

Schaller, Daria et al.

Article

Künstliche Intelligenz: Chance oder Gefahr? Wie verändert der Einsatz von KI unsere Gesellschaft?

ifo Schnelldienst

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Schaller, Daria et al. (2023) : Künstliche Intelligenz: Chance oder Gefahr? Wie verändert der Einsatz von KI unsere Gesellschaft?, ifo Schnelldienst, ISSN 0018-974X, ifo Institut - Leibniz-Institut für Wirtschaftsforschung an der Universität München, München, Vol. 76, Iss. 08, pp. 03-28

This Version is available at:

<https://hdl.handle.net/10419/279729>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Künstliche Intelligenz: Chance oder Gefahr?

Wie verändert der Einsatz von KI unsere Gesellschaft?

Es ist unbestritten, dass der Einsatz von Künstlicher Intelligenz (KI) die Gesellschaft, insbesondere die Arbeitswelt, verändern wird. Große Unsicherheit besteht aber über das »Wie«. Ist ihr Einsatz eher eine Chance oder doch eher eine Gefahr? Auf der einen Seite werden größere Effizienz, Dynamik und neue Geschäftsmodelle erwartet. Auf der anderen Seite sind mit diesen Möglichkeiten Ängste verbunden – für die Wirtschaft wie auch für die Gesellschaft. Entwickelt sich die KI zum Jobkiller? Wird durch die Komplexität und mangelnde Nachvollziehbarkeit dieser Technologie die KI sogar unkontrollierbar? Ein geeigneter Rahmen für den Umgang mit KI-Technologien ist unerlässlich.

Daria Schaller, Klaus Wohlrabe und Anna Wolf

KI, Cloud Computing und Blockchain – wo steht die deutsche Wirtschaft?

Digitale Technologien verändern nicht nur die Effizienz und den Ablauf von Produktionsprozessen, sondern haben auch einen tiefgreifenden, disruptiven Einfluss auf unsere Wirtschaft, der oft mit dem Begriff »digitaler Wandel« umfasst wird. Der digitale Wandel bezeichnet die Integration digitaler Technologien in die wirtschaftlichen Arbeitsabläufe, aber auch ihre Auswirkungen auf die Lebensverhältnisse und die Gesellschaft insgesamt. Zu den zahlreichen neuen Technologien, die in den vergangenen Jahren einen beachtlichen Fortschritt und eine breite Akzeptanz erfahren haben, gehören digitale Plattformen, das Internet der Dinge (Internet of Things, IoT), Robotik, Cloud Computing, Blockchain und nicht zuletzt Künstliche Intelligenz. Durch ihre breitgefächerte Anwendung und tiefreichende Verbreitung entstehen innovative Produkte,

Dienstleistungen und Geschäftsmodelle, die in einer Vielzahl von Branchen, von Logistik und Energie über Landwirtschaft, Handel, Telekommunikation, Finanzdienstleistungen, Verarbeitendes Gewerbe bis hin zur Gesundheitsversorgung, Anwendung finden. Sie haben gar das Potenzial, das Leben der Menschen und die Gesellschaft nachhaltig zu verändern.

Cloud Computing, Künstliche Intelligenz und Blockchain zählen heute zu den wichtigsten Innovationen im Bereich der Informations- und Kommunikationstechnologie. Diesen drei scheinbar sehr unterschiedlichen Technologien ist es gemein, dass sie auf eine Modernisierung und Optimierung der Dateninfrastruktur in Unternehmen abzielen, indem sie eine effiziente Speicherung, Verarbeitung und Analyse großer Datenmengen ermöglichen. Alle drei Techno-



Daria Schaller

ist Fachreferentin im ifo Zentrum für Makroökonomik und Befragungen.



Dr. Klaus Wohlrabe

ist Stellvertretender Leiter des ifo Zentrums für Makroökonomik und Befragungen und Leiter Befragungen.



Dr. Anna Wolf

ist Fachreferentin im ifo Zentrum für Industrieökonomik und neue Technologien.

logien haben das Potenzial, bestehende Geschäftsmodelle und die Art und Weise, wie Wissen, Produkte und Dienstleistungen entstehen und ausgetauscht werden, maßgeblich zu verändern.

CLOUD COMPUTING

Cloud Computing hat die Art und Weise, wie Unternehmen, Privatpersonen und der öffentliche Sektor IT-Ressourcen nutzen, grundlegend verändert und sich als eine wichtige IT-Infrastruktur im modernen, digitalen Arbeiten etabliert. Beim Cloud Computing werden die IT-Ressourcen (wie Speicherplatz und Rechenleistung) den Nutzern über das Internet, meist auf Basis von nutzungsabhängigen Gebühren, zur Verfügung gestellt. Das Prinzip des Cloud Computing ermöglicht neben Aufwands- und Kostenersparnis, eine höhere Flexibilität, Sicherheit, Skalierbarkeit und Effizienz beim Einsatz von Rechenleistung und Datenspeicherung. Die wohl bekanntesten Formen des Cloud Computing sind die Speicherdienste der US-Technologiekonzerne Google (Google Drive), Amazon (Amazon Web Services – AWS) und Microsoft (OneDrive). Für das Cloud Computing gibt das politische Programm der EU »Europas digitale Dekade« konkrete Zielvorgaben. Demnach soll bis zum Jahr 2030 Cloud Computing bei mindestens 75% der Unternehmen der EU im Einsatz sein.

KÜNSTLICHE INTELLIGENZ

KI-basierte Systeme nutzen Techniken wie maschinelles Lernen und Deep Learning, um mit Hilfe großer Datenmengen komplexe Aufgaben wie Mustererkennung, Sprachverarbeitung, Trendanalyse und Entscheidungsfindung zu automatisieren. Höhere Rechenleistungen, Verfügbarkeit großer Datenmengen und neue Algorithmen haben zu einer rasanten Weiterentwicklung der KI-Technologie geführt und breite Anwendung über sämtliche Wirtschaftsbereiche hinweg ermöglicht. Zuletzt hat insbesondere die Verbreitung von Chatbot-Anwendungen, wie beispielsweise ChatGPT des Unternehmens OpenAI, KI in den Fokus des öffentlichen Interesses gerückt. Seit 2019 steht KI auf der Prioritätenliste der europäischen Industriepolitik (Europäisches Parlament 2020).

BLOCKCHAIN

Blockchain basiert auf der Distributed Ledger Technologie (DLT) und beinhaltet Algorithmen und Systeme für digitale, unveränderbare und dezentrale Dateninfrastrukturen. Blockchains können zu einer vertrauenswürdigen und manipulationssicheren Aufbewahrung unterschiedlicher Informationen eingesetzt werden. Ursprünglich als Grundlage für Kryptowährungen (Bitcoin) entwickelt, hat sich die Blockchain-Technologie zu einer vielseitigen DLT weiterentwickelt, die das Potenzial hat, Sicherheit, Transparenz und Effizienz von digitalen Transaktionen zu verbessern

(Catalini et al 2019). Neben den gängigen öffentlichen Blockchains, die global von zahlreichen Teilnehmern betrieben werden (wie Bitcoin, Ethereum u.a.), gibt es unternehmenseigene DLT, die primär im betriebswirtschaftlichen und industriellen Bereich für einen bestimmten Zweck verwendet werden. Typische Anwendungsfelder sind im Supply Chain Management, Finanzwesen oder in der öffentlichen Verwaltung. Blockchain-Technologie ist ebenfalls ein wichtiges Thema innerhalb der digitalen Strategie der EU. So baut der europäische öffentliche Sektor eine eigene Infrastruktur für Blockchain-Dienste auf, die eine Interoperabilität mit Plattformen des privaten Sektors ermöglichen sollen (Europäische Kommission 2023).

In diesem Beitrag soll anhand von empirischen Daten aus der ifo Konjunkturumfrage die Verbreitung der digitalen Technologien – KI, Blockchain und Cloud Computing – in der deutschen Wirtschaft untersucht werden. Die Ergebnisse basieren auf einer repräsentativen Umfrage, die das ifo Institut im Rahmen seiner monatlichen Konjunkturumfragen im Juni 2023 durchführte. Die Frage nach der Einstellung der Unternehmen zu diesen digitalen Technologien wurde im Auftrag des Hanseatic Blockchain Institute e.V. erhoben. Die Erhebung erfolgte als Teil der vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) geförderten W3NOW-Studie, die vom Hanseatic Blockchain Institute durchgeführt wird, um die Durchdringung der Blockchain-Technologie in der deutschen Wirtschaft zu untersuchen.¹

ERGEBNISSE DER IFO KONJUNKTURUMFRAGE

In Tabelle 1 sind die Ergebnisse für die drei Schlüsseltechnologien für die jeweiligen Sektoren aufgelistet. Neben den vier abgefragten Kategorien wird auch die Summe aus »im Einsatz« und »geplant« dargestellt. Dies kann als aktuelle *Durchdringungsquote* interpretiert werden. Detaillierte Ergebnisse für die Industrie und den Dienstleistungssektor sind in Abbildung 1 dargestellt. Es zeigt sich, dass gegenwärtig Cloud Computing am weitesten verbreitet ist. 46,5% der Unternehmen nutzen diese bereits, und weitere 11% planen das in naher Zukunft. Die entsprechenden Anteile sind bei Künstlicher Intelligenz (13,3 bzw. 9,2%) und Blockchain (3,2 bzw. 3,7%) teilweise deutlich geringer. Dies könnte daran liegen, dass Cloud Computing für die breite Masse schon länger zur Verfügung steht als die beiden anderen Technologien. Für alle drei Technologien gilt, dass sie deutlich häufiger in Großunternehmen im Einsatz sind, im Vergleich zu kleineren und mittelständischen Unternehmen. Min-

¹ Die ifo Konjunkturumfragen basieren auf ca. 9 000 monatlichen Meldungen von Unternehmen des Verarbeitenden Gewerbes, des Dienstleistungssektors, des Handels und des Bauhauptgewerbes. Die Sonderfrage haben rund 6 600 Unternehmen beantwortet. Die Ergebnisse berechnen sich auf Basis der gewichteten Anteile der Unternehmen, die auf die Antwortoptionen »im Einsatz«, »geplant«, »diskutiert« und »kein Thema« der entsprechenden Frage in der ifo Konjunkturumfrage entfallen.

destens eine der drei Technologien wird gegenwärtig bei 42% der Firmen eingesetzt.

Ein Blick auf die Sektoren zeigt, dass beim Cloud Computing vor allem die Unternehmen aus der Industrie und dem Dienstleistungssektor vorangehen. Dort berichten rund 60% von entsprechenden Einsatzmöglichkeiten. In der Industrie sind es vor allem die Pharmaindustrie und die Autobauer, und bei den Dienstleistern sind es vor allem die unternehmensnahen Dienstleistungen, wie die IT-Unternehmen, Unternehmen der Telekommunikationsbranche sowie Rechts-, Steuer und Unternehmensberatungen, die Cloud Computing bereits einsetzen. Im Handel und im Bausektor liegen die entsprechenden Werte unter 50%. Lediglich für etwa ein Viertel der Unternehmen spielt das Thema »Cloud Computing« gegenwärtig keine Rolle.

Die Industrie ist auch Vorreiterin beim Einsatz von Künstlicher Intelligenz. Jedes dritte Industrieunternehmen nutzt diese bereits oder plant ihren Einsatz. Bei den Dienstleistern und im Handel liegt der Wert bei 20%, im Baugewerbe bei 15%. Wie bereits bei Cloud Computing sind es vor allem die Automobil- und die Pharmaindustrie, die sich mit diesem Thema aktiv beschäftigen, aber auch die Hersteller von Elektro- und Informations- und Kommunikationstechnik sowie die Maschinenbauer. Innerhalb des Dienstleistungssektors ist die KI-Technologie vor allem bei den unternehmensnahen Dienstleistungen im Einsatz, oder es wird darüber diskutiert (vgl. Abb. 1). Bei den IT-Dienstleis-

tungen gibt es nahezu kein Unternehmen, das sich nicht mit dem Thema beschäftigt oder keine Anwendungsmöglichkeiten sieht. Das ist nicht verwunderlich, denn große Sprachmodelle, wie das von ChatGPT, wurden auf riesigen Codebeständen trainiert, um unter anderem auch Software-Code zu generieren. Das Thema »Künstliche Intelligenz« ist allerdings noch nicht überall vorgedrungen. Rund 40% der Unternehmen geben an, dass KI gegenwärtig kein Thema ist. Im Baugewerbe liegt dieser Anteil bei 60%.

Die Blockchain-Technologie kommt gegenwärtig im Vergleich zu den anderen beiden Technologien auf eine geringere Anwendungsquote. Rund 7% der Unternehmen nutzen sie oder planen es, rund 19% diskutieren eine mögliche Einführung. Für rund drei Viertel der befragten Unternehmen ist Blockchain gegenwärtig noch kein Thema. Eine breitere Anwendung findet die Blockchain-Technologie vor allem in der Industrie, wo der Durchdringungswert bei 12% liegt, d.h., 12% der Unternehmen setzen bereits Blockchain-Technologie ein oder planen einen Einsatz, und fast jedes vierte Unternehmen (24%) diskutiert darüber. Besonders weit in Sachen Blockchain ist die chemische Industrie. Dort nutzen 14% der Unternehmen die Technologie für ihre Geschäftsprozesse. In der Automobilindustrie setzen 13% der Unternehmen Blockchain-Technologie ein, und rund jedes fünfte Unternehmen plant den Einsatz. Bei den Herstellern von Nahrungs- und Futtermitteln, Druckerzeugnissen, Datenverarbeitungsgeräten, elektrischen Ausrüstungen, Möbeln und im

Tab 1

KI, Blockchain und Cloud Computing in Unternehmen^a

	Deutschland	Industrie	Dienstleistungen	Großhandel	Einzelhandel	Bauhauptgewerbe
Künstliche Intelligenz						
im Einsatz	13,3	17,3	12,2	9,1	12,2	7,1
geplant	9,2	12,9	7,2	10,0	7,3	7,9
Summe	22,5	30,2	19,4	19,1	19,5	15,0
diskutiert	36,7	38,7	38,4	35,2	25,8	25,3
kein Thema	40,8	31,1	42,2	45,7	54,7	59,7
Blockchain						
im Einsatz	3,2	5,5	2,0	2,2	2,9	2,7
geplant	3,7	6,5	2,4	2,6	2,0	2,7
Summe	6,9	12,0	4,4	4,8	4,9	5,4
diskutiert	18,7	24,0	16,1	18,3	15,0	17,0
kein Thema	74,5	64,0	79,6	76,9	80,0	77,6
Cloud Computing						
im Einsatz	46,5	48,7	49,0	39,1	35,7	32,9
geplant	11,1	9,8	12,6	10,9	6,5	9,4
Summe	57,5	58,5	61,6	50,0	42,2	42,3
diskutiert	18,2	21,6	15,5	23,0	16,1	20,0
kein Thema	24,3	19,8	22,9	27,1	41,7	37,8

^a Anteile der jeweiligen Antwortkategorie in%.

Quelle: ifo Konjunkturumfragen, Juni 2023.

Maschinenbau wird die Blockchain-Technologie von mindestens jedem zehnten Unternehmen eingesetzt, oder ihr Einsatz ist geplant. Im Dienstleistungssektor, im Handel und Bauhauptgewerbe sind es hingegen deutlich weniger. Dort schwanken die Werte um die 5%. Ausnahmen sind hier wieder die IT-Dienstleister, wo immerhin rund 15% der Unternehmen die Technologie nutzen oder den Einsatz planen.

Weniger als 1% der Unternehmen haben gegenwärtig alle drei Technologien gleichzeitig im Einsatz. Dieser Anteil erhöht sich geringfügig auf 2,5%, wenn der geplante Einsatz berücksichtigt wird. Mindestens zwei Technologien werden gegenwärtig von rund 8% der Unternehmen eingesetzt. Dabei ist insbesondere die Kombination Cloud Computing und KI am häufigsten. Der Anteil steigt auf rund 15%, wenn die

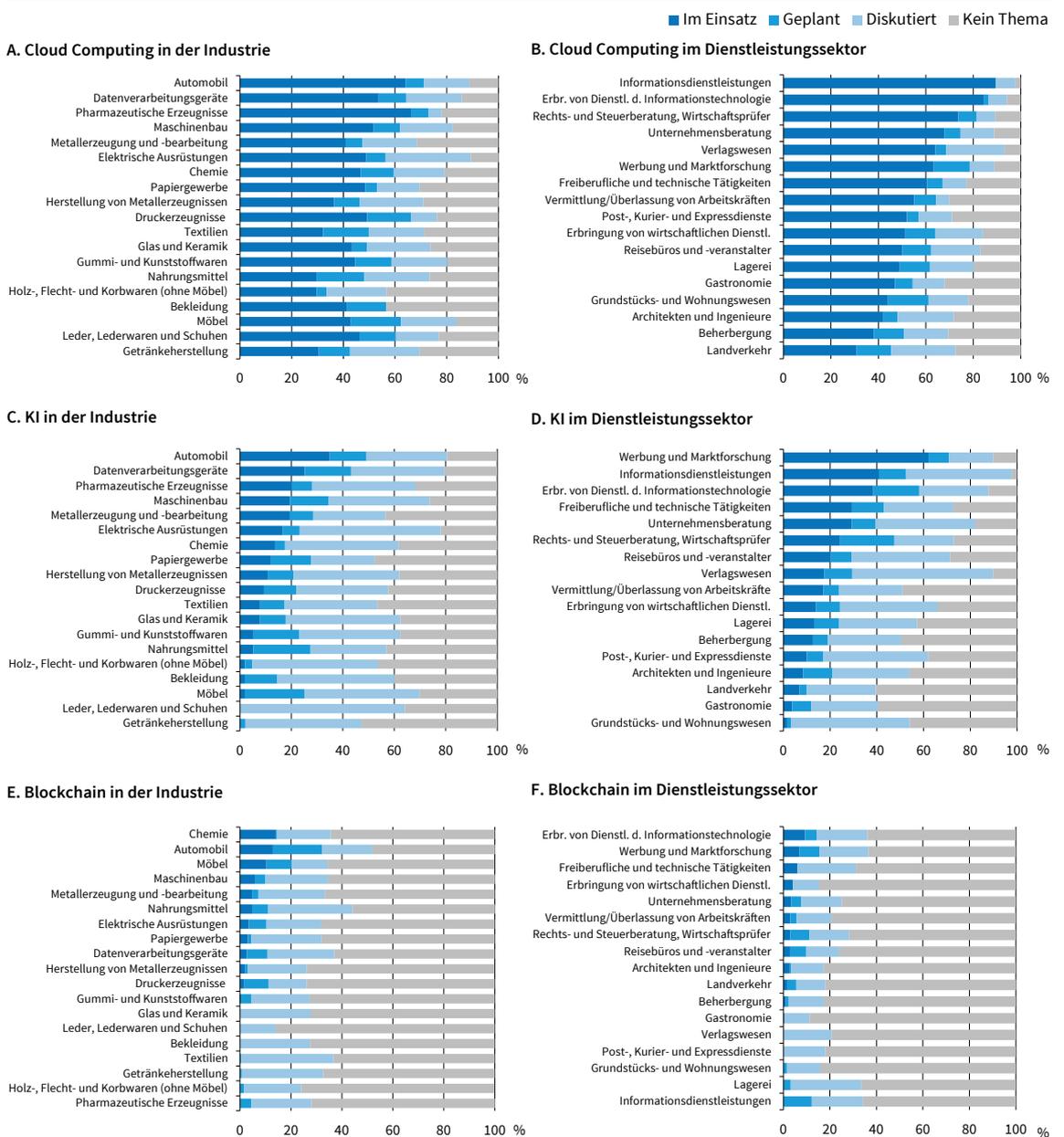
geplanten Einsätze hinzugerechnet werden. Bei rund drei Viertel der Unternehmen ist mindestens eine der Technologien in der Diskussion. Die Anzahl der Antworten verteilt sich gleichmäßig über die verschiedenen Sektoren.

