

AUS POLITIK UND ZEITGESCHICHTE

Künstliche Intelligenz

Katharina A. Zweig

DROHT KI DEN MENSCHEN
ZU ERSETZEN?

Ralf Otte

INTELLIGENZ
UND BEWUSSTSEIN.
ODER: IST KI WIRKLICH KI?

Bernd Carsten Stahl

GRAUZONEN ZWISCHEN
NULL UND EINS.
KI UND ETHIK

Thorsten Thiel

KI UND DEMOKRATIE:
ENTWICKLUNGSPFADE

Hannah Ruschemeier

REGULIERUNG VON KI.
ANSÄTZE, IDEEN, PLÄNE

Christian Kellermann ·

Cornelius Markert

KI IN DER ARBEITSWELT

Florian Nuxoll

KI IN DER SCHULE

APuZ

ZEITSCHRIFT DER BUNDESZENTRALE
FÜR POLITISCHE BILDUNG

Beilage zur Wochenzeitung Das **Parlament**



Künstliche Intelligenz

APuZ 42/2023

KATHARINA A. ZWEIG

DROHT KI DEN MENSCHEN ZU ERSETZEN?

Um die Frage nach der Ersetzbarkeit des Menschen durch KI zu beantworten, ist es hilfreich, sich die Technologie hinter den Systemen anzuschauen. Wozu sind maschinelles Lernen, neuronale Netze und generative KI imstande – und wozu nicht?

Seite 04–08

RALF OTTE

INTELLIGENZ UND BEWUSSTSEIN.

ODER: IST KI WIRKLICH KI?

KI ist hervorragend geeignet, um in einem künstlichen Datenraum zu agieren. Sie hat jedoch eklatante Schwächen in natürlicher Umgebung, denn Verständnis und Wahrnehmung werden nur simuliert. Gleichwohl schickt KI sich an, ein rudimentäres Bewusstsein zu erlangen.

Seite 09–16

BERND CARSTEN STAHL

GRAUZONEN ZWISCHEN NULL UND EINS.

KI UND ETHIK

Worin bestehen ethische Probleme in der KI? Was müssen wir beachten, wenn wir über Ethik in der KI sprechen? Und wie können damit verbundene Fragen bearbeitet werden? Statt nach einfachen Antworten zu suchen, sollte man KI als Teil eines komplexen Ökosystems sehen.

Seite 17–22

THORSTEN THIEL

KI UND DEMOKRATIE: ENTWICKLUNGSPFADE

Welche Effekte hat die Verbreitung von KI auf die Struktur und Dynamik demokratischer Öffentlichkeit? Welche Möglichkeiten und Risiken ergeben sich für demokratisches Regieren? Und inwiefern verändert sich die Teilhabe von Bürgerinnen und Bürgern?

Seite 23–28

HANNAH RUSCHEMEIER

REGULIERUNG VON KI.

ANSÄTZE, IDEEN, PLÄNE

In der raschen Verbreitung von KI-Technologie manifestiert sich eine neue Form von Macht.

KI-Regulierung ist deshalb die Regulierung von Macht und deshalb eine genuin rechtsstaatliche Frage. Wie ist der Entwurf für eine umfassende KI-Verordnung der EU zu bewerten?

Seite 29–34

CHRISTIAN KELLERMANN · CORNELIUS MARKERT

KI IN DER ARBEITSWELT

In den vergangenen 200 Jahren hat sich die Produktivität im verarbeitenden Gewerbe dank des Einsatzes von Maschinen um ein Vielfaches gesteigert. Durch KI könnte die Automatisierung nun auch in die Domänen kognitiver Arbeit vordringen und neue Potenziale freisetzen.

Seite 35–40

FLORIAN NUXOLL

KI IN DER SCHULE

KI bietet sowohl Lehrkräften als auch Schülerinnen und Schülern große Chancen, Lernprozesse zu optimieren. Die Rolle der Lehrkräfte bleibt dabei zentral: Eine Schule ohne Lehrer wird auch in Zukunft eine schlechte Schule sein – eine Schule ohne KI aber auch.

Seite 41–46

EDITORIAL

„Künstliche Intelligenz hilft gegen Waldbrände“, „KI könnte deutscher Wirtschaft Milliarden einbringen“, „Krebs: Künstliche Intelligenz erkennt schädliche Mutationen“, „Gewinnt KI den nächsten Krieg?“ Schlagzeilen wie diese sind jeden Tag zahlreich zu lesen. Und schon die kleine Auswahl zeigt, welche Bandbreite und welche hohen Erwartungen mit dem Thema Künstliche Intelligenz verbunden sind. Spätestens seit das Unternehmen OpenAI im November 2022 seinen Chatbot ChatGPT vorgestellt hat, der die Fähigkeiten von KI für viele direkt erlebbar macht, ist das Thema dauerhaft in einer breiten Öffentlichkeit angekommen.

