artificial intelligence (AI) robot market from 2020 to 2030* [Online]. Verfügbar unter <https://www.statista.com/forecasts/1449861/ai-robot-market-worldwide> (Abgerufen am 27 Juni 2024).
- Statista (2024c) *Size of computer vision market worldwide from 2020 to 2030* [Online]. Verfügbar unter <https://www.statista.com/forecasts/1449859/worldwide-computer-vision-market-size> (Abgerufen am 27 Juni 2024).

- Statista (2024d) *Size of the generative artificial intelligence (AI) market in Europe from 2020 to 2030* [Online]. Verfügbar unter <https://www.statista.com/forecasts/1449848/generative-ai-market-size-europe> (Abgerufen am 27 Juni 2024).
- Statista (2024e) *Size of the natural language processing (NLP) market worldwide from 2020 to 2030* [Online]. Verfügbar unter <https://www.statista.com/forecasts/1449869/nlp-market-size-world> (Abgerufen am 27 Juni 2024).
- Thormundsson, B. (2024) *Expected impact of artificial intelligence (AI) on the world between 2025 to 2028, by life aspect* [Online], Statista. Verfügbar unter <https://www.statista.com/statistics/1449200/ai-impact-of-life-aspects-globally/> (Abgerufen am 27 Juni 2024).
- Turing, A. M. (1937) "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society*, s2-42, No. 1, S. 230–265.
- Turing, A. M. (1950) "Computing Machinery and Intelligence", *Mind*, Vol. 59, No. 236, S. 433–460.
- Ulich, E. (2013) „Arbeitssysteme als Soziotechnische Systeme – eine Erinnerung“, *Psychologie des Alltagshandelns*, Vol. 6, No. 1, S. 4–12.
- van den Berg, G. J., Kunaschk, M., Lang, J., Stephan, G. & Uhlendorff, A. (2024) *Predicting Re-Employment: Machine Learning Versus Assessments by Unemployed Workers and by Their Caseworkers* (IAB-Discussion Paper) [Online], Nürnberg, Institut für Arbeitsmarkt- und Berufsforschung.
- Winkler, J. K. & Haenssle, H. A. (2022) „Bildererkennung mittels künstlicher Intelligenz in der Hautkrebsdiagnostik“, *Die Dermatologie*, Vol. 73, S. 838–844.
- Wirtz, B. W., Weyerer, J. C. & Geyer, C. (2019) "Artificial Intelligence and the Public Sector: Applications and Challenges", *International Journal of Public Administration*, Vol. 42, No. 7, S. 596–615.
- Wissenschaftlicher Beirat für Digitale Transformation der AOK Nordost (2018) *Anmerkungen zu einer KI-Strategie für eine gesetzliche Krankenkasse* [Online]. Verfügbar unter <https://www.aok.de/pp/nordost/wissenschaftlicher-beirat/aok-nordost-entwickelt-strategie-zum-thema-kuenstliche-intelligenz/> (Abgerufen am 16 Juli 2024).
- Zaber, M., Casu, O. & Brodersohn, E. (2024) *Artificial Intelligence in Social Security Organizations* [Online], Genf, International Social Security Association; United Nations University.
- Zajko, M. (2023) "Automated Government Benefits and Welfare Surveillance", *Surveillance & Society*, Vol. 21, No. 3, S. 246–258.

Zandt, F. (2023) *Die größten Befürchtungen in Bezug auf KI* [Online]. Verfügbar unter <https://de.statista.com/infografik/30013/risiken-von-kuenstlicher-intelligenz/> (Abgerufen am 26 August 2024).

Zarges, L., Garnitz, J., von Maltzan, A. & Wohlrabe, K. (2023) *Der Investitionsstandort Deutschland aus Sicht der Familienunternehmen: Jahresmonitor der Stiftung Familienunternehmen* [Online], München.