Industrie und Dienstleistungssektor als Vorreiter

Zusammenfassend kann gesagt werden, dass die Industrie und der Dienstleistungssektor, hier vor allem die unternehmensnahen Dienstleistungen, Vorreiter beim Einsatz neuer Technologien sind. Der Handel und insbesondere der Bausektor zeigen hingegen noch eine gewisse Zurückhaltung. Wenig überraschend sind es vor allem größere Unternehmen, die sich mit ent-

Abb. 1

Status der Technologie nach Wirtschaftsbereich



Quelle: ifo Konjunkturumfragen, Juni 2023.

© ifo Institut

sprechenden Anwendungen beschäftigen oder bereits in der Praxis einsetzen. Eine relative Einordnung der Zahlen insgesamt und im Zeitablauf ist gegenwärtig noch nicht möglich, da es die erste umfangreiche Erhebung des ifo Instituts zur technologischen Durchdringung der deutschen Wirtschaft ist. Andere Erhebungen zu einzelnen Technologien aus den Vorjahren, etwa des Statistischen Bundesamts zur Verbreitung des Cloud Computing und der KI-Anwendungen aus dem Jahr 2021, deuten jedoch darauf hin, dass die technologische Durchdringung in der Wirtschaft seit 2021 deutlich zugenommen hat: Im Jahr 2021 nutzten 28 % der Unternehmen Cloud Computing und 11 % der Unternehmen Künstliche Intelligenz. Mit dieser Durchdringungsquote lag Deutschland im europäischen Mittelfeld. Zur Verbreitung der DLT bzw. Blockchain-Technologie liegen derzeit keine Daten der amtlichen Statistik vor. Das ifo Institut plant, in einem Jahr die Unternehmen erneut nach dem Einsatz der digitalen Technologien zu befragen, um die Veränderung der Durchdringungsquote im zeitlichen Verlauf zu untersuchen.

DISKUSSION

Zwar lassen die Umfrageergebnisse keine Schlüsse zu, in welchen Bereichen die Technologien eingesetzt werden. Sie zeigen jedoch, dass die Einsatzbereiche von KI, Cloud Computing und Blockchain sowohl branchenspezifisch als auch branchenübergreifend sehr vielfältig sein dürften. Im Folgenden seien einige der sich bereits im Einsatz befindenden Anwendungen kurz vorgestellt.

Industrielle KI-Anwendungen bieten die Möglichkeit, mit Hilfe des maschinellen Lernens große Datenmengen aus industriellen Prozessen effizient zu verarbeiten und zu analysieren. KI kann z.B. verwendet werden, um mit Hilfe von vernetzten Geräten und Anlagen sowie gesammelten Daten industrielle Prozesse zu überwachen und zu steuern, um rechtzeitig Wartung, Reparatur oder Umrüstung einzuleiten (Hecker 2017). Laut einer Studie unter Unternehmen im produzierenden Gewerbe, die KI bereits nutzen, gaben 34 % an, KI für die Qualitätssteigerung von Produkten und Dienstleistungen einzusetzen (Harlacher et al. 2023). Auch findet KI vermehrt in der industriellen Analyse Anwendung. KI-gestützte Lösungen helfen bspw. Fahrzeugherstellern dabei, Produktionsprozesse zu verbessern, aber auch bei der Weiterentwicklung der Fahrsicherheit und des autonomen Fahrens. Dafür nutzen KI-Systeme die Möglichkeiten des IoT in Fahrzeugen. IoT-Systeme helfen auch bei der Verfolgung des Echtzeit-Zustands von Fahrzeugen, indem die fortlaufend anfallenden Fahrzeugdaten automatisch ausgewertet werden.

Auch im Pharma- und Biotechbereich finden zunehmend KI-Analysetools Anwendung. Zum Einsatz kommen hier unter anderem Deep Learning Modelle für diagnostische Anwendungen, die Bilder (z.B. Fotos

oder radiologische Bilder) in Kombination mit Pathologie- und Behandlungsdaten analysieren. Auch pharmazeutische Wissenschaftsliteraturdatenbanken können inzwischen mit KI-Methoden effizienter und schneller ausgewertet werden, was die Medikamentenentwicklung deutlich beschleunigen dürfte. Daneben kommt auch die Blockchain-Technologie im Gesundheitswesen zum Einsatz: So werden in Estland elektronische Patientenakten bereits auf einer Blockchain abgesichert. Unter anderem arbeitet das Unternehmen IBM an zahlreichen Blockchain-Lösungen für das Gesundheitswesen und die Biowissenschaften (IBM 2023).

Ein branchenübergreifendes Einsatzgebiet für die digitalen Technologien ist die Steuerung und Überwachung von *Lieferketten*. Nicht zuletzt durch die Erfahrung der Lieferengpässe der vergangenen Jahre ist es für viele Hersteller wichtig, jede Etappe des Liefer- und Herstellungsprozesses zu überwachen und genau zu wissen, wann die benötigten Komponenten im Produktionsstandort eintreffen. Aus diesem Grund stützen sich moderne Lieferketten inzwischen auf IoT-, Blockchain- und KI-Technologien. Im Fall des Einsatzes der KI werden auf Basis von datengestützten Annahmen und Vorhersagen automatisierte Anpassungen in Echtzeit möglich (Schünemann 2023). Dadurch kann die Transparenz von Lagerbewegungen verbessert und mögliche Lieferunterbrechungen besser vorhergesagt werden. Laut einer Studie von Bitkom nutzte im Oktober 2022 bereits jedes fünfte Logistikunternehmen in Deutschland KI, weitere 26 % planten dies oder diskutierten darüber (Bitkom 2022). Auch die Blockchain-Technologie wird in den Unternehmen zur Rückverfolgbarkeit von Vorprodukten eingesetzt, meistens mit den Zielen der Kostensenkung, Verbesserung der Liefergeschwindigkeit, Wahrnehmung der unternehmerischen Sorgfaltspflichten oder der Nachhaltigkeit bzw. des Umweltschutzes. Durch den Einsatz der Blockchain-Technologie können zwischenbetriebliche Prozesse, die sonst mit viel Personalaufwand verbunden sind, automatisiert werden. Im Falle eines Rückrufs können so Problemquellen und betroffene Produkte schneller identifiziert werden, und sowohl Zutatenherkunft als auch Lieferwege von Produktchargen schneller nachvollzogen werden. So können Rückrufe zielgenauer und schneller durchgeführt werden. Lückenlose Rückverfolgbarkeit von Vorprodukten ist ein wichtiges Thema etwa in der Nahrungs- und Futtermittelindustrie, in der Chemie- und Pharmaindustrie, aber auch im Automobilbau.

Im *Einzelhandel* und bei den *personenbezogenen Dienstleistungen* kommen KI und Blockchain häufig in der Front-End-Kommunikation zum Einsatz (Fraunhofer 2022). Mit Hilfe von KI-Assistenten werden etwa personalisierte, auf den Kunden zugeschnittene Angebote (Recommender Systeme²) in den Bereich Produkt, Content und Service angeboten. Die auf Block-

² Recommender Systeme sind Algorithmen, die Interessen von Nutzern quantifizieren und auf Grundlage historischer Daten Vorhersagen treffen.

chain basierenden Smart Contracts und Non-fungible Token (NFT) werden als eine Form von Eigentumsnachweis bzw. Echtheitszertifikat für digitale oder auch physische Güter verwendet.

Ein weiteres bereits genutztes und branchenübergreifendes Einsatzgebiet von KI und Blockchain in Unternehmen ist die *Finanzplanung und -verwaltung*. Banken und Versicherungen, aber auch andere Unternehmen nutzen KI in der intelligenten Betrugs-erkennung. Auch hier liegen selbstlernende KI-Verfahren zugrunde, die mit historischen Daten trainiert werden. Zur Anwendung kommen diese sowohl für Internal Fraud, wie betrügerische Handlungen eigener Mitarbeiter in der Buchhaltung, als auch für External Fraud, z.B. Kreditkartenbetrug (Trabold et al. 2021). Auch Blockchain-Anwendungen werden zum Zwecke der Transparenz und Rückverfolgbarkeit von Zahlungsströmen bzw. zur gesicherten Übertragung von Vermögens- und Vertragsdaten verwendet.

GENERAL PURPOSE TECHNOLOGY?

Mehrere Forschungsarbeiten haben sich mit der Frage befasst, ob KI, Blockchain und Cloud Computing die Anforderungen erfüllen, um als General Purpose Technology (GPT) eingestuft zu werden. GPT sind bestimmte Schlüsseltechnologien, die wirtschaftliche und soziale Strukturen so gravierend verändert haben, dass dadurch ganze Epochen technologischen Fortschritts und Wirtschaftswachstums eingeleitet wurden. Typischen Beispiele für GPT sind Elektrizität, die Erfindung des Telefons, des Computers und des Internets. Hauptmerkmale der GPT sind, dass sie branchenübergreifend wirken, zu Kostensenkungen und Produktivitätsgewinnen sowie Produkt- und Prozessinnovationen beitragen. Im Fall Blockchain haben Catalini et al. (2019) zwei zentrale Kostenpositionen, die durch die Blockchain-Technologie erheblich gesenkt werden können, identifiziert: Verifizierungskosten (costs of verification) und Netzwerkeffektkosten (cost of networking). Für KI diskutierte Crafts (2021) über eine mögliche Einstufung als GPT, da KI die Produktivität der Forschung und Entwicklung auf zahlreichen Gebieten erhöht und damit die Grundlage für eine tiefgreifende globale Transformation der Wirtschaft legt. Auch die empirischen Ergebnisse der ifo Konjunkturumfrage stützen die These, dass die diskutierten digitalen Technologien angesichts ihres breiten branchenspezifischen sowie branchenübergreifenden Einsatzes das Potenzial haben, sich zu GPT zu entwickeln.

REGULIERUNGSBESTREBUNGEN VON KI, BLOCKCHAIN UND CLOUD COMPUTING SEITENS DER POLITIK

Bereits im Jahr 2018 hat die Bundesregierung eine »Strategie Künstliche Intelligenz« sowie im Jahr 2019 eine umfassende Blockchain-Strategie vorgelegt. Auch

der konsequente *Einsatz von Cloud Computing* wird als zentrale Basistechnologie für die digitale Transformation und damit zur Sicherung der Wettbewerbsfähigkeit der deutschen Wirtschaft gesehen.

Die europäischen Regulierungsbehörden haben sich in den vergangenen Jahren intensiv mit dem Einsatz von KI, Blockchain und Cloud Computing auseinandergesetzt. Diese Bestrebungen zielen darauf ab, den Schutz der Bürgerrechte, die Sicherheit und die ethische Verwendung dieser Technologien sicherzustellen. Im Bereich KI hat sich das EU-Parlament im Juni 2023 auf einen Vorschlag für ein Gesetz geeinigt, das unter anderem eine Einstufung verschiedener KI-Anwendungen und ihrer Risiken vorsieht. Zu »Hochrisiko-KI-Systemen« gehören solche, die zum Beispiel in den Bereichen der biometrischen Identifizierung (z.B. Gesichtserkennung), der kritischen Infrastruktur wie Straßenverkehr, Wasser-, Gas-, Wärme- und Stromversorgung oder in der Strafverfolgung verwendet werden. Das Europäische Parlament will hiermit sicherstellen, dass die in der EU eingesetzten KI-Systeme sicher, transparent, nichtdiskriminierend, nachvollziehbar und nachhaltig sind. Die Gesetzesvorlagen werden aktuell noch kontrovers diskutiert. Bis Ende 2023 soll jedoch eine Einigung erzielt werden.

Im Bereich Blockchain hat die europäische Politik ebenfalls die Notwendigkeit erkannt, klare rechtliche Rahmenbedingungen zu schaffen. Dazu gehören neben noch laufenden Arbeiten zur Erfassung möglicher Risiken für die Finanzstabilität auch konkrete gesetzgeberische Maßnahmen auf europäischer und nationaler Ebene. 2022 hat die Europäische Union mit der Verordnung über Märkte für Kryptowerte (Regulation on Markets in Crypto-Assets – MiCA), der Transfer of Funds Regulation (TFR) und auch der Verordnung zur Schaffung eines sogenannten DLT Pilot Regimes einen europäischen Rahmen zur Behandlung von Kryptowerten geschaffen. Allerdings sind auch hier noch bei weitem nicht alle Bereiche der Blockchain-Ökonomie abgedeckt. So beinhaltet MiCA noch keine regulatorische Einordnung der NFTs. Auch Kryptowerte, die als Finanzinstrumente eingestuft werden, fallen noch nicht unter den Geltungsbereich der MiCA-Vorschriften.

Im Bereich Cloud Computing hat die EU ebenfalls regulatorische Maßnahmen ergriffen, um den Schutz der Privatsphäre und den sicheren Datentransfer zu gewährleisten. Regulatorische Fragen ergeben sich hier vor allem aus den bestehenden Datenschutz-Anforderungen, etwa der europäischen Datenschutz-Grundverordnung (DSGVO). Die DSGVO, die im Mai 2018 in Kraft trat, legt strenge Regeln für den Umgang mit personenbezogenen Daten fest, einschließlich deren Speicherung und Übertragung in die Cloud.

Während ein verbindlicher europäischer Gesetzesrahmen für die Verwendung der digitalen Technologien erst im Entstehen ist, wächst mit einer zunehmenden Marktdurchdringung der Ruf nach einer rechtlichen Einordnung. Denn erst wenn Rechtssicherheit über Einsatz und Verwendung vorliegt, wird

die Wirtschaft vermehrt auf digitale Technologien setzen können.

ZUSAMMENFASSUNG

KI, Blockchain und Cloud Computing sind digitale Technologien, die den Unternehmen wesentliche Kosten, Zeit- und Prozessvorteile ermöglichen und damit als Treiber für eine Vielzahl von volkswirtschaftlich relevanten Ökosystemen dienen können. Ihre ökonomischen Auswirkungen gehen jedoch weit über den Aspekt der Kosten- und Prozessoptimierung hinaus, denn sie ermöglichen ganz neue Geschäftsmodelle und Geschäftsprozesse und schaffen damit ein neues Wertschöpfungspotenzial. Der digitale Wandel ist somit im vollen Gange und stellt die Unternehmen und die Politik vor große Herausforderungen. Während ein verbindlicher europäischer Gesetzesrahmen für den Einsatz und die Verwendung der digitalen Technologien erst im Entstehen ist, ist die Wirtschaft in Deutschland bereits inmitten einer digitalen Transformation und damit auch im globalen Wettbewerb um Technologieführerschaft und Technologiesouveränität.

REFERENZEN

Bitkom (2022), »Jedes fünfte Logistikunternehmen setzt Künstliche Intelligenz ein«, verfügbar unter: <https://www.bitkom.org/Presse/Presseinformation/Logistikunternehmen-Kuenstliche-Intelligenz-KI>, aufgerufen am 27. Juni 2023.

Bundesministerium für Wirtschaft und Energie (2019), »Blockchain-Strategie der Bundesregierung – Wir stellen die Weichen für die Token-Ökonomie«, verfügbar unter: <https://www.bmwk.de/Redaktion/DE/Publikationen/Digitale-Welt/blockchain-strategie.html>, aufgerufen am 21. Juni 2023.

Catalini, C. und J. Gans (2019), »Some Simple Economics of the Blockchain«, *MIT Sloan Research Paper* Nr. 5191-16.

Crafts, N. (2021), »Artificial intelligence as a general-purpose technology: an historical perspective«, *Oxford Review of Economic Policy* 37(3), 521–536.

Destatis (2022), »Nutzung von Cloud Computing nach Beschäftigtengrößenklassen«, <https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/IKT-in-Unternehmen-IKT-Branche/Tabellen/iktu-06-cloud-computing.html>, aufgerufen am 21. Juni 2023.

Die Bundesregierung (2018), *Strategie Künstliche Intelligenz der Bundesregierung*, verfügbar unter: <https://www.bundesregierung.de/resource/blob/997532/1550276/3f7d3c41c6e05695741273e78b-8039f2/2018-11-15-ki-strategie-data.pdf>, aufgerufen am 21. Juni 2023.

Europäische Kommission (2023), *Blockchain Strategie der EU*, verfügbar unter: <https://digital-strategy.ec.europa.eu/de/node/34>, aufgerufen am 27. Juni 2023.

Europäisches Parlament (2023), »KI-Gesetz: erste Regulierung der künstlichen Intelligenz«, verfügbar unter: <https://www.europarl.europa.eu/news/de/headlines/society/20230601STO93804/ki-gesetz-erste-regulierung-der-kuenstlichen-intelligenz>, aufgerufen am 23. Juni 2023.

Europäisches Parlament (2020), »Was ist künstliche Intelligenz und wie wird sie genutzt?«, verfügbar unter: <https://www.europarl.europa.eu/news/de/headlines/society/20200827STO85804/was-ist-kuenstliche-intelligenz-und-wie-wird-sie-genutzt>, aufgerufen am 27. Juni 2023.

Fraunhofer (2022), »KI revolutioniert die Supply Chain«, verfügbar unter: <https://www.iis.fraunhofer.de/de/profil/jb/2021/ki-revolutioniert-supply-chain.html>, aufgerufen am 27. Juni 2023.

Harlacher, M., N. Feggeler, Y. Pfeifer und N. Ottersböck (2023), »Produzierendes Gewerbe auf internationalem Niveau«, *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 118(3), 173–177.

Hecker, D., I. Döbel, S. Rüping und V. Schmitz (2017), Künstliche Intelligenz und die Potenziale des maschinellen Lernens für die Industrie, *Wirtschaftsinformatik & Management* 9, 26–35.

IBM (2023), »Blockchain-Lösungen für das Gesundheitswesen und die Biowissenschaften«, <https://www.ibm.com/de-de/blockchain/industries/healthcare>, aufgerufen am 24. Juli 2023.