Wie so häufig in der langen Geschichte des technischen Fortschritts sind auch die Entwicklungen auf dem Feld der KI von großen Hoffnungen und mindestens ebenso großen Befürchtungen begleitet. Was tatsächlich neu ist: Erstmals fragen sich auch akademisch ausgebildete „Wissensarbeiter“, ob es ihre Jobs noch lange geben wird. Waren es früher körperlich anstrengende Arbeiten, die Menschen an Maschinen übertrugen, sind es heute Übersetzungen, Recherchen oder komplexe Analysen. Was viele dabei gleichermaßen fasziniert und gruselt, ist die Fähigkeit der KI, zu lernen und das erworbene Wissen direkt anzuwenden. In einer populären Vorstellung ist der Schritt zum „Erwachen“ der KI, zur Herausbildung eines eigenen Willens, dann nur noch ein kleiner.

Dies mag eine dystopische Übertreibung sein. Doch auch ohne autonomes Bewusstsein ist KI ein mächtiges Instrument. Mit ihrer Hilfe lassen sich Bilder und Stimmen täuschend echt fälschen, was die Gefahr von Deep Fakes erhöht und damit dem demokratischen Diskurs empfindlich schaden kann. Auch können bestehende Vorurteile algorithmisch fortgeschrieben werden, was mit Blick auf *predictive policing* oder soziale Scoring-Verfahren höchst problematisch ist. Neben der politischen Regulierung ist es daher vor allem wichtig, den mündigen Umgang mit KI zu fördern. Ob Künstliche Intelligenz am Ende „Gutes“ oder „Schlechtes“ bewirkt, liegt in unseren eigenen Händen. Denn es ist die Maschine, die von uns lernt, nicht andersherum.

Johannes Piepenbrink

ESSAY

DROHT KI DEN MENSCHEN ZU ERSETZEN?

Katharina A. Zweig

Es ist eine der Fragen, die KI-Expertinnen und -Experten am häufigsten hören: Wird Künstliche Intelligenz uns ersetzen? Und als Nächstes: Welche Arbeitsplätze werden wegfallen? Diese Fragen sind beileibe nicht neu und wurden schon früher mehrfach gestellt. Zu Beginn der 1980er Jahre versuchte man, Computer intelligent zu machen, indem man versuchte, ihnen die Welt zu erklären. Heraus kamen große Regelsysteme, zum Beispiel für Übersetzungen: Da gab es Systeme mit 50 000 Regeln und riesigen Vokabellisten, die vorgaben, welches Wort wie zu übersetzen sei. Die Ergebnisse waren überraschend gut, und schon machten die Menschen sich Sorgen. Fantastischste Ideen kamen auf, wie in kurzer Zeit alles von Maschinen übernommen werden würde.

Wenig später merkte man, dass die Systeme doch nicht so intelligent waren: Ein Programm übersetzte zum Beispiel den Satz „Wir packen das“ mit „Let’s pack it“ und den Satz „Das passt mir nicht in den Kram“ mit „That does not usually suit me in the stuff“. Es waren also insbesondere die figurativen Sätze, mit denen diese regelbasierten Systeme Schwierigkeiten hatten. Das größte Problem an den regelbasierten Systemen aber war ihre mangelnde Wartbarkeit: Wenn ein Fehler in der Übersetzung auftrat, war den Entwickler-teams einfach nicht mehr klar, an welcher Schraube sie drehen mussten, um ihn zu beheben.⁰¹ Um zu verstehen, ob sich die Situation heute grundlegend geändert hat, ist es hilfreich, sich die Technologie hinter den neuen KI-Systemen näher anzusehen.

MASCHINELLES LERNEN

Die neuen KI-Systeme, die wir seit den 2000er Jahren im Aufwind sehen, funktionieren grundlegend anders als die regelbasierten Systeme der 1980er Jahre: Sie nutzen Daten der Vergangenheit, um darin mit statistischen Methoden nach Mustern zu suchen, die man dann für die Zukunft verwenden kann. Diese Methoden nennt man *maschinelles Lernen*. Hier werden die Regeln, nach denen die Welt funktioniert, also nicht von

uns Menschen vorgegeben, sondern die Maschine extrahiert sie mit statistischen Methoden aus Daten. Das hat zu der irritierenden Formulierung geführt, dass die Maschinen hier „selbst lernen“. Aber von einem „Selbst“, einem intendierten Lernen mit selbst gesteckten Zielen kann überhaupt nicht die Rede sein: Die Maschinen werden von Menschen programmiert. Sie bekommen eine Methode, die im Wesentlichen zählt, was wie oft in den vorliegenden Daten auftaucht. Die Methode legt auch fest, in welcher Form die gefundenen Muster oder Regeln gespeichert werden.