Schünemann, H. (2023), »Kundenstory – Renfert GmbH«, verfügbar unter: <https://www.agile-im.de/2023/03/31/aim-kundenstory-renfert-gmbh/>, aufgerufen am 27. Juni 2023

Trabold, D., R. Blankertz und L. Schrader (2021), »Effiziente Betrugserkennung durch Maschinelles Lernen«, Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, verfügbar unter: https://www.iais.fraunhofer.de/content/dam/iais/gf/bda/Downloads/Fraunhofer_IAIS_Whitepaper_Fraud.pdf, aufgerufen am 28. Juni 2023.

Vera Demary und Armin Mertens

Zwischen Chance und Herausforderung: KI in Unternehmen

Künstliche Intelligenz (KI) polarisiert. Auf der einen Seite stehen Versprechen von Effizienz, Dynamik und neuen Geschäftsmodellen. Auch wenn es KI per se schon lange gibt – plötzlich wird Wirtschaft, Gesellschaft und Öffentlichkeit klar, was diese Versprechungen eigentlich bedeuten. Bislang war KI für viele, selbst für Unternehmen, oft abstrakt und weit weg. Dies änderte sich spätestens seit der breiten Verfügbarkeit von generativen KI-Sprachmodellen (LLM, Large Language Models) wie ChatGPT oder dessen Google-Äquivalent Bard. Jeder kann Arbeitsaufträge eingeben und verwundert, begeistert oder entsetzt feststellen, wie gut die KI diese meistern kann. Plötzlich ist klar, dass KI-Anwendungen ein »Game Changer« sein können.

Auf der anderen Seite sind mit diesen Möglichkeiten auch Ängste verbunden – für die Wirtschaft wie auch für die Gesellschaft. Diese beziehen sich vor allem auf zwei Aspekte: Zum einen besteht die Sorge vor Komplexität und mangelnder Nachvollziehbarkeit



Dr. Vera Demary

leitet das Cluster »Digitalisierung und Klimawandel« am Institut der deutschen Wirtschaft, Köln.



Dr. Armin Mertens

leitet das Cluster »Big Data Analytics« am Institut der deutschen Wirtschaft, Köln.

von Entscheidungen und Aktionen, die durch die KI ausgelöst werden. Diese müssen in der Konsequenz zwar nicht falsch im klassischen Sinne sein, können aber nicht intendierte oder nicht absehbare Effekte nach sich ziehen. Der zweite Aspekt ist die Sorge vor einer weitreichenden Veränderung von Arbeit, verbunden mit einer notwendigen Veränderung von Kompetenzen und Tätigkeiten bis hin zu dem Verlust von Arbeitsplätzen.

WIE UNTERNEHMEN IN DEUTSCHLAND KI NUTZEN

Vor diesem Hintergrund beschäftigen sich Unternehmen inzwischen umfassend mit KI. Laut Daten des KI-Monitors setzten im Jahr 2022 knapp 19% der Unternehmen aus Industrie und industrienahe Dienstleistungen KI ein (Rusche et al. 2022, S. 14 f.). Zwei Jahre zuvor waren es nur rund 10%. Diesen Trend zu einem Mehreinsatz der Technologie im unternehmerischen Umfeld bestätigen – aufgrund der Befragungsmethodik und Stichprobe auf einem geringeren Niveau – beispielsweise auch Daten des ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung (Rammer 2022, S. 11). Der KI-Monitor zeigt zudem auf, dass Unternehmen KI trotz der eingangs beschriebenen Herausforderungen im Zusammenhang mit der neuen Technologie eher als Chance, denn als Risiko für die Volkswirtschaft bewerten: Auf einer Skala von 0 (nur Risiko) bis 100 (nur Chance) liegt ihre Einschätzung im Jahr 2022 bei knapp 65 (Rusche et al. 2022, S. 14 f.).

Eine wesentliche Voraussetzung für den Einsatz von KI in Unternehmen oder gar deren eigene Entwicklung ist die betriebliche Fähigkeit zur Bewirtschaftung von Daten (siehe auch Engels 2023). Diese sogenannte Data Economy Readiness umfasst Kenntnisse des Unternehmens in den Bereichen Datenmanagement, Datenspeicherung und Datennutzung, die wesentlich für die erfolgreiche Beteiligung an Datenökosystemen und auch die Nutzung Künstlicher Intelligenz sind. Eine Befragung von Unternehmen aus Industrie und industrienahe Dienstleistungen im Jahr 2022 ergab, dass 31% »data economy ready« waren (Büchel und Engels 2022, S. 1 f.). Besonders kleine Unternehmen mit weniger als 50 Beschäftigten hatten die erforderlichen Kompetenzen noch nicht, was auch die Implementierung von KI im Unternehmen erschweren dürfte. Hoffnung gibt, dass sich die Data Economy Readiness im Vergleich zum Vorjahr – wenn auch nur leicht – verbessert hat. Im Jahr 2021 zählten noch 2 Prozentpunkte weniger zu den Unternehmen, die »data economy ready« sind.

WELCHE KI-KOMPETENZEN UNTERNEHMEN NACHFRAGEN

Für die Nutzung von KI in Unternehmen benötigen diese Mitarbeitende mit passenden Kompetenzen, oft in der Entwicklung, aber vor allem auch in der Anwen-

dung von KI. Der Einsatz von KI in Unternehmen verändert also auch den Arbeitsmarkt. Einerseits verändern sich bestehende Berufe und Tätigkeitsbereiche. Aufgaben, die zuvor manuell von Arbeitskräften erledigt wurden, können (stärker) automatisiert und sogar autonom übernommen werden. Andererseits können auch gänzlich neue Berufsbilder und -bezeichnungen entstehen. So wird etwa vom kürzlich erstmals nachgefragten »Prompt Engineer« erwartet, kreativ und effizient mit Eingaben für generative KI-Modelle wie ChatGPT oder Bard umgehen zu können (Büchel und Mertens 2023).

Diese Entwicklungen gehen häufig mit der eingangs beschriebenen Angst vor negativen Beschäftigungseffekten von KI einher, die im Aggregat bisher allerdings nicht nachgewiesen werden konnten (Acemoglu et al. 2022). Denn gleichzeitig können auch positive Effekte erwartet werden: Unternehmen können durch den Einsatz von KI befähigt werden, neue Geschäftsmodelle zu entwickeln und durch einen höheren Automatisierungsgrad Einsparungen zu erzielen und ihre Produktivität zu steigern. Dadurch können Kapazitäten entstehen, die auch in nicht-automatisierbaren Aufgabenbereichen eine erhöhte Nachfrage zur Folge haben (siehe Acemoglu und Restrepo 2018).

Die Nachfrage nach KI-Kompetenzen im deutschen Arbeitsmarkt hat in den letzten Jahren deutlich zugenommen. Auf der einen Seite ist der Anteil an Stellenanzeigen, in denen Kompetenzen rund um den Umgang mit Daten gefordert werden, seit 2019 kontinuierlich gewachsen (Büchel et al. 2023, S. 58). Datenkompetenzen, die als Grundlage für den Einsatz von KI gesehen werden können, werden mittlerweile in mehr als jeder fünften Stellenanzeige nachgefragt. Auf der anderen Seite wächst auch die Anzahl an KI-Stellengesuchen auf dem deutschen Arbeitsmarkt kontinuierlich (Büchel und Mertens 2022, S. 7 f.). Doch nicht nur die generelle Nachfrage nach KI-Fachkräften verändert sich, sondern auch die innerhalb dieser KI-Stellen nachgefragten Kompetenzen und Technologien. Die Dynamik dieser Veränderungen ist im Vergleich zu anderen Berufsbildern sehr hoch. Büchel und Mertens (2022) zeigen etwa, dass sich die Kompetenzanforderungen in KI-Berufen in nur wenigen Jahren deutlich verschoben haben. Besonders auffallend hierbei ist erhöhte Nachfrage nach Kenntnissen von Cloud-Technologien und -Plattformen im Zusammenhang mit Künstlicher Intelligenz.

Der Erwerb von KI-Kompetenzen erfolgt – wie auch bei anderen Kompetenzen – im Bildungssystem oder in der Praxis, in formellen Kursen oder informell »on the job«. Dies macht deutlich, dass den Bildungsinstitutionen eine besondere Rolle zukommt, die sich mit einem ansteigenden Bedarf an Kompetenzen im Umgang mit Daten und Modellen noch verstärken wird. Eine Modernisierung des Bildungssystems ist daher unerlässlich. Es müssen moderne Technologien, aber vor allem auch der Umgang damit, bereits ab

den frühen Bildungsstufen eine Rolle spielen. Auch das Lehrpersonal muss entsprechend aus- und weitergebildet sein, und die Ausstattung der Bildungsinstitutionen auf allen Stufen mit geeigneter Software und Hardware muss gegeben sein. Wie wichtig diese Aspekte sind, belegt die bereits bestehende Fachkräftelücke in den IT-Berufen. Im Durchschnitt des Jahres 2022 gab es für knapp 41 800 Stellen in diesen Berufen bundesweit keine passend qualifizierten Arbeitslosen (Flake et al. 2023, S. 2). Für Unternehmen mit Bedarf an Mitarbeitenden für die Entwicklung und Implementierung von KI ist diese Situation jetzt schon dramatisch. Die absehbar weiterwachsende Nachfrage nach entsprechendem Personal dürfte die Problematik in den nächsten Jahren noch erheblich verschärfen.

WAS GENERATIVE KI FÜR UNTERNEHMEN BEDEUTET

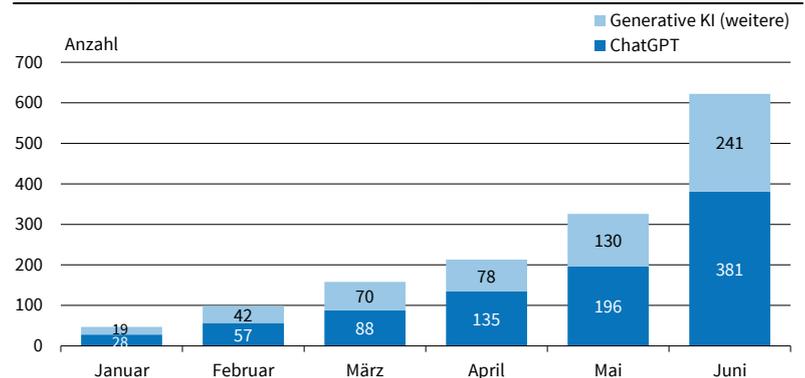
Der Begriff »generative KI« umschreibt KI-basierte Systeme, die in der Lage sind, gänzlich neue Inhalte zu erstellen, die in ihrem Ergebnis menschenähnlich sind. Ein Vorteil der aktuell viel diskutierten Modelle ist, dass die Anwendung für Beschäftigte am Computer einfach ist und ihr Nutzen schnell ersichtlich ist. Auch Unternehmen, die nicht wie oben erläutert, »data economy ready« sind, können diese Modelle nutzen. Empfehlenswert ist angesichts der rasanten Entwicklung und der umfassenden Möglichkeiten der Technologie sicherlich die Erarbeitung unternehmensinterner Handlungsleitlinien, aber darüber hinaus bestehen keine wesentlichen Voraussetzungen für den Einsatz dieser generativen KI.

Betrachtet man die Anzahl an Stellenanzeigen für den deutschen Arbeitsmarkt, die konkret Kenntnisse rund um generative KI und ChatGPT nachfragen, so lässt sich seit Beginn dieses Jahres ein deutlicher Anstieg in der Nachfrage erkennen (vgl. Abb. 1). Diese wachsende Popularität der KI-Technologien lässt sich vor allem über den noch immer anhaltenden Hype um ChatGPT erklären. Doch auch andere, neue generative KI-Modelle werden immer häufiger Teil von geforderten Kompetenzprofilen.

Die Software-Beratung Capterra hat im Sommer 2023 eine Befragung unter rund 500 Beschäftigten durchgeführt, die generative KI nutzen (Bahr 2023). Davon nutzten Anfang Juni 2023 (vor der Einführung von Bard in Deutschland) 68% ChatGPT. Vor allem wird ChatGPT zur Textbearbeitung, zur Analyse von Daten und zum Schreiben von Texten eingesetzt. Jeweils mehr als 40% der ChatGPT-Nutzerinnen und Nutzer gaben diesen Einsatzbereich des KI-Tools an. 60% prüfen alle Outputs von ChatGPT vor der Verwendung sorgfältig, weitere 35% führen Überprüfungen für viele Outputs durch. Dies zeigt, dass die Nutzung generativer KI auch Grenzen hat und in den meisten Fällen eine Unterstützung des Menschen darstellt, aber keinen Ersatz. Dies bestätigt der Befund der ge-

Abb. 1

Stellenanzeigen in Deutschland, die Kenntnisse in generativer KI nachfragen, 2023



Quelle: Textkernel; Institut der deutschen Wirtschaft.

© ifo Institut

nannten Studie zu den Nachteilen von ChatGPT: Am häufigsten wurde die Befürchtung genannt, dass Menschen den Output als immer richtig erachten und so Falschinformationen verwenden könnten (Bahr 2023). Positiv bewerteten die Befragten jedoch die Verbesserung der Arbeitsabläufe durch ChatGPT.

Es wird auch an diesem Beispiel deutlich: KI bietet große Chancen, ist aber auch in der Entwicklung, Implementierung und Anwendung eine Herausforderung. Dieser müssen Unternehmen mit Information, Schulungen und einem angemessenen Risikobewusstsein begegnen, um die Potenziale von KI heben zu können. Diese sind noch lange nicht ausgeschöpft: Die Anwendungsmöglichkeiten generativer KI gehen weit über Text oder auch Bild hinaus, und auch andere KI-Technologien können für Unternehmen sehr hilfreich sein.

WELCHE REGULIERUNG KI BENÖTIGT

Das beschriebene Zusammenspiel aus Chancen und Risiken von KI verdeutlicht, dass ein geeigneter Rahmen für den Umgang mit KI-Technologien unerlässlich ist. Auf EU-Ebene soll der AI Act, dessen Details aktuell in den Trilog-Verhandlungen von Kommission, Parlament und Rat ausgehandelt werden, diesen Rahmen schaffen. Die konkreten Inhalte des Gesetzes sind sehr umstritten, insbesondere die Klassifikation von KI nach Risikostufen bzw. die Einstufung bestimmter KI-Technologien in diese Stufen. Technologien der risikoreichsten Stufe werden in der EU verboten werden. Hier ist klar, dass mit der Einstufungsentscheidung auch die Zukunft von KI in der EU entscheidend mitbestimmt wird. Um Innovationen nicht abzuwürgen, sondern mit dem Gesetz Rechtssicherheit zu schaffen und KI zu fördern, dürfen die Vorgaben nicht zu weitreichend sein, sondern sollten stattdessen stetig evaluiert und gegebenenfalls verschärft werden, sollte die Regulierung nicht ausreichen.

Das Themenfeld der generativen KI im AI Act wird ebenfalls noch diskutiert – es hat gar den Anschein, als sei die Kommission auf den letzten Metern der Regulierung von der Technologie überrascht worden.

Forscher der Stanford University haben geprüft, inwieweit die derzeit verfügbaren LLM die Anforderungen des AI Act in seiner aktuellen Version erfüllen (Bommasani et al. 2023). Das ChatGPT zugrunde liegende LLM kann von den 48 identifizierten Anforderungen lediglich 25 erfüllen, also 52%. Das LLM von Google, das hinter Bard steckt, liegt bei 27 erfüllten Anforderungen (56%). Keines der Modelle erfüllt alle Anforderungen. Dies kann zweierlei Implikationen haben: Zum einen kann der Befund ein Hinweis auf die Notwendigkeit des AI Act sein, weil vorhandene LLM offenbar wesentliche Qualitätskriterien wie etwa Urheberrecht oder Transparenz nicht freiwillig erfüllen. Allerdings sagt diese Untersuchung nichts darüber aus, ob es mit den derzeitigen LLM und ihren Geschäftsmodellen überhaupt möglich wäre, die Anforderungen zu erfüllen. Somit könnte der AI Act zum anderen auch dazu führen, dass Innovationen über LLM in Europa erschwert oder gar verhindert werden und diese Art KI-Technologie zukünftig vielleicht hierzulande eine geringere Rolle spielt als in anderen Weltregionen. Eine solche Verdrängung von Technologien in Europa durch Regulierung gilt es, tunlichst zu vermeiden.

Gut umgesetzt kann der AI Act aber auch die Rechtssicherheit von Unternehmen im Zusammenhang mit KI erhöhen und so die Nutzung der Technologie vorantreiben. Unabhängig von der finalen Ausgestaltung des AI Act ist heute schon klar, dass sich das Gesetz durch eine hohe Komplexität auszeichnen wird. Es ist zudem zu befürchten, dass die Regulierung Begrifflichkeiten und Bestimmungen enthält, die ausgelegt werden müssen, weil eine klare Definition fehlt. Hier liegt immenser Beratungsbedarf von Un-

ternehmen vor, wenn das Ziel sein soll, dass der AI Act ein Enabler für Künstliche Intelligenz in Europa werden soll.

REFERENZEN

- Acemoglu, D., D. Autor, J. Hazell und P. Restrepo (2022), »Artificial Intelligence and Jobs: Evidence from Online Vacancies« *Journal of Labour Economics* 40(S1), S293–S340.
- Acemoglu, D. und P. Restrepo (2019), »Artificial Intelligence, Automation, and Work«, in: A. Agrawal, J. Gans und A. Goldfarb (Hrsg.), *The Economics of Artificial Intelligence: An Agenda*, The University of Chicago Press, Chicago und London, 197–236.
- Bahr, I. (2023), »Studie zur Nutzung von ChatGPT in Unternehmen«, verfügbar unter: Studie zur Nutzung von ChatGPT in Unternehmen ([capterra.com.de](https://capterra.com/de), aufgerufen am 21. Juli 2023).
- Büchel, J. und B. Engels (2022), »Viele Unternehmen sind nicht bereit für die Datenwirtschaft«, *IW-Kurzbericht* Nr. 96, Institut der deutschen Wirtschaft, Köln.
- Büchel, J., J. Engler und A. Mertens (2023a), »The Demand for Data Skills in German Companies: Evidence from Online Job Advertisements«, *EconPol Forum* 24(2), 56–61.
- Büchel, J. und A. Mertens (2022), *KI-Bedarfe in Deutschland. Regionale Analyse und Entwicklung der Anforderungsprofile in KI-Stellenanzeigen*, Gutachten im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz, Berlin.
- Engels, B. (2023), »KI in der deutschen Wirtschaft: Ohne Digitalisierung und Daten geht nichts«, *Wirtschaftsdienst* 104(8), im Erscheinen.
- Flake, R., J. Tiedemann und A. Jansen (2023), *Fachkräftemangel in IT-Berufen – gute Chancen für Auf- und Quereinsteigende*, Studie im Rahmen des Projektes Kompetenzzentrum Fachkräftesicherung (KOFA) in Zusammenarbeit mit dem Bundesministerium für Wirtschaft und Klimaschutz (BMWK), Köln.
- Rammer, C. (2022), *Kompetenzen und Kooperationen zu Künstlicher Intelligenz. Ergebnisse einer Befragung von KI-aktiven Unternehmen in Deutschland*, Gutachten im Auftrag des Bundesministeriums für Wirtschaft und Klimaschutz, Berlin.
- Rusche, C., V. Demary, H. Goecke, E. Kohlisch, A. Mertens, M. Scheufen und J. M. Wendt (2022), *KI-Monitor 2022. Künstliche Intelligenz in Deutschland*, Gutachten im Auftrag des Bundesverbandes Digitale Wirtschaft (BVDW) e.V., Berlin.

Marie-Christine Fregin und Michael Stops

Künstliche Intelligenz: Wer mit KI-Technologien erfolgreich sein will, sollte die Wirkungen valide abschätzen können

Die jüngsten Innovationen im Bereich der Künstlichen Intelligenz (KI) befeuern gerade wieder die schon länger geführte Automatisierungsdiskussion, bei der es um die Frage geht, ob und inwiefern Maschinen menschliche Arbeit ersetzen. Die nun teils kostenlos verfügbaren KI-Systeme scheinen sich dabei in mehrfacher Hinsicht von früheren technologischen Innovationen zu unterscheiden:

Erstens scheint derzeit die Geschwindigkeit, mit der die immer mächtigeren (Weiter-)Entwicklungen im KI-Bereich verfügbar gemacht werden, deutlich schneller zu sein.

Zweitens lassen die neuen KI-Systeme erwarten, dass das Spektrum der Aufgaben, die automatisiert

oder unterstützt werden können, erheblich erweitert wird. Während Software- und Robotik-Systeme vor allem spezifische manuelle Tätigkeiten und Routinetätigkeiten automatisieren oder unterstützen, können KI-Systeme potenziell auch kognitive Nicht-routinetätigkeiten übernehmen und dabei menschliche Fähigkeiten weit übertreffen. Im Fall von Large-Language-Modellen (also großen Sprachmodellen bzw. Neuronalen Netzen, die auf Basis sehr großer Datenmengen menschliche Sprache verstehen), die hinter ChatGPT und anderen KI-Anwendungen stehen, geschieht das z. B. in Bereichen, in denen die Arbeit auf der Sammlung und (Neu-)Zusammenstellung einer Vielzahl von Informationen beruht, bspw. von Bildern, Texten oder Programmcode.

Drittens versprechen die KI-Anwendungen in der Summe einen breiten Anwendungsbereich – und damit einen Wandel der Arbeitswelt, der weniger inkrementell ausfallen könnte, als das bei früheren Technologien der Fall war. Dies hat zur Folge, dass sich Berufe in nahezu allen Sektoren und über alle Qualifikationsniveaus hinweg verändern können.

Grenzen der Einsetzbarkeit lassen sich (heute) insbesondere dort ziehen, wo es im Kern um menschliche Interaktionen geht oder wo Intuition, ethisch-moralische Abwägungen, menschliche Sozialisierung oder auch persönliche Präferenzen eine Rolle spielen. Beispiele hierfür lassen sich in nahezu allen Branchen finden, aber sicher spielt das für einige Branchen eine größere Rolle, wie in den Bereichen Gesundheit und Soziales oder im Hotel- und Gaststättengewerbe. Was im Allgemeinen als »kreative Tätigkeit« beschrieben wird, kann von neueren KI-Systemen jedoch oftmals schon übernommen werden, beispielsweise die Erstellung digitaler Gemälde oder Gedichte. Dennoch gilt, dass häufig unklar ist, ob sich das Nutzenpotenzial von KI-Systemen in der praktischen Anwendung überhaupt voll entfalten wird. Bei vielen Systemen, wie bspw. ChatGPT oder anderen Large-Language-Modellen, die ihre Aufgaben durch adaptive Entscheidungsfindung und Mustererkennung erledigen, ist das Nutzenpotenzial noch nicht einmal abgesteckt: Wir wissen nicht, wo die Grenzen liegen. Was KI-Systeme in Zukunft können werden, lässt sich derzeit nicht mit Bestimmtheit vorhersagen.

Aus derzeitigen Studien zur Einführung von KI-Systemen und ihren tatsächlichen Effekten lässt sich eine Entwicklung, in der Gesellschaften mit der massenhaften Freisetzung von Beschäftigten durch Automatisierung rechnen müssen, eher nicht ableiten. Für vorausschauende Gesellschaften und Ökonomien liegt es jedoch trotzdem nahe, eine Situation in der Zukunft zu reflektieren, in der in einigen Bereichen möglicherweise deutlich weniger oder auch gar keine Menschen mehr beschäftigt sein werden; siehe beispielsweise den Insight Report »Positive AI Economic Futures« des Weltwirtschaftsforums aus dem Jahr 2021.

KI-KOMPETENZEN WERDEN NICHT IN ALLEN BRANCHEN GLEICHERMASSEN NACHGEFRAGT – DABEI VERÄNDERTEN SICH BISLANG EHER TÄTIGKEITSINHALTE VON BERUFEN

Die Nachfrage der Betriebe nach Arbeitskräften mit Kompetenzen, die für die Entwicklung, Einführung oder Nutzung von KI-Systemen notwendig sind, lässt darauf schließen, dass sich Deutschland eher noch in einem frühen Stadium der Nutzung von KI-Systemen befindet. Bis dato sind vor allem Betriebe der Branchen Informations- und Kommunikationstechnologien und der Unternehmensdienstleistungen besonders aktiv –

also Betriebe in den beiden Branchen, die vorwiegend KI-Systeme entwickeln und/oder (bei anderen Unternehmen) einführen. Mit zunehmender Nachfrage nach KI-Systemen sollten diese Branchen eher noch Beschäftigung aufbauen als abbauen. Hier sind daher auch positive Effekte auf die Löhne zu erwarten. In anderen Branchen ist der Bedarf an Arbeitskräften mit KI-bezogenen Kompetenzen (derzeit) noch relativ gering; gegenwärtig werden solche Kompetenzen am häufigsten von Betrieben des produzierenden Gewerbes, von den Finanzdienstleistungen oder auch im Einzelhandel nachgefragt. Bislang ist jedoch kein Zusammenhang mit der Entwicklung der Beschäftigung in diesen Betrieben erkennbar – nicht nur in Deutschland veränderte KI vor allem Tätigkeitsinhalte innerhalb von Berufen. Gleichwohl schüren einige Unternehmen Befürchtungen von Entlassungen, indem sie Personalabbau ankündigen und dies mit der Einführung von KI begründen.

Dass KI bisher insgesamt wenig quantitative Beschäftigungseffekte verursacht, mag mehrere Gründe haben. KI-Systeme sind in der deutschen Wirtschaft noch recht wenig verbreitet und die Systeme, die Marktreife erlangt haben und in Unternehmen im Einsatz sind, verfügen oftmals über eher spezifische Kompetenzen. Einige KI-Anwendungsbeispiele, die im Rahmen des Forschungsprojektes »ai:conomics« vom Research Centre for Education and the Labour Market der Universität Maastricht (ROA), dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) in Nürnberg sowie der Innovationsagentur zukunft zwei GmbH näher untersucht werden, zeigen, dass derzeit eingeführte KI-Systeme in ihrer Reichweite, ihrem Anwendungsbereich und in ihrem Zweck eher begrenzt sind; das heißt, dass es lediglich spezifische Tätigkeiten sind, die mit KI unterstützt oder durch KI ersetzt werden. Dadurch ist auch eher eine geringe Anzahl von Beschäftigten betroffen – oft nicht zuletzt deshalb, weil die Einführung von KI der nächste Schritt in einem länger dauernden Prozess der Automatisierung ist und KI-Neuerungen vor allem dort zum Einsatz kommen, wo bereits vorher viele Tätigkeiten digitalisiert und/oder automatisiert wurden. Bei der KI-Einführung ergeben sich zudem gegensätzliche Effekte: Einerseits müssen Beschäftigte bei der Einführung neuer Systeme oftmals neue Tätigkeiten ausführen und teilweise erlernen; hier ist daran zu denken, dass



Dr. Marie-Christine Fregin

ist Forschungsleiterin am Research Centre for Education and the Labour Market an der Universität Maastricht.



Dr. Michael Stops

ist wissenschaftlicher Mitarbeiter am Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

KI-Systeme in ihrer Funktionsweise und Zielsetzung verstanden, trainiert und gewartet werden müssen. Andererseits ist erwartbar, dass bestimmte Tätigkeiten, die bisher den Beschäftigten vorbehalten waren, von der KI unterstützt und manchmal sogar übernommen werden können. Und schließlich bleibt den Beschäftigten (noch lange) die Aufgabe, die Ergebnisse und Vorschläge, die ein KI-System erzeugt, zu verstehen, zu bewerten und die richtigen, zweckgerichteten Schlüsse zu ziehen.

IN EINER WELT, IN DER KI IMMER WICHTIGER WIRD, IST ES RATSAM, SICH GANZHEITLICH MIT DEN WIRKUNGEN AUSEINANDERZUSETZEN

Aus den bisherigen Befunden lässt sich ableiten: Dort, wo KI-Systeme einen höheren Innovationsgrad und eine bessere Qualität verbunden mit einer höheren Produktivität gewähren, sind Betriebe gut beraten, in Digitalisierung und die Nutzung Künstlicher Intelligenz zu investieren. In einer Welt, in der innovative Technologien immer mehr an Bedeutung gewinnen, wird dies nicht zuletzt für die Wirtschaftlichkeit und internationale Wettbewerbsfähigkeit von Unternehmen relevant sein.

Dabei ist naheliegend, dass KI-Systeme in den Bereichen, in denen Fachkräfte besonders gesucht werden, den Bedarfsdruck reduzieren könnten. Dies ist jedoch kein Selbstläufer. KI-Technologien sind bei weitem (noch) nicht so fortgeschritten, dass es nur wichtig wäre, zu wissen, wo die Fachkräftengpässe sind, um sie mit Maschinen zu überwinden. Auch die Tatsache, dass nur ein bestimmter Teil der üblicherweise in einem Beruf auszuführenden Tätigkeiten übernommen werden kann, dämpft die Hoffnung auf eine Entlastung von Fachkräftengpässen durch KI. Dennoch ist es für die Betriebe sinnvoll zu prüfen, in welchen Bereichen eine maschinelle Unterstützung möglich ist – nicht zuletzt in Bereichen, in denen Menschen gefährliche, gesundheitsschädliche oder auch sehr ermüdende Tätigkeiten ausführen müssen.

Gleichzeitig wird deutlich, dass die Konsequenzen des Einsatzes von KI-Systemen für Beschäftigte, Arbeit und Arbeitsmärkte bislang nicht hinreichend verstanden werden. Darauf weist nicht nur die Enquete-Kommission »Künstliche Intelligenz« in ihrem Abschlussbericht an die Deutsche Bundesregierung aus dem Jahr 2020 hin. Wenn die Stakeholder des technologischen und gesellschaftlichen Wandels nicht besser verstehen, welche Konsequenzen der KI-Einsatz für Beschäftigte, Arbeit und Arbeitsmärkte hat, wird es schwierig bleiben, Arbeitswelten von Morgen besser einzuschätzen und die Zukunft der Arbeit proaktiv zu gestalten. Nicht zuletzt für eine Regulierung innovativer Systeme in Betrieben und Volkswirtschaften, die Chancen der Technologien ergreifen und negativen Konsequenzen vorbeugen lässt, ist ein besseres Verständnis kausaler Effekte des KI-Einsatzes notwendig. Dies gelingt nur auf der Basis evidenzbasierter Evaluation.

DIE STAKEHOLDER DES TECHNOLOGISCHEN UND GESELLSCHAFTLICHEN WANDELS MÜSSEN KI-EFFEKTE VERSTEHEN, UM ARBEITSWELTEN DER ZUKUNFT GUT ZU GESTALTEN

Die Erfahrungen aus dem derzeit laufenden Forschungsprojekt ai:conomics zeigen, dass evidenzbasierte Evaluationen im Vorfeld der Technologieeinführung eher selten in einem erforderlichen Umfang bedacht wurden. Dabei wird in die Entwicklung von KI und in ihre Einführung sehr viel Geld investiert. Eine datenbasierende Analyse wäre von großem Vorteil für die Unternehmen – gerade zum Beginn des technologischen Wandels mit KI –, um die Erfolgswahrscheinlichkeit von Investitionen zuvor besser abschätzen und den Erfolg im Anschluss besser beurteilen zu können. Hierzu gehört ein umfassender Blick auf positive und negative Effekte sowohl für die Belegschaft und die Produktivität. Hürden für eine evidenzbasierte Evaluation liegen z. B. in der Verfügbarkeit geeigneter Daten und ausreichend performanter Technik – aber auch in rechtlichen, organisatorischen und kulturellen Gegebenheiten. Bereits im Vorfeld der KI-Einführung sollten Unternehmen Kenntnis darüber haben, ob die für ein erfolgreiches Training von KI benötigten Daten zur Verfügung stehen. Und/oder ob diese Daten gegebenenfalls an KI-Dienstleister übermittelt werden dürfen und können, wenn KI-Technologie eingekauft wird. Unternehmen sollten auch wissen, ob ihre Belegschaft für die Einführung von KI-Systemen bereit ist und über die notwendigen Kompetenzen verfügt oder wie diese aufgebaut werden können.

Befragungen im Rahmen unserer Forschung legen nahe, dass Beschäftigte und Führungskräfte die Einführung von neuen Technologien eher akzeptieren, wenn sie gut über Ziele und Wirkungsweisen der Technologien informiert sind. Unternehmen sollten sich über die Technologieakzeptanz ihrer Belegschaften bewusst sein und Qualifizierungsmaßnahmen zum Zweck und zur Funktion der neuen Technologien sowie zum praktischen Umgang mitdenken. In unserer Forschung bei und mit Unternehmen haben wir beobachtet, dass veränderte Kompetenzbedarfe und Lernkultur im Betrieb manchmal nicht unbedingt zusammenpassen. Bislang handelt es sich hier um anekdotische Evidenz. Die Muster kehren bei der Arbeit mit verschiedenen Konzernen jedoch wieder. Nach mehrjähriger Forschung zu Technologieinnovationen und ihrer Einführung in Unternehmen verfestigt sich der Eindruck, dass es vielen Unternehmen ähnlich geht.

»BUSINESS CASE« UND »PEOPLE CASE« GEHÖREN ZUSAMMEN: EINE UMFASSENDE EVALUATION DER KI-EFFEKTE HILFT ALLEN

Unternehmen sollten jedoch wissen, wie der Erfolg einer Technologieeinführung valide messbar ist, das ist nicht zuletzt relevant für kommende KI-Projekte.

Dabei geht es nicht nur darum, ob das Unternehmen dank der KI-Einführung wirtschaftlich erfolgreicher agiert. Um besser zu verstehen, welchen Anteil eine Technologieeinführung am Erfolg hat, sollte dem Unternehmen klar sein, mit welchen Zielen es das KI-System einführen will: Soll die Produktivität an einer bestimmten Stelle im Unternehmen erhöht werden? Geht es um Entlastung von bestimmten Tätigkeiten zugunsten anderer Tätigkeiten, die weniger durch Technik unterstützbar sind? Soll die Diversität und Qualität eines bestimmten Produktes erhöht werden? Das Set an Erfolgsindikatoren ist vielfältig. Es bezieht sich nicht nur auf den »Business Case«, sondern lässt sich auch entlang der Unternehmenswerte fortsetzen; z. B. ist hier auch an Zufriedenheitsindikatoren für Belegschaften, Kund*innen und andere Stakeholder zu denken (der »People Case«). Bislang gibt es kaum entsprechende Studien auf der Betriebsebene, obwohl das Potenzial, entsprechende Daten zu erschließen, aufgrund der zunehmenden Digitalisierung noch nie so groß war wie gegenwärtig. Befragungen der Belegschaften könnten vorhandene Informationslücken weiter schließen.

In der Forschung bei und mit multinationalen Konzernen im Projekt ai:conomics und in dessen Vorläuferprojekten zeigte sich, dass bei der wissenschaftlichen Begleitung von Technologieeinführungen in Betrieben einige Hürden auftreten – obwohl sich alle Beteiligten grundsätzlich einig darüber waren, dass die Einführung von KI-Systemen wissenschaftlich begleitet werden sollten. So gibt es z. B. keine Standards für die Erhebung von oder den Zugang zu einschlägigen Daten wie Key Performance Indicators. Dies erschwert die Erhebung, Nutzung und Zusammenführung solcher Daten aus verschiedenen Quellen erheblich. Unter anderem mag dies daran liegen, dass bei der Entwicklung der entsprechenden Produktivsysteme die Möglichkeit einer evidenzbasierten Evaluation nicht vorgesehen wird. Bei der Messung »weicherer« Faktoren (wie z. B. Lernkultur im Unternehmen, Wohlbefinden der Beschäftigten) oder der Identifikation konkreter Skills und Kompetenzen, die aufgebaut bzw. nachgefragt werden müssen, fehlt Praktiker*innen in den Unternehmen teilweise das notwendige Wissen, um entsprechende Erhebungen zu designen, durchzuführen und die Daten auszuwerten.

DIE ERFORSCHUNG VON KI-EFFEKTEN IST NOTWENDIG, ABER ANSPRUCHSVOLL – UND GELINGT AM BESTEN, WENN ALLE STAKEHOLDER DES WANDELS ZUSAMMENARBEITEN

Die Messung der Konsequenzen von Technologieeinführungen, wie die von KI-Systemen, erfordert anspruchsvolle Forschungsdesigns. Dabei stellen Feldexperimente den »goldenen Standard« dar. Sie können genutzt werden, um kausale Evidenz zu generieren und auch um Brücken zwischen Wissenschaft und Unternehmen zu bauen. Sie können z. B. wertvolle Erkenntnisse für Unternehmen liefern, die vor der Herausforderung einer strategischen Personalplanung stehen, sowie für Sozialpartner, die den regulatorischen Rahmen für den Einsatz von KI im Betrieb schaffen müssen. Aus dieser Perspektive können Feldexperimente den F&E-Prozess zu KI in Unternehmen um eine sozioökonomische Dimension ergänzen. Sie sind aber voraussetzungsvoll: Bei einem Feldexperiment in einem Unternehmen nimmt eine zufällig ausgewählte Gruppe von Beschäftigten die Arbeit mit KI-Unterstützung auf; dies ist die sogenannte Experimentgruppe (»Treatment-Gruppe«). Für die Dauer des Experiments arbeiten die anderen Beschäftigten ohne KI-Unterstützung weiter; das ist die Kontrollgruppe. Eine solche »Versuchsordnung« ermöglicht die Beobachtung kausaler Auswirkungen von KI auf Beschäftigte und Arbeitsplätze. Dafür bedarf es einer detaillierten Planung und genauer Absprachen über den Ablauf der Einführung der KI-Systeme; andernfalls lassen sich Experiment- und Kontrollgruppe nicht vergleichen, und eine valide Wirkungsanalyse ist nicht möglich.