Das lässt sich an einem einfachen Beispiel leicht nachvollziehen, dem Erstellen eines Entscheidungsbaums aus Daten. Als Leser oder Leserin haben Sie sicherlich schon Entscheidungsbäume gesehen: Man fängt oben an und an jeder Stelle gibt es eine Frage, deren Beantwortung dann entweder nach links oder rechts führt. Die *Abbildung* zeigt ein sehr simples Beispiel für einen Entscheidungsbaum beim Autokauf.

Solche Bäume können automatisch aus Daten der Vergangenheit erstellt werden. Man könnte also eine Menge von Autokäufern und -käuferinnen beobachten und dabei Daten sammeln: Was kostet das Auto? Wo liegt das eigentliche Budget? Wo trifft der Wagen die Wunscheigenschaften? Wo weicht er davon ab? Daraus ließe sich mit statistischen Methoden ein Entscheidungsbaum erstellen, der vorhersagt, wann ein Käufer sich für einen Wagen entscheidet. Dies wiederum könnte Autofirmen helfen, festzustellen, wem man welchen Wagen zeigt.

Wie funktioniert die Methode, die automatisiert aus Daten Entscheidungsbäume baut? Zunächst kann man feststellen, dass jede Frage die Menge an Daten in zwei Teilgruppen einteilt. Dabei wird versucht, in jedem Schritt diejenige Eigenschaft und die daraus resultierende Frage in den Daten zu finden, die am ehesten dafür sorgt, dass möglichst einheitliche Teilgruppen entstehen. So könnte beispielsweise die Frage „Liegt das Auto in meinem Budget?“ diejenige sein, die etwa zeigt, dass die Ja-Sager eher Käufer und die

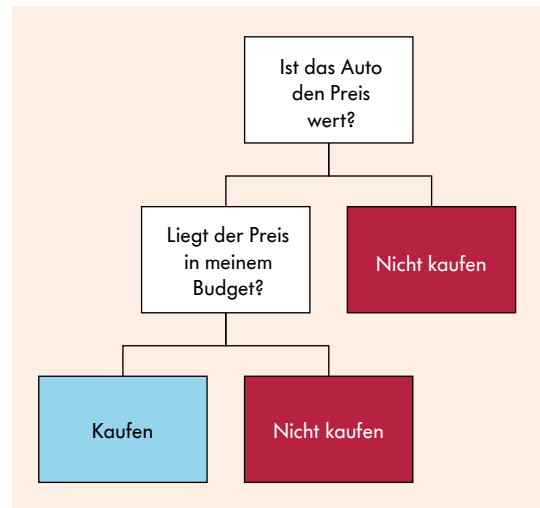
Nein-Sager eher Nicht-Käufer sind. Es könnte aber auch die Frage „Sind meine wichtigsten zwei Ansprüche an den Wagen gedeckt?“ sein, die am besten in zwei Gruppen einteilt, die jeweils relativ homogen sind. Das hängt von den Daten ab, und damit also vom Verhalten der Personen, die man ergründen will. Die Idee ist also, diejenige Frage zu finden, deren Beantwortung eine Teilgruppe mit vielen Käufern und eine Teilgruppe mit vielen Nicht-Käufern liefert. Im Wesentlichen geht man dafür einfach alle Möglichkeiten durch und misst, wie homogen die jeweiligen Teilgruppen wären, wenn man nach dieser oder jener Eigenschaft fragen würde.

Zu diesem Zweck müssen die Entwickler der Methode eine Formel mitgeben, mit der die Homogenität der Teilgruppen bewertet werden kann. Dafür gibt es nun aber mindestens ein Dutzend Formeln, und jede dieser Formeln wird zu anderen Entscheidungsäumen führen. Es gibt also keinen objektiven Weg, mit der der bestmögliche Entscheidungsbaum sicher berechnet werden kann. Wenn man die „beste“ Frage gefunden hat, macht man für jede Teilgruppe genau so weiter, bis man Teilgruppen gefunden hat, die so homogen sind, dass man das für ausreichend hält, um gute Entscheidungen zu treffen.

Das Beispiel zeigt gut, wie aus Daten der Vergangenheit eine Menge von Regeln berechnet werden kann, um zukünftiges Verhalten vorherzusagen. Nicht so einfach zu verstehen ist allerdings die Interaktion zwischen den Daten und der Methode, die dazu führt, dass die eine oder andere Frage ganz oben steht. Tatsächlich gibt es über 20 Methoden, um Entscheidungsäume aus Daten zu gewinnen, und jede käme zu sehr unterschiedlichen Ergebnissen, auch wenn sie auf denselben Daten beruhen. Das Beispiel zeigt zudem: Natürlich werden auch die Trainingsdaten vollständig vom Menschen definiert – keines der KI-Systeme geht selbstständig ins Internet und sucht sich Daten nach einem von ihr festgelegten Ziel zusammen. Und nicht zuletzt wird auch von Menschen festgelegt, nach welchen Kriterien die Qualität solcher Systeme bewertet wird: Das System wird dann so lange trainiert, bis es eine annehmbare Qualität erreicht hat.