Für die Durchführung eines Feldexperiments bedarf es nicht nur der grundsätzlichen Bereitschaft der Betriebe, an der Erforschung der KI-Effekte für Arbeitnehmer*innen und Arbeit mitzuwirken. Die Organisationen müssen auch bereit sein, temporär und partiell von bisherigen Entscheidungsparadigmen zugunsten eines gültigen Forschungsdesigns abzuweichen. Das schließt ein, dass Entscheidungsträger*innen auf mehreren Ebenen sowie Personalvertretungen ebenfalls bereit sein müssen, die Umsetzung des Forschungsdesigns mitzutragen. Letzteres ist auch eine wichtige Voraussetzung für die Realisierung valider Befragungsdesigns. Diese können Mehrwerte für alle Stakeholder des technologischen und gesellschaftlichen Wandels stiften.

Andreas Gillhuber

Die wichtigsten Trends im KI-Einsatz – und worauf es jetzt in Europa ankommt



Andreas Gillhuber

verantwortet als Geschäftsführer und Co-CEO der Alexander Thamm GmbH das operative Geschäft in allen Kundenprojekten und ist Vorstand in der German Data Science Society e.V.

Mit ChatGPT ging der Hype los: Künstliche Intelligenz (KI) rückte Ende letzten Jahres vom Expertenthema in den Fokus von Politik, Unternehmen und Endverbrauchern. KI ist Trendsetter, und entsprechend manifestierten sich 2023 auch für den Einsatz von KI-Anwendungen zahlreiche Trends. Einige der wichtigsten stellen wir vor – zusammen mit den einhergehenden politischen und gesellschaftlichen Fragestellungen.

TREND 1: GENERATIVE KI UND EXPLAINABLE AI – MEGATREND UND FORSCHUNGSGBIET

Der Einsatz von sehr großen Sprachmodellen (Large Language Models, LLMs) ist mit ChatGPT im öffentlichen Bewusstsein angekommen. ChatGPT steht dabei stellvertretend für die Potenziale von Sprachmodellen, die in Kombination mit Reinforcement Learning menschliche Fragen – ungeachtet noch vorhandener Schwächen – gut beantworten können. So können LLMs für Unternehmen im Bereich des Kundenservice neue Möglichkeiten einer schnelleren und automatisierten Kundenkommunikation eröffnen, die Customer Experience verbessern und zugleich Kosten senken. Die große Stärke von LLMs sind die Interaktion in natürlicher Sprache bei Berücksichtigung des jeweiligen Kontextes, wie die Bezugnahme zu vorherigen Fragen und Antworten.

LLMs sind ein Beispiel von »Generativen KI« (Generativen AI, GenAI)-Modellen. In den letzten Jahren erzielten Unternehmen und Forschungseinrichtungen in den USA und China bahnbrechende Resultate mit großen KI-Modellen, aber auch deutsche Wissenschaftler und Forscher in Start-ups sind ganz vorne mit dabei. Dazu gehören Stable Diffusion von Prof. Björn Ommer (LMU)¹, die fotorealistischen 3D-Avatare von Prof. Matthias Nießner (TUM)² und auch viele weitere Forschungsgebiete der rund 50 KI-Lehrstühle, die sich im Munich Center of Machine Learning (MCML)³ zur interdisziplinären Zusammenarbeit organisiert haben. Aus etlichen dieser For-

¹ Prof. Dr. Björn Ommer, Head of Computer Vision & Learning Group, Ludwig Maximilian Universität München (LMU), <https://ommer-lab.com/people/ommer/>.

² Prof. Dr. Matthias Nießner, Technische Universität München (TUM), Department of Informatics: <https://niessnerlab.org/index.html>.

³ Munich Center of Machine Learning (MCML) ist eines von sechs nationalen KI-Kompetenzzentren: <https://mcml.ai>.

schungsgebiete gehen Start-ups hervor, die den Vergleich mit ihren US-Pendants nicht scheuen brauchen.

Trainiert mit riesigen Datenmengen auf den weltweit größten Superrechnern sind diese GenAI-Modelle zu spektakulären Anwendungen fähig – von automatisierter Texterstellung über die Kreation fotorealistischer Bilder bis zur Entwicklung von Programm-Code aus natürlicher Sprache. GenAI kombiniert hierfür KI und maschinelles Lernen, um weitgehend automatisiert neuen digitalen Video-, Audio- oder Text-Content zu entwickeln bzw. zu generieren – daher der Name. Die möglichen Anwendungen sind vielfältig: So können kleine und mittlere Unternehmen, die keine eigene Werbeabteilung haben, GenAI im Marketing einsetzen. GenAI hat somit das Potenzial, mit unterschiedlichsten Anwendungen die Berufswelt nachhaltig zu verändern.

Noch offen ist die Frage, in welchem Ausmaß Unternehmen Tools wie DALL-E 2 oder Mindverse in ihre Arbeitsprozesse integrieren werden. Die Erklärbarkeit der Modelle wird hierbei eine entscheidende Rolle spielen. Auch der europäische AI-Act (s.u.) verlangt detaillierte Dokumentation und Transparenz, um z. B. Bias zu vermeiden und sogenannte »Black-Box-Modelle« nachvollziehbar zu machen. Hierfür gibt es bereits einige Methoden im Forschungsbereich »Explainable AI« (XAI), um besser zu verstehen, welche Daten und Variablen in den komplexen Algorithmen primär genutzt werden. LLMs können helfen, automatisiert zu dokumentieren und standardisiert Code zu schreiben, was die Lesbarkeit und Nachvollziehbarkeit vereinfacht. Die Fortentwicklung von XAI geht damit Hand in Hand mit der Entwicklung von GenAI.

TREND 2: DATA CENTRIC AI – DIE DATEN MACHEN DEN UNTERSCHIED!

Die Entwicklungen der letzten Monate haben zudem deutlich gemacht: Die Demokratisierung der KI schreitet voran. Kostspielige und aufwändig trainierte KI-Modelle werden durch Open Source leichter zugänglich. Diese stehen damit den Unternehmen als Grundlage zur Verfügung, die sie nun feintunen, mit spezifischen Abfragen/Kontexten (sog. Prompts) füttern, durch eigene Modelle (je nach geschäftlicher Fragestellung, zu lösender Problematik etc.) ergänzen sowie individualisieren und sich somit am Markt differenzieren können.

Ein wichtiger weiterer Differentiator sind die Daten des eigenen Unternehmens und/oder seiner spezifischen Partner und Kunden. Deshalb ist es schon jetzt notwendig, in die Qualität der eigenen Daten und den Aufbau einer umfassenden Data Governance zu

investieren. Wie ein solches Investment aussieht, zeigt das Beispiel eines bei einem großen Verkehrs- und Logistikkonzern genutzten Datenkatalogs, bei dessen Entwicklung und Implementierung die Alexander Thamm GmbH unterstützt hat.

Klarheit über die eigenen Daten mit einem Datenkatalog für den Data Lake

Millionen Reisende oder Güterverkehre, tausende Transportmittel sowie ungezählte Daten über Infrastruktur und Fahrzeuge – jedes große Transport- und Logistikunternehmen verfügt über einen enormen Datenschatz, der allerdings erst effektiv nutzbar gemacht werden muss. Einige dieser Unternehmen entwickeln daher seit einigen Jahren einen Datenkatalog für ein modernes Management der Daten in den angeschlossenen Datenbanken, Data Warehouses und Data Lakes. Diese Datenkatalog-Software fungiert als neues Werkzeug und erster konkreter Schritt hin zu einer sicheren, effizienten und konzernübergreifenden Bereitstellung von Daten.

Das operative Tool gibt einen Überblick über den Datenbestand, ermöglicht mit einem Regelungssystem den konzernweiten Zugriff auf diese Daten und legt so die Basis für eine effektive Zusammenarbeit. Damit ein solcher zentraler Datenkatalog funktioniert, muss er in eine effektive Data Governance eingebettet sein, einem zeitgemäßen Data Access Management, einem intelligenten Verrechnungsmodell für alle nutzenden Geschäftsbereiche und Gesellschaften, die über ein Datenökosystem und die passende Infrastruktur miteinander datentechnisch verbunden sind.

TREND 3: KI & REGULATORIK – INNOVATION BRAUCHT EINEN RAHMEN, ABER KEIN KORSETT!

Spätestens mit dem Start von ChatGPT ist neben den Möglichkeiten von KI auch die Frage nach einer angemessenen Regulierung von KI-Anwendungen im öffentlichen Bewusstsein angekommen. In der Tat stellen gesetzliche Regulierungen wie der AI Act der Europäischen Union oder die ESG-Regulierung im Finanzsektor die Unternehmen vor große Herausforderungen. Mittelständler sehen sich kaum in der Lage, die aufkommenden zusätzlichen Compliance-Regeln umzusetzen.

Doch unabhängig von diesen übergreifenden Überlegungen steht im Hinblick auf eine mögliche Regulatorik fest: Wenn wir in Deutschland und in der EU nicht den Anschluss verlieren wollen, müssen Unternehmen, Politik und Gesellschaft hier gemeinsam einen KI-Standort schaffen, der Innovationsgeist fördert und Risikobereitschaft honoriert. Wir werden im globalen Wettbewerb nur bestehen können, wenn sich die innovativsten Köpfe dafür entscheiden, in der Europäischen Union zu forschen, zu lehren und zu gründen. Zusammen mit den Mitgliedern des Bundesverbands Künstliche Intelligenz macht sich die Alexander

Thamm GmbH deshalb dafür stark, die Potenziale der Künstlichen Intelligenz im Sinne europäischer Werte zu entwickeln und die digitale Souveränität der EU zu stärken und gemeinsam mit der German Data Science Society e.V.⁴ die Interaktion von Wissenschaft, Forschung und Wirtschaft zu fördern.

In der Abstimmung über den geplanten AI Act im Europäischen Parlament in diesem Juni sehen wir noch nicht das erhoffte Zeichen zum Aufbruch. Denn der aktuell verabschiedete Vorschlag bremst eher Innovationsfreude und Technologieoffenheit und stellt die Risiken in den Vordergrund. Dementsprechend sehen wir beim AI Act vor allem in drei Bereichen Verbesserungsbedarf:

Die erste Problematik ist die Definition von Künstlicher Intelligenz, die dem Gesetzentwurf zugrunde liegt. Diese entspricht zwar der OECD-Definition, beschränkt sich aber nicht auf die wesentlichen Merkmale einer KI und kann somit theoretisch auch sämtliche elaboriertere Software-Systeme, wie etwa regelbasierte Tabellenkalkulationsprogramme, regulieren. Eine solche pauschale Regulierung würde KI-Innovationen innerhalb der EU im Keim ersticken und traditionelle Expertensysteme, die seit Jahrzehnten z. B. in der Versicherungswirtschaft entwickelt und angewendet werden, ebenso in den Fokus rücken.

Ebenso kritisch ist die geplante Regulierung von Foundation Models, der oben beschriebenen hoch innovativen Basistechnologie für Generative KI, die zahlreiche neue Anwendungsfelder und wirtschaftliche Vorteile ermöglicht. Mit den aktuell vorgesehenen zusätzlichen Anforderungen für Foundation Models werden bestimmte KI-Modelle unverhältnismäßig unter den Generalverdacht einer Hochrisikoanwendung gestellt und damit in ihrer Entwicklung dauerhaft gefährdet. Deutlich hilfreicher wäre dagegen eine verpflichtende Zertifizierung für Foundation Models, die in High-Risk-Systemen eingesetzt werden. Bei Anwendungen in unkritischen Bereichen wäre dann eine Zertifizierung nicht erforderlich.

In einem offenen Brief von über hundert führenden Wirtschaftsvertretern werden weitere gute Vorschläge gemacht. Es sollte ein »Transatlantischer Rahmen« sicherstellen, dass Europas Regelwerk synchronisiert wird mit den Vorschlägen wichtiger Akteure der Vereinigten Staaten, die ebenfalls an einer KI-Regulierung arbeiten. Zudem drängen die Autoren auf flexiblere Regeln für Generative KI. Diese sollte nicht in einem starren Gesetz verankert werden. Stattdessen sollte die EU nur allgemeine Grundsätze festlegen und die konkrete Umsetzung dann laufend einem Expertengremium überlassen, um mit der rasanten Technologieentwicklung Schritt zu halten.

Die dritte große Baustelle des Gesetzentwurfs tut sich schließlich bei der Klassifizierung und Defi-

⁴ Die German Data Science Society (GDS) e.V. ist die führende Vereinigung der akademisch ausgebildeten Data Scientists und fördert den Wissens- und Erfahrungsaustausch zwischen Wissenschaft und Forschung mit der Wirtschaft: <https://gds-society.de>.

dition von Hochrisiko-KI auf: Es ist unstrittig, dass bestimmte KI-Anwendungen, z. B. in kritischer Infrastruktur, strengen Auflagen unterliegen müssen. Die im Gesetz geschaffene Kategorie für Hochrisiko-Anwendungen spannt den Bogen jedoch viel zu weit. Die Folge: KI-Anwendungen, die kein oder nur ein geringes gesellschaftliches Risiko darstellen könnten, werden überreguliert – ein Ansatz, der weder für die KI-Entwicklung innerhalb der EU noch für die Perspektiven der gesamten Wirtschaft hilfreich ist und letztlich Know-how und Innovationsgeist aus dem EU-Markt vertreiben wird.

Überregulierung von Innovation: Beispiel 3D-Umgebungssensorik

Was die EU beim Thema KI braucht, sind nicht ideologisch motivierte Schnellschüsse, sondern ein verhältnismäßiger und vernünftiger gesetzlicher Rahmen. Ein Rahmen, der eine geschützte Entwicklung von KI-Innovationen ermöglicht und fördert, von Innovationen wie einem von der Alexander Thamm GmbH entwickelten 3D-Umgebungssensoriksystem für Roboter. Das System wird in der Produktionslogistik und der stationären Pflege eingesetzt. Es soll die Mensch-Maschine-Interaktion effizienter gestalten und zugleich frühzeitig Gefahren erkennen, indem Personen, Gegenstände und Flüssigkeiten in der unmittelbaren Umgebung zuverlässig detektiert und gegebenenfalls sicherheitsrelevante Maßnahmen eingeleitet werden.

Unser System ist von ANNEX II und der Richtlinie 2006/42/EG betroffen und könnte daher erst nach einem umfassenden Risikomanagementprozess implementiert werden. Hierzu zählen u. a. die Bereitstellung zahlreicher Protokolle von Modellierungsaktivitäten und -ergebnissen, eine technische Dokumentation für die Prüfung durch Dritte, die Gewährleistung einer angemessenen Modellgenauigkeit, Robustheit und Cybersicherheit sowie eine Konformitätsbewertung gemäß den sektoralen Rechtsvorschriften. Der AI Act würde also die Kosten und die Komplexität der Entwicklung eines Produkts, das die Arbeitssicherheit und die Arbeitsbelastung hart arbeitender Menschen in Krankenhäusern, Pflegeheimen und Lagerhäusern unterstützt, deutlich erhöhen. Mit der Einführung dieser Compliance-Anforderungen geht die EU demnach bewusst das Risiko ein, die Entwicklung neuer Technologien zu behindern, die eigentlich zur Verbesserung der Sicherheit und Arbeitsbelastung am Arbeitsplatz beitragen.

IN FÜNF SCHRITTEN ZU TRAGFÄHIGEN KI-STRUKTUREN

Schon dieser Blick auf drei wichtige KI-Trends gibt einen Hinweis auf die Dynamik, mit der sich KI-Themen derzeit entwickeln. Dieser Dynamik muss sich allerdings nicht allein der gesetzliche Rahmen innerhalb der EU anpassen. Die Staaten und die Unternehmen müssen zusammen vielmehr auch in die notwendi-

gen Strukturen investieren – und zwar physisch und organisatorisch.

Ein erster Meilenstein ist dabei der Aufbau europäischer Hochleistungsrechenzentren, speziell für die Anwendungsentwicklung und Forschung im Bereich der Künstlichen Intelligenz. Große KI-Modelle werden derzeit hauptsächlich in den USA und China entwickelt. Den Forschern und Entwicklern in Europa fehlt aktuell ausreichender Zugang zu den erforderlichen Rechenkapazitäten. Für eine dauerhaft erfolgreiche und unabhängige KI-Entwicklung fordert die Initiative LEAM (Large European AI Models) des KI Bundesverbands darum von den Staaten und Unternehmen innerhalb der EU die Bereitschaft, jetzt entschieden in die KI-Zukunft und die digitale Souveränität zu investieren. Es ist jetzt Zeit zu handeln: Denn große KI-Modelle verdrängen wegen ihrer Leistungsfähigkeit zunehmend bestehende Lösungen, sind aber nur in den USA über Schnittstellen verfügbar. In Europa stehen wir darum vor der Herausforderung, den Anschluss an diese Entwicklung nicht zu verpassen.

Die Schaffung eines europäischen KI-Hochleistungsrechenzentrums ist allerdings nur ein erster wichtiger Schritt. Sozusagen ein Leuchtturm, um den herum ein vitales Ökosystem aus Entwicklern, Forschern und Unternehmern entstehen muss. Damit dies gelingt, sind noch vier weitere Schritte nötig: der Aufbau von ausreichenden Personalkompetenzen, um die ersten großen KI-Modelle innerhalb der EU zu erschaffen; eine Auswahl von Daten und Algorithmen, die den besonderen Bedürfnissen der europäischen Gesellschaften und Ökonomien gerecht wird; die Entwicklung einer eigenständigen Organisation, die die neuen KI-Modelle Industrie und Forschung zur Verfügung stellt; und die Berücksichtigung europäischer Identitäten und Werte bei der Generierung von KI-Modellen in Verbindung mit einem Höchstmaß an Nachhaltigkeit.

FAZIT

KI wird unsere Wirtschaft und unsere Arbeitsweisen grundlegend verändern. Die aktuellen KI-Trends geben uns eine erste Vorstellung, wie diese Veränderung ablaufen wird. Wie bei jeder industriellen Revolution gilt für die EU und die hier ansässigen Unternehmen auch im Hinblick auf die KI: Es ist besser, Subjekt und nicht Objekt des Wandels zu sein.

Damit die europäischen Ökonomien diesen Wandel erfolgreich gestalten können, brauchen wir darum einen gesetzlichen Rahmen, der Innovation fördert und nicht fürchtet. Und wir brauchen Strukturen, in denen sich das Know-how und die Innovationskraft unserer Unternehmen auch im Wettbewerb mit den USA und aus China erfolgreich entfalten kann.

Die Chancen hierfür sind da. Wir müssen uns als Gesellschaft und Ökonomie nur in die Lage versetzen, diese zu ergreifen, und uns jetzt bereit machen für den Wettbewerb der Zukunft.

REFERENZEN

Abbou, D., L. Alexandre, L., Ardant, R. et al. (Hrsg.) (2023), »Offener Brief an die Vertreter der Europäischen Kommission, des Europäischen Parlaments und des Europäischen Rats. Künstliche Intelligenz: Europas Chance, wieder zur technologischen Avantgarde aufzuschließen«, verfügbar unter: <https://www.handelsblatt.com/downloads/29231390/1/ki-brief.pdf>.