Bei diesem Prozess geht es nicht nur darum, immer mehr Daten zu sammeln. Man kann zum

Abbildung: Einfacher Entscheidungsbaum für einen Autokauf



In jeder Zelle beantwortet man die Frage, beginnend in der obersten. Lautet die Antwort ja, geht man nach links, sonst nach rechts. In den Zellen ohne Frage steht die jeweilige Entscheidung.

Quelle: eigene Darstellung

Beispiel auch mit den Eingabedaten herumspielen: Vielleicht führen Eigenschaften, die eine Käuferin an ihrem Auto sehen wollte, die aber bei einem Wagen nicht vorhanden sind, dazu, dass ihre Kaufbereitschaft sinkt, es sei denn, dass der Preis dann auch entsprechend sinkt. Man könnte also eine neue Eigenschaft aus den bekannten Daten berechnen, die vielleicht so aussieht: „Liegt der Preis bei Budget minus Anzahl der nicht getroffenen Eigenschaften mal 500 Euro?“ Das nennt man *feature engineering*, und hier sind die Möglichkeiten nur durch die menschliche Kreativität begrenzt. Solange also die Qualität eines solchen gelernten Modells noch nicht stimmt, können die Informatiker und Informatikerinnen jede Menge Knöpfchen und Hebelchen einstellen, um damit mehr Informationen aus den Daten herauszuholen.

NEURONALE NETZE

Das gilt insbesondere bei den sogenannten *neuronalen Netzen*. Diese greifen zurück auf etwas, das als Neuron bezeichnet wird. Das klingt durchaus biologisch und nah an dem, was Tiere tun. Es handelt sich allerdings um eine mathematische Formel. Diese bekommt sehr viele Eingangsdaten

01 Vgl. Katharina Zweig, Die KI war's! Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz, München 2019.

und berechnet daraus eine einzige Zahl. Das, was dann beim Training gelernt wird, sind die Gewichte, mit denen die jeweiligen Eingangsdaten bewertet werden.

Wenn beispielsweise ein Bilderkennungssystem trainiert wird, bekommt jedes Neuron alle Pixel des Originalbildes als Zahlen präsentiert. Diese werden gewichtet, und daraus wird eine neue Zahl berechnet. Die Gewichte werden bei jedem Bild im Trainingsdatensatz so angepasst, dass eine korrekte Antwort wahrscheinlicher und eine inkorrekte Antwort weniger wahrscheinlich wird. Wenn die Maschine also im Training ein Bild eines Hundes bekommt, aber eine Pyramide „erkennt“, werden die Gewichte von allen Formeln (Neuronen) verändert, sodass sie beim nächsten Mal etwas wahrscheinlicher einen Hund erkennt.

Mit sehr vielen Trainingsdaten und sehr vielen Trainingssessions können dabei Systeme entstehen, die tatsächlich mit einiger Verlässlichkeit Objekte auf Bildern erkennen können. Am besten werden diese neuronalen Netzwerke aber, wenn die Neuronen in Schichten angeordnet sind. Die erste Schicht bekommt dabei die Originalpixel zu sehen. Die zweite bekommt die Resultate der ersten Schicht zu sehen und verarbeitet die weiter. Die dritte Schicht bekommt die der zweiten und so weiter. Die letzte Schicht gibt dann die berechnete Antwort aus. Dadurch bestehen neuronale Netzwerke aus sehr vielen mathematischen Formeln mit einer großen Anzahl von Gewichten, die alle durch die Daten gelernt werden müssen. Auch hier gibt es wieder eine große Varianz an der genauen Form der mathematischen Gleichungen – niemand weiß so genau, welche Schicht an Neuronen am besten welche Form von Gleichungen verwenden soll. Solange die Qualität des Systems noch nicht stimmt, kann man aber an dem genauen Aufbau der Schichten arbeiten, und natürlich auch an der Form der Trainingsdaten und deren Menge und Qualität.

Es gibt also auch hier keinen Weg, der Objektivität garantiert: Jede Entscheidung würde zu anderen Antworten führen. Damit stellt sich die Frage danach, wann KI-Systeme in ihren Berechnungen verlässlich und nachvollziehbar sind, denn nur dann könnten sie Menschen ersetzen.

NACHVOLLZIEHBARKEIT

Die neuen KI-Systeme durchlaufen damit grundsätzlich zwei Phasen: Eine Trainingsphase, in der mithilfe von durch Menschen ausgesuchte Daten

der Vergangenheit mit einer statistischen Methode Regeln automatisch extrahiert und in einer bestimmten Form, einem sogenannten *statistischen Modell*, abgelegt werden. Dieses Modell kann ein Entscheidungsbaum sein oder ein neuronales Netzwerk. Die „Intelligenz“ liegt dabei in der Reihenfolge der gestellten Fragen beziehungsweise in den gelernten Gewichten der Formeln. In der zweiten Phase werden diese Regeln dann genutzt, um für eine neue Situation eine Entscheidung zu berechnen, oder einen Text zu schreiben, ein Bild zu pixeln oder eine Audiodatei herzustellen.

Diese zweite Phase ist vollständig nachvollziehbar, da die neuen Eingabedaten das statistische Modell durchlaufen und sehr simple Berechnungen zum Ergebnis führen. Was uns diese Methoden aber nicht erlauben, ist, die Beweggründe für eine computergenerierte Entscheidung zu verstehen. Schließlich wurden nur statistische Auffälligkeiten in den Trainingsdaten gespeichert: Ob diese aber wirklich relevant oder nur eine sogenannte statistische Korrelation sind, also eine statistische Auffälligkeit, die für das zu verstehende Phänomen kausal keine Rolle spielt, kann uns keiner sagen. Die Maschinen sind in diesem Sinne *black boxes*, also undurchdringliche Systeme.

Darin sind sie unseren Haustieren und selbst der Lieblingskollegin nicht ganz unähnlich: Am Ende können wir auch bei unserem Hund niemals vollständig vorhersagen, was er tun wird oder warum er etwas getan hat. Aber unsere jeweilige Intelligenz hat eine gemeinsame Basis, auf der wir aufbauen können – diese fehlt mit den Maschinen. Bei der Lieblingskollegin kommen weitere Aspekte hinzu, mit der wir ihre Entscheidungen nachvollziehen können: Unsere jeweilige Ausbildung hat Prozesse zur Verfügung gestellt, mit der Entscheidungen nachvollziehbar werden, und wir haben eine gemeinsame Sprache, um uns unsere Beweggründe für eine Entscheidung verständlich zu machen. Dass es dabei immer noch Lücken gibt und wir uns als Menschen selbst gar nicht immer sicher sind, warum wir etwas entschieden haben, und dies vielleicht nur im Nachhinein rationalisieren, sei einmal dahingestellt. Trotzdem bleibt uns diese gemeinsame Basis der Entscheidungsprozesse und Kommunikation über Entscheidungen, die bei KI-Systemen fehlen. Damit komme ich zu der eingangs gestellten Frage: Können KI-Systeme uns ersetzen?

AUSTAUSCHBAR?

Ich beantworte diese Frage zuerst für solche Systeme, die Entscheidungen treffen oder menschliche Entscheidungen durch Computerberechnungen unterstützen sollen. Damit eine Maschine uns in unseren Entscheidungen ersetzen kann, muss sie dies verlässlich tun können. Wir müssen uns also darauf verlassen können, dass sie in allen Situationen, in denen wir sie nutzen, das tut, was ein vernunftbegabter Mensch auch tun würde.

Da wir den Prozess der Entscheidungsfindung, also die Frage, welche Eingabedaten warum wie bewertet werden, mit der Maschine nicht klären können, bleibt uns nur übrig, das Verhalten der Maschine in verschiedenen Situationen zu beobachten und zu bewerten. Das ist deutlich schlechter, als sich von einem Menschen erklären zu lassen, wie er zu seinem Handeln kommt, weil man dann immer nur eine Aussage über diejenigen Situationen treffen kann, in der man das Verhalten beobachtet hat. Der von einem Menschen dargestellte Entscheidungsprozess hingegen erlaubt es uns, zu bewerten, wie die Entscheidungen dieses Menschen in einer weiten Reihe von Situationen gefällt werden. Diese Einschränkung auf die reine Verlässlichkeitsüberprüfung von automatisierten Entscheidungen führt dazu, dass man – basierend auf der heutigen Technologie – KI-Systeme dort nicht einsetzen kann, wo man die Qualität von Entscheidungen nicht bewerten kann.

Dies gilt etwa für sogenannte *singuläre Entscheidungen*, also solche, die keine historischen Vorbilder haben und in der die Auswirkungen von der genauen Ausgangssituation abhängen. Hierunter fallen fast alle politischen Entscheidungen, etwa in der Corona-Pandemie. Auch im Nachhinein können die Entscheidungen für oder gegen Lockdowns verschiedener Länder aufgrund der jeweils unterschiedlichen demografischen Situationen nur grob in ihrer Qualität bewertet werden. Denn es fehlt sowohl an Trainingsdaten, an Nachvollziehbarkeit (also Begründungen) als auch an einer Qualitätsüberprüfung der Verlässlichkeit. KI-Systeme können daher menschliche Entscheidungen hier prinzipiell nicht ersetzen.