Bundesverband der Unternehmen der Künstlichen Intelligenz in Deutschland e. V. (Hrsg.) (2023a), »The EU AI Act, Towards the Finish Line: Key Issues and Proposals for the Trilogue Negotiations«, verfügbar

unter: https://ki-verband.de/wp-content/uploads/2023/07/Position-Paper_AI-Act-Trilogue_GermanAIAssociation.pdf.

Bundesverband der Unternehmen der Künstlichen Intelligenz in Deutschland e. V. (Hrsg.) (2023b), »Konzeptpapier: Large European AI Models (LEAM) als Leuchtturmprojekt für Europa«, verfügbar unter: <https://ki-verband.de/wp-content/uploads/2022/06/LEAM-Konzeptpapier-V1.2-1.pdf>.

Gillhuber, A., G. Kauermann und W. Hauner (Hrsg.) (2023), *Künstliche Intelligenz und Data Science in Theorie und Praxis. Von Algorithmen und Methoden zur praktischen Umsetzung in Unternehmen*, Springer Spektrum, Berlin.

Johannes Walter

How to Tame a Dragon: Gefahren und Regulierung von KI

Die letzten Monate waren voller Warnungen über die existenziellen Gefahren, die von Künstlicher Intelligenz (KI) ausgehen: Ein offener Brief des Future of Life Institutes fordert eine sechsmonatige Pause in der KI-Entwicklung (Future of Life Institute 2023). Ein Aufruf des Center for AI Safety fordert KI als »globale Priorität«, ähnlich zu Pandemien oder Atomkriegen, zu behandeln (Center for AI Safety 2023). Beide Initiativen fanden zahlreiche und teils sehr prominente Unterzeichner. Viele Menschen sind von den rapiden Fortschritten in KI und der gesellschaftlichen Debatte in den letzten Monaten verunsichert.

Gibt es tatsächlich gute Gründe anzunehmen, dass von KI eine existenzielle Gefahr für die Menschheit ausgeht? Schließlich ist diese Debatte sehr alt.¹ Bereits in den 1950er Jahren warnte der britische Mathematiker Alan Turing vor der Möglichkeit, dass Maschinen die Kontrolle übernehmen könnten. Über die Jahrzehnte ebten die Warnungen nicht ab, doch die KI-Apokalypse blieb aus. Dies erinnert an Äsop Fabel vom Jungen, der zu oft »Wolf« rief. Andererseits kam der Wolf in der Fabel am Ende eben doch. Daher diskutiert dieser Artikel zunächst, wie stichhaltig die Argumente für KI als existenzielle Gefahr sind und was Kritiker dagegen vorbringen.

Adressieren die globalen KI-Regulierungen die Gefahren, die von KI ausgehen? Unbestritten ist, dass KI Gefahren mit sich bringt – ob diese nun von existenzieller Natur sind oder nicht. Die global ambitionierteste Bemühung, KI zu regulieren, ist zweifellos der AI Act der Europäischen Union. Nach Jahren der Diskussion steht das Gesetzesvorhaben kurz vor seinem Abschluss; die Trilogverhandlungen zwischen EU-Parlament, EU-Kommission und EU-Rat begannen im Juni, und der Diskussionsbedarf reißt nicht ab. Doch auch andere Länder führen zunehmend Regeln für den Umgang mit KI ein. Ob und inwiefern diese

Regulierungen auf Gefahren tatsächliche und potenziell existenzielle Gefahren eingehen, wird im zweiten Teil dieses Artikels besprochen.

WIE KÖNNTE ES ZU EINER EXISTENZIELLEN GEFAHR DURCH KI KOMMEN?

Zunächst lohnt es sich zu klären, was in der Debatte als »existenzielle Gefahr« verstanden wird. Erstens sind damit Gefahren gemeint, die die Menschheit als Ganzes betreffen, wie die oben erwähnten Beispiele eines Atomkriegs oder Pandemien. Zweitens werden damit meistens auch Gefahren gemeint, die nicht die ganze Menschheit, aber einzelne Gesellschaften betreffen (Critch und Russell 2023), wie z. B. die Destabilisierung einer Demokratie.

Trotz der atemberaubenden Fortschritte, die im Bereich KI in den letzten Monaten erzielt wurden, mag die Vorstellung, dass KI eine existenzielle Bedrohung für die Menschheit darstellt, für viele Menschen absurd erscheinen. Und doch gibt es die oben beschriebene Debatte. Wie genau könnten wir uns also auf den Punkt zubewegen, an dem KI zu einer Gefahr für die gesamte Menschheit wird?

Zwar existiert keine allgemein anerkannte Theorie, doch es lassen sich einige Gemeinsamkeiten in den Argumenten derjenigen ziehen, die das Potenzial für eine solche Gefahr sehen:

- Die große Mehrheit aller Forschenden ist sich einig, dass menschenähnliche Intelligenz auf nicht-biologischer Grundlage möglich ist.
- Die Geschwindigkeit mit der KI neue Fähigkeiten erlernt, hat sich in den letzten Jahren beschleunigt: Wir bewegen uns mit großen Schritten auf eine siliziumbasierte menschenähnliche Intelligenz zu.



Johannes Walter

ist wissenschaftlicher Mitarbeiter am ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung, Mannheim, und Doktorand am KIT.

¹ Bemerkenswerterweise stand eine Diskussion über die Risiken von KI nicht auf der Agenda des Dartmouth Summer Research Project on Artificial Intelligence von 1956, das gemeinhin als Wiege der KI-Forschung betrachtet wird (Yudkowsky 2008).

- Eine KI, die vordergründig erscheint, als ob sie mit menschenähnlichen Fähigkeiten ausgestattet wäre, sei tatsächlich bereits sehr viel intelligenter als ein Mensch aufgrund der überlegenen Speicher-, Rechen- und Datentransfermöglichkeiten von Computern.
- Eine KI mit menschenähnlichen Planungsfähigkeiten und trainiert auf dem gesamten Internet hätte das Wissen und die Möglichkeit, sich immer weiter zu verbessern. Entweder unter Mithilfe von Menschen oder möglicherweise sogar autonom.

An diesem Punkt kommt das sogenannte Alignment-Problem auf: Wie stellen wir sicher, dass eine KI, die intelligenter ist als wir und die möglicherweise sogar autonom agiert, sich so verhält, dass ihre Entscheidungen auf einer Linie mit menschlichen Normen und Werten sind? Dieses Problem ist ungelöst und aktiver Forschungsgegenstand (Taylor et al. 2020).

Aber selbst eine KI, deren Fähigkeiten und Autonomie noch nicht an die von Menschen heranreicht, könnte schon ausreichen, um große Probleme zu verursachen. Kriminalität, Terrorismus und Krieg können Gründe liefern, weshalb Menschen KI schädlich einsetzen wollen. Aber eine Gefahr könnte auch unabsichtlich entstehen: Viele einzelne Akteure, z. B. in der Open Source Community (Meta 2023), arbeiten individuell rational an technischen Neuerungen; dabei möchte keiner eine außer Kontrolle geratene Superintelligenz. Letztlich tragen aber dennoch alle zu ihrem Entstehen bei, da sie die Externalitäten ihrer Beiträge nicht berücksichtigen (ähnlich dem Problem, das in der Ökonomie als »Tragik der Allmende« bekannt ist).

KRITIK AN DER THESE DER EXISTENZIELLEN BEDROHUNG

Es herrscht nicht nur eine angeregte Debatte darüber, wie es zu einer existenziellen Gefahr kommen könnte, sondern auch über Sinn und Unsinn der Debatte an sich. Aber warum sollte es schlecht sein, sich KI als existenziellem Risiko zu widmen? Selbst wenn die Wahrscheinlichkeit, dass KI das Überleben der Menschheit bedroht, gering ist, wäre es nicht besser, sich mit dieser Möglichkeit auseinanderzusetzen? Im Wesentlichen gibt es drei Argumente, die gegen eine intensive Beschäftigung mit dieser Frage vorgebracht werden.

Das erste Argument lautet, dass selbst die aktuell besten KI-Systeme noch nicht »echt« intelligent sind und es noch sehr lange dauern werde, bis dieser Zustand erreicht werden könnte. Bei der Frage nach der verbleibenden Dauer, bis eine mögliche Superintelligenz erreicht ist, herrscht nach wie vor Uneinigkeit. Allerdings mangelt es auch nicht an Ideen, welche Zutaten noch fehlen (LeCun 2022).

Das zweite Argument ist, dass wir uns besser den heute schon gegebenen Problemen durch KI widmen sollten (Tucker 2023). Wir haben nur limitierte Mengen

an Geld, Aufmerksamkeit und Hysterie. Den Gefahren, die von einer Superintelligenz ausgehen, wird man sich besser widmen können, wenn es soweit ist (falls es überhaupt dazu kommt), so das Argument. Bis dahin sollten wir uns also besser den heute schon gegebenen Gefahren durch KI widmen. Welche Gefahren sind das? Zu den größten aktuellen Problemen durch KI gehören:

- Ermöglichung und Verstärkung von Diskriminierung benachteiligter Gruppen,
- Verbreitung von Desinformation und Deepfakes,
- die hohen CO₂-Ausstöße während des Trainings großer KI-Modelle,
- Cybersicherheitsrisiken: Große Sprachmodelle können genutzt werden, um Schwachstellen aufzufinden und Malware zu schreiben,
- Ermöglichung von staatlicher Massenüberwachung,
- traumatisierende Arbeitsbedingungen für Menschen, die Trainingsdaten erstellen,
- Konzentration von Macht in den Händen einiger weniger Unternehmen.

Das dritte Argument ist schließlich, dass eine übermäßige Fokussierung auf die potenziellen Gefahren der KI dazu führt, die Vorteile zu verpassen, die diese Technologie in Bereichen wie Produktivität, Medizin, Unterhaltung und ähnlichem mit sich bringen kann. Unser Lebensstandard könnte länger als notwendig hinter einem möglichen Optimum zurückbleiben.

Es liegt in der Natur der Sache, dass sich diese Debatte (noch) nicht empirisch beantworten lässt. Dagegen lässt sich bereits analysieren, ob und wie die aktuelle Regulierung von KI auf die oben beschriebenen möglichen und realen Gefahren eingeht.

EUROPÄISCHE UND INTERNATIONALE KI-REGULIERUNG

Im Frühjahr 2021 legte die Europäische Kommission den ersten Entwurf des AI Act vor (EU-Kommission 2021). Im Sommer 2023 befindet sich der AI Act nun im Trilog, dem Verhandlungsverfahren zwischen der Europäischen Kommission, dem Europäischen Parlament und dem Rat der Europäischen Union. Die zentralen Bestimmungen des AI Act umfassen:

- Eine risikobasierte Regulierung: Das Prinzip lautet, je höher das Risiko durch die KI, desto strenger die Regulierung. Dies reicht von keiner Regulierung für harmlose Anwendungen, über zunehmend strengere Regulierung bis hin zu vollständigen Verboten.
- Vollständig verboten sind: Die Bewertung der Vertrauenswürdigkeit von Personen, basierend auf ihrem Sozialverhalten (sogenannte Social Scoring Algorithmen), unbewusste Verhaltensänderung durch KI-Systeme, Ausnutzen von Schwächen von Menschen mit Behinderungen und der Einsatz von

Erkennungssystemen aufgrund von biometrischen Daten (z. B. Gesichtserkennung) in Echtzeit im öffentlichen Raum (Art. 5 AI Act).

- Umfassende Dokumentations- und Auskunftspflichten: Fairnessabwägungen müssen bereits im Entwicklungsprozess berücksichtigt und Sorgfaltspflichten dokumentiert werden (Art. 17 und Art. 18 AI Act).
- Sichere Hochrisiko-KI-Anwendungen sollen vor allem durch zwei Mechanismen sichergestellt werden: durch menschliche Aufsicht über KI (Human Oversight, Art. 14 AI Act) und durch Konformitätsbewertungen (Conformity Assessments, Art 19 und Art. 43 AI Act).

Weitere Probleme werden in der EU von begleitender Gesetzgebung adressiert: Die zunehmende Marktkonzentration von einigen wenigen Unternehmen im Digital Markets Act (European Union 2022), der Effekt von KI als Empfehlungsmechanismus von Inhalten im Digital Services Act (European Union 2022). Außerdem wird die AI Liability Directive (European Commission 2022) Regeln zur zivilrechtlichen Haftung von KI-Anwendungen einführen.

In den USA wurde 2022 vom Weißen Haus der Blueprint for an AI Bill of Rights (White House 2022) veröffentlicht. Im Gegensatz zum AI Act sind die dort formulierten Regeln allerdings nicht verpflichtend. Außerdem lud US-Präsident Joe Biden im Februar die Chefs von Meta, Google und OpenAI zu einer Krisensitzung ins Weiße Haus (White House 2023). Ein Ergebnis des Treffens ist die Veranstaltung eines AI Hackathons, also dem Versuch, KI-Modelle in einer sicheren Umgebung zu attackieren, um mögliches Fehlverhalten der KI aufzudecken (AI Village 2023). US-Präsident Joe Biden und UK-Premierminister Rishi Sunak kündigten im Juni 2023 zudem eine multilaterale Initiative zu KI-Sicherheit an (Sunak und Biden 2023).

Auch China hat eine Reihe von Regulierungsvorhaben zu KI auf den Weg gebracht (Sheehan 2022). Insbesondere sind hier die Regeln zur Entwicklung und dem Einsatz generativer KI (Digichina Stanford 2023) zu nennen, wobei China einen anderen Schwerpunkt als die EU setzt. Dies zeigt sich z. B. bei der umfassenden anlasslosen Massenüberwachung im öffentlichen Raum in China.

ADRESSIERT DER AI ACT MIT DIESEN PUNKTEN DIE OBEN BESCHRIEBENEN GEFAHREN?

Es liegt in der Natur der Sache, dass der AI Act intensiver auf konkrete, bereits heute existierende Gefahren eingehen kann als auf existenzielle Gefahren, die möglicherweise weiter in der Zukunft liegen. Zum Beispiel können die Verbote von biometrischen Identifikations- und Social-Scoring-Systemen EU-Bürger schützen.

Weniger konkrete Gefahren sind dagegen natürlich schwieriger zu regulieren. Wenn es noch keine autonome Superintelligenz gibt, wie soll man dann

Regeln schaffen, die für Sicherheit sorgen? Doch der AI Act geht auch hier in die richtige Richtung. Umfassende Dokumentations- und Auskunftspflichten für Unternehmen, die KI entwickeln, können dafür sorgen, dass der Prozess auf dem Weg zu immer intelligenteren Systemen so abläuft, dass Forscher, Zivilgesellschaft und Politik zu jedem Zeitpunkt Einblick erhalten und möglicherweise intervenieren können.

Natürlich bleiben viele Fragen auch mit der Verabschiedung des AI Acts offen. Wie kann man funktionierende menschliche Aufsicht über KI herstellen? Aktuelle empirische Studien legen nahe, dass noch viel anwendungsnahe Forschung notwendig ist, um diese Frage befriedigend beantworten zu können (Biermann et al. 2022). Wie müssen die Konformitätsbewertungen ausgestaltet sein, um einen sinnvollen Eindruck über die Qualität eines KI-Modells zu gewinnen? Auch hier werden Politik und Unternehmen noch vieles dazulernen müssen. Initiativen, wie der oben erwähnte AI Hackathon der US-Regierung, und weitere Aufmerksamkeit für dieses Problem in der Forschung Community in diesem Bereich (Raji et al. 2022) sind notwendig. Doch insgesamt geht vor allem der AI Act der Europäischen Union, aber auch die weiteren angesprochenen internationalen Regulierungen auf die Gefahrenlage ein.

FAZIT: FLEXIBLE UND INTERNATIONALE REGULIERUNG NOTWENDIG

Gibt es tatsächlich gute Gründe anzunehmen, dass von KI eine existenzielle Gefahr für die Menschheit ausgeht? So rasant die technische Entwicklung aktuell auch ist, so unvorhersehbar ist sie auch. Niemand weiß, ob sie auch in den nächsten zehn Jahren so schnell weitergehen wird wie in den letzten zehn Jahren oder ob die Entwicklung abflacht. Die Wahrscheinlichkeit, dass eine Superintelligenz den Fortbestand der Menschheit gefährdet, scheint aktuell immer noch klein, jedoch sind die Argumente der Befürworter zumindest konsistent. Doch selbst auf dem aktuellen technischen Stand stellen KI-Systeme zweifellos eine gesellschaftsweite Gefahr dar.

Adressieren die globalen KI-Regulierungen diese Gefahren? Der EU AI Act macht vieles richtig. Es ist nun notwendig, die oft vagen Vorgaben des AI Act zu konkretisieren: Wie kann man funktionierende menschliche Aufsicht über KI herstellen? Wie müssen die Konformitätsbewertungen ausgestaltet sein, um einen sinnvollen Eindruck über die Qualität eines KI-Modells zu gewinnen? Die Regulierung hat mit dem AI Act gerade erst begonnen und muss mit der technischen Entwicklung Schritt halten. Dafür sollte der AI Act mit der Zeit flexibel weiterentwickelt werden. Unternehmen und auch Forschende sollten dies als Chance begreifen.

Europa ist ein wichtiger Markt für KI-Anwendungen, doch die wichtigsten technischen Innovationen kommen aus den USA und aus China. Insofern ist es

zu begrüßen, dass auch in diesen Ländern KI-Regulierungen auf den Weg gebracht werden. Darauf sollte aufgebaut werden und in den Bereichen, in denen Einigkeit herzustellen ist, sollte eine internationale Lösung gesucht werden. Ziel könnte eine Art Klimaabkommen von Paris für KI sein. Die Initiative von US-Präsident Joe Biden und UK-Premierminister Rishi Sunak sollte daher von der EU aufgegriffen werden.

Wenn die KI-Regulierung global koordiniert wird und sich flexibel an die technische Entwicklung anpasst, stehen die Chancen gut, dass wir die Gefahren von KI eindämmen und die Möglichkeiten dieser faszinierenden Technologie optimal nutzen.

REFERENZEN

AI Village (2023), »DEFCON 31«, verfügbar unter: <https://aivillage.org/defcon31/>.

Biermann, J., J. Horton und J. Walter (2022), »Algorithmic Advice as a Credence Good«, ZEW Discussion Paper No. 22-07, Mannheim.

Center for AI Safety (2023), »Statement on AI Risk«, verfügbar unter: <https://www.safe.ai/statement-on-ai-risk>.

Critch, A. und S. Russell (2023), »TASRA: A Taxonomy and Analysis of Societal-Scale Risks from AI«, verfügbar unter: <https://arxiv.org/pdf/2306.06924v2.pdf>.