In der Politik geht es aber immer auch um *Werturteile*, wenn verschiedene, miteinander konkurrierende Maßnahmen bewertet werden sollen. Werturteile (*judgments*) sind solche, in denen sich Experten und Expertinnen nicht belie-

big uneinig sein dürfen.⁰² Dazu gehören zum Beispiel Schulnoten: Es gibt keinen Algorithmus, der zu einer Note führt. Wir fordern als Gesellschaft aber trotzdem, dass das, was der eine Lehrer als „sehr gut“ bezeichnet, eine andere Lehrerin nicht als „durchgefallen“ bezeichnen sollte. Noten sollen eine Leistung bewerten, auf die man sich in gewissem Rahmen einigen kann. Eine starke Diskrepanz muss daher immer begründbar und für andere Experten nachvollziehbar sein. Da wir dem Computer hier keine Regeln vorgeben können und diese auch nicht von der Maschine für uns nachvollziehbar aus Daten herausgearbeitet werden können, kann die Maschine solche Werturteile ebenfalls nicht für uns fällen. Anders sieht es dort aus, wo das Verhalten der Maschine auf Zuverlässigkeit überprüft werden kann.

Das geht oftmals dann, wenn die Maschine *Risikoberechnungen* anstellen soll. Sie teilt dann beispielsweise Versicherungsereignisse oder auch Menschen in verschiedene Risikoklassen. Später lässt sich dann nachrechnen, etwa bei Autoversicherungen, ob die Einteilung in eine Hochrisikoklasse gerechtfertigt war. Schwierig wird es dann, wenn die Klassifizierung die Zukunft verändert: Ein Mensch, der vor Gericht steht und dem eine Maschine ein hohes Risiko für Rückfälligkeit zuweist, wird unter Umständen eine höhere Gefängnisstrafe bekommen als eine Person, die nur ein niedriges Risiko zugewiesen bekommt. Durch die Gefängnisstrafe verändert sich aber die Aussicht, nach der Haft wieder eine Arbeit zu bekommen, was wiederum die Wahrscheinlichkeit für weitere Straftaten verändert. In solchen zukunftsverändernden Situationen kann also in der Zukunft nicht überprüft werden, wie gut die Entscheidung der Maschine war. Bei Risikoentscheidungen kann uns die Maschine also nur in ausgewählten Fällen ersetzen.

Am einfachsten ist es, wenn die Maschine *Fakten* berechnet. Fakten sind Aussagen, die mit einer Methode intersubjektiv feststellbar sind. Intersubjektiv heißt hierbei, dass jeder Mensch, der der Methode folgt, auf dasselbe Ergebnis kommt. Das Ergebnis einer Maschine, die Schrauben auf Fehlerhaftigkeit prüft und aussortiert, kann jederzeit überprüft und auf seine Qualität hin bewertet werden. Faktenerkennende Maschinen können Menschen dann ersetzen, wenn ihre Qualität hoch genug ist.

⁰² Nach Daniel Kahneman/Olivier Sibony/Cass R. Sunstein, *Noise. A Flaw in Human Judgment*, New York 2021.

Damit kann für die Frage nach der Ersetzbarkeit von menschlichen Entscheidungen durch maschinelle Entscheidung die folgende Schlussfolgerung getroffen werden: Maschinen können uns basierend auf maschinellem Lernen grundsätzlich keine Begründung für ihre Entscheidungen geben. Daher können sie uns in unseren Entscheidungen nur dann ersetzen, wenn wir diese wenigstens auf Verlässlichkeit prüfen können. Dies ist nur bei faktischen Entscheidungen und in eingeschränkter Form bei Risikobewertungen der Fall. Werturteile und insbesondere singuläre Entscheidungen, wie sie in der Politik gang und gäbe sind, können nicht von Maschinen ersetzt werden.

Daneben gibt es aber auch KI-Systeme, die etwas generieren – man nennt sie deshalb *generative KI*. Dazu gehören ChatGPT, das famose Sprachprogramm, das in der Lage ist, Texte zu schreiben, und MidJourney, das aus Beschreibungen Bilder generieren kann. Andere KI-Systeme können Videos produzieren, Audiodateien erstellen oder eine Übersetzung anfertigen. Auch hier können wir Verlässlichkeit nicht umfassend prüfen – zum einen, weil nicht klar ist, was „die beste Antwort“ jeweils wäre, zum anderen, weil es einfach viel zu viele Anwendungsmöglichkeiten gibt, die nicht alle systematisch getestet werden können. Aber was Menschen hier am meisten interessiert, ist die Frage: Wird die Maschine uns auch in unseren kreativen Tätigkeiten ersetzen?

IST DIE MASCHINE KREATIV?