Digichina Stanford (2023), »Translation: Measures for the Management of Generative Artificial Intelligence Services«, verfügbar unter: <https://digichina.stanford.edu/work/translation-measures-for-the-management-of-generative-artificial-intelligence-services-draft-for-comment-april-2023/>.

EU-Kommission (2021), »AI Act Draft«, verfügbar unter: <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0206>.

European Commission (2022), »Liability Rules for Artificial Intelligence«, verfügbar unter: https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en.

European Union (2022), »Digital Markets Act«, verfügbar unter: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925>.

European Union (2022), »Digital Services Act«, verfügbar unter: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065&qid=1666857835014>.

Future of Life Institute (2023), »Pause Giant AI Experiments: An Open Letter«, verfügbar unter: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

LeCun, Y. (2022), »A Path Towards Autonomous Machine Intelligence«, verfügbar unter: <https://openreview.net/forum?id=BZ5a1r-kVsf>.

Meta (2023), »Meta and Microsoft Introduce the Next Generation of Llama«, verfügbar unter: <https://about.fb.com/news/2023/07/llama-2/>.

Raji, I. D., P. Xu, C. Honigsberg und D. E. Ho (2022), »Outsider Oversight: Designing a Third Party Audit Ecosystem for AI Governance«, verfügbar unter: <https://doi.org/10.48550/arXiv.2206.04737>.

Sheehan, M. (2022), »China's New AI Governance Initiatives Shouldn't Be Ignored«, verfügbar unter: <https://carnegieendowment.org/2022/01/04/china-s-new-ai-governance-initiatives-shouldn-t-be-ignored-pub-86127>.

Sunak, R. und J. Biden (2023), »President Biden and U.K. Prime Minister Rishi Sunak Hold News Conference at White House«, BCS News, verfügbar unter: <https://www.youtube.com/watch?v=SzMYpYwPpul&t=750s>.

Taylor, J., E. Yudkowsky, P. LaVictoire, A. Critch (2020), »Alignment for Advanced Machine Learning Systems. Ethics of Artificial Intelligence«, verfügbar unter: <https://intelligence.org/files/AlignmentMachineLearning.pdf>.

Tucker, I. (2023), »Signal's Meredith Whittaker: 'These Are the People Who Could Actually Pause AI if They Wanted To'«, verfügbar unter: <https://www.theguardian.com/technology/2023/jun/11/signals-meredith-whittaker-these-are-the-people-who-could-actually-pause-ai-if-they-wanted-to>.

White House (2022), »Blueprint for an AI Bill of Rights«, verfügbar unter: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>.

White House (2023), »Readout of White House Meeting with CEOs on Advancing Responsible Artificial Intelligence Innovation«, verfügbar unter: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/05/04/readout-of-white-house-meeting-with-ceos-on-advancing-responsible-artificial-intelligence-innovation/>.

Yudkowsky, E. (2008), »Artificial Intelligence as a Positive and Negative Factor in Global Risk«, *Global Catastrophic Risks*, 308–345.

Dirk Heckmann

ChatGPT, XR und Co. vor Gericht

Plädoyer für einen Rechtskulturwandel bei digitalen Innovationen

Der berühmte Netzphilosoph Rafael Capurro hat einmal formuliert:

»Die Innovation, das Neue, ist eine Anomalie gegenüber dem Alten. Eine Anomalie widerspricht dem Alten, dem Bisherigen und wirkt wie ein Fehler. Daher kommt auch das Problem der Innovation, dass sie sich gegenüber dem Bewährten erstmal durchsetzen muss.«

Vielleicht erklärt dies den Umstand, warum wir uns schwertun, technische Innovationen im Prozess der digitalen Transformation zu akzeptieren, weil mit ihnen Risiken verbunden sind. Nur: Wer betrachtet die Risiken, die dadurch entstehen, dass Digitalisierung ausgebremst wird und Innovationen sich nicht entfalten können? Ausgebremst nicht zuletzt

durch rechtliche Bedenken, allen voran »wegen des Datenschutzes«.

Besonders schön zu beobachten war und ist dies derzeit bei der Etablierung generativer Künstlicher Intelligenz (KI), etwa zur automatisierten Texterstellung. Hier erleben wir durch ChatGPT von dem Unternehmen OpenAI eine Art »iPhone-Moment«. Genauso wie dieses Smartphone 2007 mit seiner intuitiven Bedienung den IT-Markt revolutionierte und auch weniger technikaffinen Menschen die Welt der Apps auf Mobiltelefonen erschloss, hat ChatGPT gute 15 Jahre später eine KI-Anwendung in zahllose Haushalte gebracht. Dabei dient ChatGPT nicht nur der passiven Nutzung, wie dies zuvor schon bei einfachen Chatbots oder automatisierter Spracherkennung (Alexa, Siri

etc.) der Fall war. Vielmehr können KI-Textgeneratoren nun von allen Menschen aktiv und kreativ gestaltend genutzt werden. Ein Grund zur Freude, Neugier und Ansporn? Nicht unbedingt für Juristen. Der erste Gedanke solcher Bedenkenträger ist eher: Geht das überhaupt? Ist das nicht rechtswidrig? Muss man das nicht verhindern?

CHATGPT IM SPIEGEL DES DATENSCHUTZRECHTS

Gelegenheit für solche Bedenken bietet par excellence gerade bei digitalen Innovationen das Datenschutzrecht, die Datenschutzgrundverordnung (DSGVO). Und so dauerte es nicht lange, bis die Datenschutzaufsicht in Italien Ende März 2023 ChatGPT verbot, weil es datenschutzwidrig sei. Begründet wurde dies unter anderem mit einem Verstoß gegen Art. 5 Abs. 1 lit. d DSGVO, den Grundsatz der sog. Datenrichtigkeit. Danach müssen personenbezogene Daten sachlich richtig und erforderlichenfalls auf dem neuesten Stand sein. Dieser Vorwurf ist insofern bemerkenswert, als generative KI wie ChatGPT gar nicht den Anspruch erhebt (und erheben kann), immer »richtig« zu sein. Solche Systeme des Maschinellen Lernens (hier: von KI-Textgeneratoren) basieren auf neuronalen Netzen, die darauf trainiert wurden, Muster in großen Datensätzen zu erkennen. Dieser komplexe Vorgang beruht letztlich auf einer Wahrscheinlichkeitsberechnung (»Welches Wort folgt höchstwahrscheinlich auf das vorhergehende?«). Dies führt angesichts der (zumindest derzeit noch unvermeidbaren) Unvollkommenheit der Datenbasis immer wieder zu sogenannten Halluzinationen, also der Ausgabe fehlerhafter Informationen. Hierin einen Verstoß gegen den Grundsatz der Datenrichtigkeit anzuprangern, zeugt von einem Missverständnis der Funktionsweise solcher KI, die eben kein erweitertes Online-Lexikon ist. Natürlich gibt es weitere, durchaus diskussionswürdige datenschutzrechtliche Fragen gegenüber OpenAI und seinem Produkt ChatGPT, auf die auch die deutsche Datenschutzkonferenz hingewiesen hat. Nur haben die Datenschutzbeauftragten einiger Länder (allen voran Schleswig-Holstein) diese offenen Fragen – anders als ihre Kollegen aus Italien – nicht als Anlass für ein sofortiges Verbot, sondern für einen Fragenkatalog genommen, der OpenAI übersandt wurde. Die Antworten hierauf stehen noch aus. Die italienische Datenschutzaufsicht hat Ende April 2023 ihr Verbot zurückgenommen, was genauso übereilt wirkte wie dessen ursprünglicher Erlass.

RECHTSKULTURWANDEL HIN ZU EINER KONSTRUKTIV-ABWÄGENDEN HALTUNG

Mir geht es an dieser Stelle nicht um ein abschließendes Urteil zur Datenschutzkonformität von ChatGPT und vergleichbarer generativer KI. Vielmehr möchte ich auf ein Phänomen hinweisen und hierzu den verfassungsrechtlichen Rahmen abstecken. Bedarf es eines Rechtskulturwandels – weg von destruktiven

rechtlichen Bedenken, hin zu einem konstruktiv-abwägenden Verhalten? Ich spreche bewusst von einem Wandel der Rechtskultur und nicht des Rechts, denn das Recht hat vielfach diese Agilität, die ich in der Rechtspraxis, also der Anwendung des Rechts im Einzelfall – zumal bei digitalen Innovationen – vermisste. Es geht um eine generelle Haltung. Der rechtskulturelle Wandel hin zu dieser konstruktiv-abwägenden Haltung gegenüber dem Neuen, dem Unbekannten, wäre ein Wandel in doppelter Hinsicht: »Konstruktiv« in dem Sinne, dass man auch als Jurist nicht einfach Bedenken in den Raum stellt, sondern sofort Lösungen anbietet oder zumindest anstrebt – und dabei zugleich den Wert der digitalen Innovation anerkennt. Und »abwägend« in dem Sinne, dass man bei dem (im Übrigen notwendigen) Rechtsgüterschutz nicht nur das eine Rechtsgut benennt und verteidigt, um das man sich sorgt, sondern zugleich kollidierende Rechtsgüter in den Blick nimmt, deren Wert und Wichtigkeit ebenso auf die Waagschale gehören.

ANWENDUNGSBEISPIEL VIDEOKONFERENZSYSTEME

Besonders deutlich wurde das in der Pandemie, als praktisch »von heute auf morgen« alle Einrichtungen, in denen bislang menschliche Begegnungen den Alltag prägten, ihre Aktivitäten aufgrund der Kontaktbeschränkungen in den virtuellen Raum verlagern mussten. Damit geschäftliche Besprechungen oder Schulunterricht überhaupt noch stattfinden konnten, musste dies über Videokonferenzsysteme organisiert werden – und zwar schnell und reibungslos. Die Herausforderung: Ausgerechnet jene Anbieter, deren Produkte sich durch hohe Funktionalität, gute Performance und Skalierbarkeit auszeichneten (wie Microsoft Teams oder Zoom), haben ihren Unternehmenssitz in den USA (aus datenschutzrechtlicher Sicht also einem »unsicheren Drittstaat«), was die Frage nach einem potenziellen Zugriff dortiger Geheimdienste und Sicherheitsbehörden auf jene Daten aufwarf, die im Zuge von Videokonferenzen erfasst und verarbeitet werden. Auch hier war die erste Reaktion mancher (nicht aller) Datenschutzbehörden: verbieten oder zumindest warnen vor der Anwendung der Videokonferenzsoftware. Abgesehen davon, dass solche Produktwarnungen ihrerseits rechtlich fragwürdig sind (weil sie den strengen Anforderungen des Bundesverfassungsgerichts für wettbewerbslenkende staatliche Warnungen kaum genügen): Könnte man nicht mit einer anderen Haltung an diese Herausforderungen herangehen – eben konstruktiv-abwägend? Konstruktiv in dem Sinne, dass man



Prof. Dr. Dirk Heckmann

ist Inhaber des Lehrstuhls für Recht und Sicherheit der Digitalisierung an der TU München. Dort leitet er auch das TUM Center for Digital Public Services. Nebenamtlich ist er Direktor am Bayerischen Forschungsinstitut für Digitale Transformation und Mitglied des Bayerischen Verfassungsgerichtshofs.

erst nachdenkt und dann auch verlaublich: Wir haben eine große Herausforderung, trotz Kontaktbeschränkungen unternehmerisches bzw. schulisches Wirken aufrechtzuerhalten; wir haben technische Lösungen, die wohl funktionieren und in dieser Situation hilfreich sind – wie können wir das nun so gestalten, dass die digitale Innovation zum Einsatz kommt? Abwägend in dem Sinne, dass man erst nachdenkt und dann so entscheidet: Das Recht auf informationelle Selbstbestimmung mag gefährdet sein; aber auf der anderen Seite stehen auch weitere Grundrechte wie das Grundrecht auf Bildung und Chancengleichheit oder auch die unternehmerischen Freiheitsrechte. In einer Abwägung können Gefährdungen des einen Grundrechts in Kauf genommen werden, wenn die Interessen der anderen Grundrechte im Einzelfall überwiegen. »Geht nicht wegen Datenschutz« ist schon deshalb als Grundsatz untauglich, weil kein Grundrecht per se überwiegt, sondern Grundrechtskollisionen durch Abwägung aufgelöst sind.

GRUNDRECHTSGEFÄHRDUNG DURCH UNTERLASSENE DIGITALISIERUNG

Das war in der Pandemie vielfach aufgrund der vitalen Interessen des Grundrechts auf Schutz von Leben und Gesundheit der Fall, weshalb andere Freiheitsrechte, auch das Recht auf informationelle Selbstbestimmung zurücktreten mussten. Genau genommen geht es selbst dann nicht um eine Bedeutungslosigkeit des Datenschutzes, sondern um eine angemessene, zuweilen pragmatische Angleichung dieser Interessen. Das Verfassungsrecht spricht von »praktischer Konkordanz« (Hesse). Wir dürfen uns nicht nur fragen, ob Digitalisierung vielleicht den Datenschutz gefährden kann, sondern müssen zugleich fragen, welche Grundrechte gefährdet sind, wenn wir digitale Innovationen untersagen, sie nicht fördern oder vor ihnen gewarnt wird (wie das leider bei der elektronischen Patientenakte der Fall war). Es gibt auch ein Untermaßverbot, also den verfassungsrechtlichen Appell an den Staat, nicht zu wenig zu tun. Hinsichtlich der defizitären Klimaschutzpolitik leitet das Bundesverfassungsgericht hieraus ein intertemporales Freiheitsrecht zum Schutz künftiger Generationen her.

Bei alledem soll nicht übersehen werden, dass es auch Missbrauch und unredliche Geschäftsmodelle in digitalen Kontexten gibt, denen man regulatorisch und durch strengen Vollzug vorhandener Schutznormen begegnen muss. Dies aber zum Anlass zu nehmen, hinter jeder digitalen Innovation primär Rechtsverletzungen zu sehen, schafft unterdessen eine Rechtskultur permanenter Bedenklichkeit, die jeglichem Rechtsschutz einen Bärendienst erweist. Wenn nämlich hierzulande digitale Innovationen ausgebremst werden, verlieren wir nicht nur digitale Souveränität. Am Ende setzen sich am Markt digitale Anwendungen aus weniger regulierten Ländern durch, deren Einsatz nicht minder Grundrechte hierzulande gefährdet.

VORGESCHOBENE RECHTLICHE BEDENKEN ALS INNOVATIONSHEMMNIS

Aber es gibt eine noch größere Gefahr, die durch die fast reflexhafte Abwehr digitaler Innovationen hervorgerufen wird: Verfestigt sich nämlich das Narrativ der latent rechtswidrigen digitalen Technologien, hemmt das die Innovationsfreude besonders bei Start-ups und sehen sich auch jene »Bewahrer des Status quo« bestätigt, die eine Veränderung per se nicht wollen, so dass die Berufung auf »Compliance« gerade recht kommt.

Der Einsatz generativer KI wie ChatGPT ist da ein gutes Beispiel, gehen mit ihm doch auch Sorgen um Arbeitsplätze oder das Beharren auf lieb gewonnene (ineffiziente) Geschäftsprozesse einher. Dass man diese Innovation freilich weder durch Schlechttreden noch durch Vorwürfe der Rechtswidrigkeit aufhalten kann, zeigt der soeben veröffentlichte Hype Cycle Report von Gartner 2023, der wie folgt eingeleitet wird: »Generative künstliche Intelligenz und ChatGPT haben die KI-Diskussion auf ein neues Niveau gehoben und das Interesse auf Vorstandsetagen und Staatsoberhäupter gleichermaßen gelenkt. Die Verantwortlichen für Daten und Analysen müssen die Trends im Auge behalten und die Entwicklung von Innovationen verfolgen, um glaubwürdige Argumente für Investitionen zu schaffen.« Und weiter: »Die generative KI hat wie keine andere Technologie in den letzten zehn Jahren Auswirkungen gehabt. Die gesteigerte Produktivität von Entwicklern und Wissensarbeitern, die Systeme wie ChatGPT nutzen, ist sehr real und hat Unternehmen und Branchen dazu veranlasst, ihre Geschäftsprozesse und den Wert der menschlichen Ressourcen zu überdenken. Im Gegenzug haben die offensichtlichen Fähigkeiten generativer KI-Systeme die Debatte über den sicheren Einsatz von KI und die Frage, ob eine allgemeine künstliche Intelligenz erreicht werden kann oder sogar schon vorhanden ist, neu entfacht.«

ANWENDUNGSBEISPIEL EXTENDED REALITY

Die Debatte sollte sich freilich nicht auf generative KI beschränken. Auch andere digitale Innovationen werden durch unüberlegt oder zumindest undifferenziert vorgebrachte rechtliche Bedenken ausgebremst. So gilt Ähnliches in Bezug auf Anwendungen der Virtual Reality oder Augmented Reality (zusammen Extended Reality, XR). Trotz der unbestreitbaren Vorteile von XR nutzen Unternehmen bisher kaum die Möglichkeiten dieser neuen Technologie. In der vbw-Studie »Extended Reality – Zukunftstechnologie mit breitem Anwendungsspektrum« von September 2022 wurde festgestellt, »dass weniger als jeder zehnte befragte (bayerische) Unternehmensvertreter bereits Erfahrungen mit XR im privaten oder beruflichen Kontext gemacht hat. Ein Grund hierfür ist, dass Unsicherheiten in Bezug auf die rechtlichen Rahmenbedingungen beim Einsatz von XR bestehen.« Dabei lassen sich alle rechtlichen Bedenken gegenüber XR-Technologien im

Unternehmen zerstreuen, wie ich in dem im Juli 2023 erschienenen Rechtsmonitor XR¹ aufgezeigt habe. Datenschutzrechtlich kommt es etwa auf einen gut gestalteten Einwilligungsprozess zwischen Unternehmen und betroffenen Beschäftigten an, der die Freiwilligkeit des Einsatzes unterstreicht. Ähnlich kann auch dem Arbeitsschutz Rechnung getragen werden, um Nachteile wegen »Motion Sickness« zu begrenzen; hier können Betriebsvereinbarungen hilfreich sein. Nicht nur für diese Risiken, sondern auch für die Bedeutung des Schutzes von Geschäftsgeheimnissen ist zu sensibilisieren. All dies ist nicht trivial, aber umsetzbar, wenn ein Unternehmen professionell arbeitet.

ERMUTIGUNG ZU SINNSTIFTENDER VERÄNDERUNG

Und das zeigt ein weiteres Mal, wie wichtig dieser Rechtskulturwandel ist. Wenn man stärker auch die Chancen der Digitalisierung betont, den rechtlichen Herausforderungen weniger mit Bedenken als mit Hinweisen auf rechtliche Gestaltungsmöglichkeiten (auch im Sinne eines »Legal Design Thinking«) begegnet und insgesamt lösungsorientiert statt problemorientiert agiert, dann ermutigt man sowohl die Entwickler als auch deren potenzielle Nutzer zu Fortschritt und sinnstiftender Veränderung. Dieser Wille zu Veränderung braucht auch zuweilen Mut. Aktuell beobachten wir hierzulande hingegen immer wieder diese »German Angst«, mit der Veränderung etwas falsch zu machen, was auch mit unserem Haftungs-

¹ Siehe: <https://www.vbw-bayern.de/vbw/Themen-und-Services/Digitalisierung/Technologien-und-Geschäftsfelder/Rechtssicher-Einsatz-von-Extended-Reality-Anwendungen.jsp>.

system zusammenhängt: Man haftet in der Regel nur für Schäden, die auf einem aktiven, fehlerhaften Tun beruhen. Für Schäden, die auf ein Unterlassen zurückgeführt werden können, haftet man nur, wenn das Unterlassen pflichtwidrig ist, also eine Pflicht zum Handeln bestand. Eine solche findet sich in unserer Rechtsordnung aber nur selten.