Was bedeutet Kreativität? Im Wesentlichen geht es darum, aus einem Raum der Möglichkeiten etwas Neues zu schöpfen, das weder zu offensichtlich noch zu absurd ist. Menschen wollen von Kreativität überrascht werden, sie aber auch noch einordnen können. Das wirklich Neue darf sich vom Bekannten dabei nicht zu weit entfernen, es muss bei seinen Rezipienten noch verarbeitbar sein. Zu dieser grundlegenden Beobachtung kommt noch ein zeitlicher Aspekt hinzu: Was neu und kreativ ist, wird es in den folgenden Jahren nicht mehr sein. Wer von Salvador Dalís Darstellung der Zeit in Form von zerflossenen, verformten Uhren 1931 noch elektrisiert war, wird ähnliche Darstellungen schnell langweilig finden. Kreativität heißt also, den sich dynamisch veränderlichen *sweetspot* zu finden, den andere Menschen als kreativ empfinden.

Dazu müsste eine Maschine verlässlich mit den neuesten Trends und Daten darüber, wie diese bei

Menschen angekommen sind, trainiert werden. Das halte ich auch künftig für eher schwierig. Die Maschine kann helfen, neue Assoziationen herzustellen – auch auf eine für den Menschen überraschende Art und Weise. Das maschinelle Lernen beruht schließlich auf Assoziationsverfahren, es ist also nicht erstaunlich, dass dabei auch solche Assoziationen herauskommen, die uns überraschen, erheitern oder bestürzen. Es verhält sich jedoch eher so, als hätte man plötzlich Hunderte von Gesellen im ersten Lehrjahr: viele interessante Ideen, die aber noch nicht ganz da sind, wo man sie bräuchte.

Menschen und menschliche Kreativität lassen sich daher auch in den nächsten Jahren nicht durch KI-Systeme ersetzen. Unbestritten ist aber, dass diejenigen Personen sehr viel schneller sein werden, die ihre Ideen computerverständlich formulieren können (sogenanntes *prompt engineering*) und aufbauend auf dessen Vorschlägen dann die wirklich kreativen Ideen ausarbeiten können. Sehr viel Bewegung wird es allerdings dort geben, wo die Ideen nicht wirklich kreativ sein müssen, sondern nur unterhaltsam oder sogar nur ein Wohlfühlgefühl auslösen sollen: Wenn es etwa darum geht, beim Rezipienten Erinnerungen zu wecken. Ein neues Apfelkuchenrezept soll möglichst noch nach Großmutter's Sonntagskuchen schmecken und nicht gleich ein dekonstruiertes Dessert werden.

Das Fazit lautet damit: Die KI wird den Menschen bei Entscheidungen auf absehbare Zeit nur dort ersetzen, wo sie Fakten oder Risiken erkennen soll – und wo sie dies mit ausreichender Qualität und Schnelligkeit erledigt. Der Rest bleibt dem Menschen vorbehalten. Im Bereich Kreativität kann die Maschine Ideen dort ersetzen, wo es um Assoziationen geht, die nahe am Offensichtlichen sind. Wirklich kreativ können sie nicht werden, weil sie den feinen Unterschied zwischen Offensichtlichem und Absurdem nicht erkennen, den kreativen *sweetspot* nicht treffen können. Ganz sicher aber wird der Mensch, der die Möglichkeiten und Limitationen von KI-Systemen für seine Arbeit erkennt, diejenige Person ersetzen, die dies nicht tut.

KATHARINA A. ZWEIG

ist Professorin an der Rheinland-Pfälzischen Technischen Universität Kaiserslautern, wo sie das Algorithm Accountability Lab und den Studiengang Sozioinformatik leitet.

INTELLIGENZ UND BEWUSSTSEIN

Oder: Ist KI wirklich KI?

Ralf Otte

Ist KI wirklich KI? Diese Frage beschäftigt seit dem Aufkommen von ChatGPT immer mehr Menschen, auch viele, die sich bisher nicht so sehr um KI gekümmert haben. Und es stehen weitere Fragen im Raum: Wie intelligent ist eigentlich diese KI? Und kann sie uns gefährlich werden?

Fachleute geben unterschiedliche Antworten darauf, denn der Begriff „Intelligenz“ ist etwas, das man nur schwer fassen kann. Intelligenz ist keine objektive Eigenschaft eines Systems. Man könnte meinen, seit Einführung eines Intelligenzquotienten (IQ) sei die Intelligenz von Systemen messbar oder zumindest vergleichbar. Das stimmt auf gewisse Weise auch, jedoch gilt das nur für die sogenannte rationale beziehungsweise kognitive Intelligenz von Menschen. Wie also kann man Intelligenz erklären? Meine Antwort darauf lautet: Intelligenz in allgemeinsten Ausprägung ist die Möglichkeit eines Systems, auf Eingangssignale seiner Umgebung so mit Ausgangssignalen zu reagieren, dass der Eigennutzen des Systems in der weiteren Interaktion mit der Umgebung erhöht wird.