So mag es zwar schwer zu berechnen sein, gleichwohl sei das Gedankenexperiment erlaubt: Welche volkswirtschaftlichen, aber zugleich individuellen Schäden entstehen tagtäglich dadurch, dass weder unsere Verwaltung noch unser Bildungswesen oder Gesundheitswesen hinreichend digitalisiert sind? Warum nimmt man diese Schäden in Kauf, während man ein einzelnes Datenschutzrisiko nicht hinnehmen will?

GLEICHKLANG VON TECHNISCHER UND RECHTLICHER ENTWICKLUNG

Um schließlich nicht missverstanden zu werden: Ich möchte demokratisch legitimes Recht keineswegs ignorieren. Wenn es gute rechtliche Argumente gegen eine bestimmte Technologie, besser gesagt: einen bestimmten Einsatz dieser Technologie gibt, dann gilt es, diese einzuhegen. Notfalls versucht man, das Recht neuen Gegebenheiten anzupassen, was (wie bei der KI-Verordnung zu befürchten ist) aber auch zu Über- oder Fehlregulierung führen kann. Juristische »Bedenkenträgerei« ist allemal unterkomplex und verliert das größere Ganze, das eine grundrechtlich ausgewogene Rechtslage auszeichnet, aus dem Blick: Digitale Innovationen sind keine Anomalie, sondern die Überwindung des Alten, das aus der Zeit gefallen ist.

Armin Grunwald

Allmähliche Disruptionen mit KI?

KI-Systeme lenken Autos, wählen Stellenbewerber aus, komponieren Musikstücke, schreiben Texte in erstaunlicher Sprachqualität und machen Roboter durch kommunikative Fähigkeiten zu neuen Mitbewohnern der menschlichen Welt. Je nach Perspektive ist es faszinierend oder erschreckend, oder beides gleichzeitig, wie gut KI viele menschliche Fähigkeiten simulieren kann und wie rasch der Fortschritt dabei ist.

Neben vielfältigen Hoffnungen auf Innovation und Effizienzsteigerung durch KI sind Sorgen bis hin zu apokalyptischen Befürchtungen verbreitet. Kontrollverlust, Machtübernahme durch Algorithmen, Zusammenbruch des Arbeitsmarkts und Manipulation durch KI sind wiederkehrende, teils aus der Science Fiction wohlbekannte Elemente. Prominente wie Elon Musk

oder Yuval Harari warnen immer wieder. Nachdem ChatGPT freigegeben war, wurden gleich zwei martialische Memoranden unter Beteiligung auch führender KI-Experten mit Forderungen nach einem weltweiten Moratorium für KI-Entwicklung publikumswirksam veröffentlicht.

An dieser Stelle soll es um einen anderen Typ von Herausforderungen gehen: um mögliche Disruptionen, die sich allmählich einstellen und deren Ursachen weniger in den Algorithmen als vielmehr in menschlichen Verhal-



Prof. Dr. Armin Grunwald

leitet das Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) und ist Professor für Technikphilosophie am Institut für Philosophie des KIT.

tensweisen liegen – und die trotz möglicher Relevanz bislang wenig beachtet werden.

ALLMÄHLICHE DISRUPTIONEN

Disruption ist zu einem prägenden Begriff der Gegenwart geworden. Zunächst geschah dies in positiver Hinsicht: Disruptive Innovation im wirtschaftlichen Wettbewerb gilt als attraktiv, jedenfalls für die Gewinner. Zu ihrer Förderung wurde durch die deutsche Bundesregierung die Agentur für Sprunginnovationen gegründet (SPRIND). Zurzeit werden jedoch mit Disruption eher aktuelle Krisenphänomene wie Pandemie und Ukraine-Krieg verbunden, die sehr schnell, fast plötzlich aufgetreten sind und praktisch sofort schwere Verwerfungen nach sich gezogen haben.

Das Wort »Disruption« verheißt nichts Gutes, Latein *disrumpere* = platzen, zerbrechen, zerreißen. Dies passt zum aktuell vielfach beschworenen Übergang in eine Zeit der permanenten Krise mit dem Zerbrechen stabiler gesellschaftlicher Zustände. Da »Disruption« über das Zerbrechen hinaus auch die Schnelligkeit dieses Vorganges meint, mag es zunächst paradox klingen, wenn im Folgenden von allmählichen Disruptionen gesprochen wird. Schließlich zerbricht nichts langsam. Allerdings kann sich ein Zerbrechen allmählich über längere Zeiträume vorbereiten und erst deutlich später zu plötzlichen Brüchen führen. Materialermüdung und Verschleiß zeigen dies in der Welt der Technik. Während Materialien allmählich ermüden und Technik sich langsam im Gebrauch abnutzt, funktioniert alles dennoch verlässlich und oft noch recht lange weiter. Erst zu einem bestimmten Zeitpunkt fällt ein Bauteil von einem auf den anderen Moment aus oder eine Brücke bricht zusammen. Das Disruptive baut sich inkrementell auf, kann lange unentdeckt bleiben und dadurch dem frühzeitig intervenierenden Eingriff, etwa einer Reparatur, entgehen.

Auch aus der Umweltbelastung sind allmähliche Disruptionen bekannt. So hat sich das Ozonloch über Jahrzehnte durch stete FCKW-Freisetzung in die Atmosphäre aufgebaut, bevor die Folgen sichtbar wurden. Auch der Klimawandel, dem heute vielfach disruptive Eigenschaften zugeschrieben werden, ist allmählich durch CO₂- und Methanemissionen entstanden. Die Versauerung der Ozeane oder der Biodiversitätsverlust sind andere, ebenfalls inkrementell verlaufende Prozesse mit Disruptionspotenzial.

Allmähliche Disruptionen haben gemeinsame Muster. Zunächst sind sie (1) *schlecht erkennbar*. In frühen Phasen sind kaum Daten verfügbar, bestenfalls schwache Signale mit begrenzter Evidenz. Erst im zeitlichen Verlauf bessern sich Datenlage und Evidenz. Es besteht das Risiko erst später Entdeckung, so dass dann bereits schwer beeinflussbare Pfadabhängigkeiten eingetreten sein können und vielleicht gar ein *Point of No Return* überschritten wurde. Sodann ist es (2) eine *ethische Herausforderung*, nur schlecht erkennbaren Entwicklungen nach Dringlichkeit und Hand-

lungsnotwendigkeit einzuordnen. Ein bloßer Verdacht reicht für die Einleitung weitreichender Maßnahmen mit z. B. erheblichem Ressourcenbedarf, Belastungen oder Freiheitseinschränkungen nicht aus. Dieses aus Vorsorgeproblemen bekannte Phänomen führt (3) auf eine spezifische *kommunikative Herausforderung*. Einerseits gibt es Verharmlosung und Abwiegeln mit dem Verweis, man müsse eine bessere Datenlage abwarten, es sei zu früh zum Handeln. Auf der anderen Seite wird dringend nach Maßnahmen gerufen, dass es ansonsten zu spät sein könnte. Gegenseitige Vorwürfe von Übertreibung, Ideologie, Spekulation, Verharmlosung, Schönrederei, Verantwortungslosigkeit oder Unterstellung versteckter Interessen sind die Folge.

ALLMÄHLICHE DISRUPTIONEN DURCH KI

Die Natur allmählicher Disruptionen bringt es mit sich, dass das Sprechen und Schreiben darüber aufgrund der in frühen Entwicklungsstadien definitionsgemäß schlechten Datenlage selbst in der erwähnten Gefahr der Spekulation stehen. Die folgenden Überlegungen sind daher tentativ. Sie fokussieren auf mögliche Disruptionen, nicht auf die bereits empirisch bekannten Gefahren von KI-Systemen wie etwa Diskriminierung, Deepfakes oder Manipulation. Das Ziel ist vorsorgeorientiert: Aufmerksamkeit für mögliche Entwicklungen mit Disruptionspotenzial zu wecken, um gegebenenfalls rechtzeitig Konsequenzen ziehen zu können. Ohne Anspruch auf Vollständigkeit:

Anpassung: Digitale Technik reguliert und ändert subtil menschliches Handeln und Verhalten, wie vielfach empirisch bestätigt. Softwaresysteme steuern explizit oder implizit menschliches Verhalten. Privat geführte digitale Infrastrukturen strukturieren die politische Kommunikation, Suchmaschinen sortieren die Welt durch die von ihnen gesetzten Filter, und Online-Plattformen strukturieren Geschäftsprozesse. Diese Beobachtungen motivieren immer wieder Befürchtungen, dass menschliches Denken und Handeln durch fortschreitende Anpassung an Software-Systeme allmählich nach deren Anforderungen und Vorgaben reguliert werden könnte. Menschliche Freiheit würde leise und unbemerkt verschwinden, ausgehöhlt durch allmähliche Gewöhnung.

Abhängigkeit: Abhängigkeiten wie beispielsweise die Abhängigkeit praktisch aller Wirtschafts- und Kommunikationsprozesse vom Funktionieren digitaler Infrastrukturen sind *latente Disruptionen*. Sie bauen sich allmählich auf, können aber im Ernstfall, wenn also z. B. Internet-Kommunikation nicht mehr funktionieren würde, abrupt eintretende und möglicherweise katastrophale Folgen haben. Bei vollständiger Abhängigkeit bräuchte es nur einen Anlass wie z. B. gesellschaftliche Verwerfungen oder einen Angriff von Hackern in bislang unbekannter Größenordnung, um die Disruption auszulösen. Abhängigkeiten von KI entstehen allmählich als Folge immer weiterer Verlagerung analoger Prozesse in die zunehmend mit KI-Syste-

men durchsetzte Digitalwelt, was ja aus vielen, z. B. wirtschaftlichen Gründen sinnvoll ist. Aufgrund des Blackbox-Charakters vieler KI-Systeme wäre es umso schwerer, bei Ausfällen wieder Menschen entscheiden zu lassen, zumal die entsprechenden menschlichen Fähigkeiten möglicherweise verlorengehen, weil sie aufgrund der Automatisierung nicht mehr benötigt werden (*ironies of automation*).

Verantwortungsdiffusion: An den neu entstehenden Schnittstellen zwischen Menschen und digitalen Systemen werden Zuständigkeiten ständig neu zwischen Mensch und KI-gesteuerter Maschine verteilt, wie etwa beim Einsatz automatisierter Entscheidungssysteme (ADM). Die Lokalisierung von Verantwortung und ihre Zuschreibung an spezifische Akteure wird dadurch zusehends komplex. Zwar verbleiben Entscheidungen und damit Verantwortung prinzipiell beim Menschen, jedoch auf eine zusehends digital vermittelte und intransparente Weise. KI-Systeme schieben sich zwischen handelnde Menschen und realweltliche Effekte. Die Verantwortung wandert von Individuen wie z. B. Autofahrern in den Hintergrund, zu Unternehmen, Programmierern, Managern, Geheimdiensten oder Regulierungsbehörden. Damit steigt die Gefahr einer allmählichen »Verantwortungsdiffusion« ins Nichts oder absichtlicher Verantwortungsverschleierung. Damit würde menschliche Verantwortung substantiell entleert und zu einer bloß formalen Hülle degradiert.

Beschleunigung: Beschleunigung ist Teil des wettbewerbsorientierten Wirtschaftssystems. Allerdings ist die Beschleunigungsspirale nicht beliebig weit überdrehbar, sondern in Gefahr, die Ressourcen zu übernutzen, aus denen sie sich speist. Eine Sorge in Bezug auf weitere, KI-gestützte Beschleunigung der Digitalisierung bezieht sich auf allmähliche Folgen dieser Beschleunigung. Dies betrifft insbesondere die Frage, ob und wann immer weitere Beschleunigung gesellschaftliche und politische Bedingungen von Reflexion und Gestaltung grundsätzlich unterminieren könnte. Bereits jetzt passt das Innovationstempo der KI mit dem Analyse-, Beratungs- und Regulierungstempo politischer Maßnahmen nicht mehr unbedingt zusammen. Allmähliche Disruption wäre eingetreten, wenn das etablierte, wenngleich immer neu auszutarierende Gleichgewicht zwischen wettbewerblicher Innovation und gemeinwohlorientierter Regulierung grundsätzlich aus der Balance geraten würde.

Diese Szenarien allmählicher Disruption sind keine Prognosen, sondern beschreiben nur mögliche Entwicklungen. Zu denken gibt, dass in zumindest einem Bereich eine derartige Entwicklung längst eingetreten ist. Während mit Internet-Kommunikation noch vor etwa 20 Jahren utopische Erwartungen Richtung Demokratie verbunden waren, ist dies gekippt. Die Möglichkeit allmählicher Disruption demokratischer Systeme, verstärkt durch KI, ist längst nicht mehr nur spekulative Befürchtung. Manipulation, *Fake News*, *Hate Speech*, kommunikative Schneebal-

efekte und die durch die Anonymität leichter gemachte Enthemmung von Kommunikation sind längst Realität geworden. Die Verbindung technischer Möglichkeiten etwa der Social Media mit menschlichem Verhalten hat offenbar ein für die Demokratie disruptives Potenzial.

MENSCHLICHE URSACHEN ALLMÄHLICHER DISRUPTION

Die geschilderten möglichen Disruptionen hängen mit technischen Eigenschaften der KI zusammen, sind jedoch entscheidend mit menschlichem Verhalten verbunden. Erst dieses nimmt Anpassungen an KI-Systeme vor oder lässt sie geschehen, schafft Abhängigkeiten oder überträgt KI-Systemen Quasi-Verantwortung mit der Folge der Verantwortungsdiffusion ins Opake oder ins Nichts. Dies ist analog zu den erwähnten allmählichen Disruptionen in Bezug auf die Umwelt: Nicht die Technik ist für Ozonloch, Biodiversitätsverlust oder Klimawandel verantwortlich zu machen, sondern die menschliche Nutzung der dadurch ermöglichen Produkte und Dienstleistungen.

Ohne Zweifel macht Technik, gerade die Digitaltechnik, vielfach das Leben angenehm und komfortabel. Ein plausibler Mechanismus für das Entstehen von Abhängigkeiten besteht daher in allmählicher Gewöhnung. Sobald Routinehandlungen in Beruf oder Freizeit an digitale Systeme adaptiert wurden, ob nun mit oder ohne KI, gehört diese Technik so zum Leben, dass es ohne sie oft kaum noch vorstellbar ist. Alternative Optionen verschließen sich, Kompetenzen gehen verloren (De-Skilling). Vermeintliche Sachzwangargumente erwecken dann oft den Anschein der Alternativlosigkeit, sind jedoch nur Ausdruck der schleichend eingetretenen Pfadabhängigkeit durch allmähliche Anpassung und Gewöhnung.

Eine besondere Rolle in der Haltung zu Digital- und KI-Systemen spielt ein psychologischer Effekt, der sogenannte und empirisch abgesicherte »automation bias«. Viele Menschen vertrauen danach algorithmisch erzeugten, auf großen Datenmengen beruhenden und mit KI-unterstützten Entscheidungsverfahren berechneten Ergebnissen stärker als menschlichen Experten mit noch so viel Berufs- und Lebenserfahrung. Der Grund liegt vermutlich in Objektivitätsunterstellungen gegenüber mathematischen und datenbasierten Verfahren einerseits und einem Subjektivitätsverdacht gegenüber Menschen andererseits.

Auch spielen menschliche – häufig übertriebene – Einschätzungen der Fähigkeit der KI eine Rolle. Die Anthropomorphisierung von KI-Systemen und autonomen Robotern ist weit fortgeschritten. Häufig wird so geredet, dass KI denken und lernen kann, dass Roboter einfühlsam sind und Emotion zeigen und dass alle diese Systeme Entscheidungen treffen. Einer genaueren Begriffsanalyse hält nichts davon Stand, sind doch KI-Systeme keine denkenden, sondern auf Basis von Daten und Algorithmen rechnenden Sys-

teme. Beispielsweise entscheiden sie nichts, sondern erzielen durch ihre statistischen Analysen Ergebnisse. Von daher sind verbreitete Redeweisen, dass etwa die KI die Herrschaft übernehmen könnte, aus der Luft gegriffen. KI hat kein Bewusstsein ihrer selbst, keinen Willen, keine Intentionen. Allmähliche Disruptionen im Zusammenhang mit KI entstehen ausgehend von Menschen, letztlich durch Selbstentmündigung.

VULNERABILITÄTEN UND RESILIENZ

In langen Zeiten hoher Stabilität war das Bewusstsein verblasst oder verlorengegangen, wie fragil und vulnerebel moderne Gesellschaften sind. Gerade in Bezug die technischen Infrastrukturen wie der Digitalisierung herrscht die Annahme eines immerwährenden reibungslosen Funktionierens vor. Vorsorge wie sorgfältiges Monitoring potenziell disruptiver Entwicklungen, die Erarbeitung von Maßnahmen zum frühzeitigen Gegensteuern, die Implementierung von Redundanzen und Puffern zur Erhöhung der Resilienz und »Pläne B« für den Fall einer katastrophalen Disruption waren kaum oder gar kein Thema.

Zurzeit ist die Bereitschaft deutlich größer, sich mit der Möglichkeit von Disruptionen, auch mit sich langsam aufbauenden Disruptionen auseinanderzusetzen, z. B. durch Analysen der Vulnerabilität. Die dargelegten Überlegungen haben gezeigt, dass allmähliche Disruption dann entstehen kann, wenn technische Möglichkeiten von KI-Systemen auf bestimmtes menschliches Verhalten treffen. Es ist gerade die sozio-technische Kombination von Effekten, nicht die Technik als solche, die dazu führt. Die Schlussfolgerung ist also, nicht den Blick auf die KI als Technik zu verengen, sondern die Wechselwirkungen mit menschlichem Verhalten in den Blick zu nehmen.

REFERENZEN

- Bainbridge, L. (1983), »Ironies of Automatization«, *Automatica* 19, 775–779.
- Deutscher Ethikrat (2023), *Mensch und Maschine. Herausforderungen durch Künstliche Intelligenz*, Deutscher Ethikrat, Berlin.
- Grunwald, A. (2019), *Der unterlegene Mensch: Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern*, RIVA, München.
- Safdar, N., J. Banja und C. Meltzer (2020), »Ethical Considerations in Artificial Intelligence«, *European Journal of Radiology* 122, 108768.