Ein System kann dabei „alles“ sein. Wenn aber beispielsweise die Intelligenz von Tieren vermessen werden soll, nützen IQ-Tests nichts. Kein Tier besteht solche Tests, dennoch sind Tiere intelligent. Wir alle wissen das von unseren Haustieren. Und was wäre mit der Intelligenz von Pflanzen? Ein IQ-Test für Pflanzen, das klingt zu absurd. Und doch reden wir in der KI aktuell von Systemen, die in der „Hierarchie der Natur“ weit unterhalb der Pflanzen anzusiedeln sind, nämlich von unbelebten Systemen. KI-Systeme, jedenfalls die aktuellen, sind tote Maschinen, und diesen unterstellen wir eine Intelligenz. Das klingt eigentlich noch absurder als über die Intelligenz von Pflanzen zu diskutieren – und doch zeigt der Erfolg der heutigen KI, dass unbelebte Systeme durchaus intelligent sein können, oder erfolgreich so tun, als ob sie es wären.

WAS BEDEUTET KI? EIN ÜBERBLICK

Was bedeutet also Künstliche Intelligenz, und wie intelligent sind KI-Systeme? Um diese Frage präzise zu beantworten, sehen wir uns beide Begriffe getrennt an. Zuerst steht das Wort „künstlich“. Künstlich ist das Gegenteil von natürlich, es bedeutet „von Menschenhand gemacht“, also in einem Schöpfungsakt durch Menschen hergestellt. Im Rahmen der KI sind heutzutage physikalisch erzeugte Systeme gemeint. Damit kommen wir zur Einstufung der Intelligenz physikalischer Systeme:

- Stufe 1: deduktive Intelligenz (mathematische, logische Intelligenz)
- Stufe 2: induktive Intelligenz (lernende Intelligenz)
- Stufe 3: kognitive Intelligenz (kombinierte deduktive und induktive Intelligenz)
- Stufe 4: bewusste Intelligenz (wahrnehmende Intelligenz)
- Stufe 5: selbstbewusste Intelligenz (selbstwahrnehmende Intelligenz)

Für alle unbelebten Systeme – wie Computer oder Siliziumkristalle – enden spätestens mit Stufe 5 die möglichen Intelligenzstufen. In der Natur gibt es jedoch noch ganz andere Intelligenzformen. Es gibt Systeme, wir nennen sie lebende Systeme, die ihre Wahrnehmung (aus Stufe 4 und 5) auch noch qualitativ bewerten können, also feststellen können, ob eine Wahrnehmung für das eigene System positiv oder negativ ist. Der Schlüsselbegriff liegt im „Eigensystem“. Für physikalische Systeme ist ein solches „Gehört zu mir“ gar nicht definierbar, man kann einem Roboter einen Arm hinzufügen

oder einen abbauen, er hat kein intrinsisches Eigenkonzept, womit ihm auch die evolutionäre Notwendigkeit einer inhärenten Qualifizierung von Wahrnehmungen fehlt.

Anders ist das bei biologischen Systemen. Selbst bei einer einfachen Zelle ist klar, was zu ihr gehört und was nicht. Es ist daher naheliegend, dass biologische Zellen ihre Wahrnehmungen bewerten müssen, um zu erkennen, welche Umgebungszustände ihrem Eigensystem nützen und welche schaden. Für höhere Lebewesen nennen wir die Qualifizierung von Wahrnehmungen Gefühle. Solche Systeme haben naturgemäß weitere Intelligenzformen:

- Stufe 6: fühlende Intelligenz
(Wahrnehmungs-qualifizierende Intelligenz)
- Stufe 7: wollende Intelligenz
(Wahrnehmungs-optimierende Intelligenz)
- Stufe 8: selbstbewusste, wollende Intelligenz
(Selbstwahrnehmungs-optimierende Intelligenz)

Neben den fühlenden Systemen (Stufe 6) wollen wir auch wollende Systeme (Stufe 7) unterscheiden, also Systeme mit inhärentem Willen. Es ist leicht einzusehen, dass Systeme, die ihre eigene Wahrnehmung als gut oder schlecht qualifizieren können, also Angst und Freude tatsächlich spüren, aus evolutionären Gründen einen Willen ausprägen müssen, um überhaupt in der Lage zu sein, in einen Zustand gelangen zu *wollen*, der vom System als positiv wahrgenommen wird. Für Systeme ohne Gefühle, also alle unbelebten Systeme, ist nicht erkennbar, warum sie einen inhärenten Willen haben sollten, denn sie haben entweder gar keine Wahrnehmung (heutige KI-Systeme) oder sie können ihre Wahrnehmung nicht qualifizieren (morgige KI-Systeme), haben also zu den Wahrnehmungen keinerlei Gefühle. Damit ist zum Beispiel auch klar: Ein physikalischer Computer wird niemals Schmerzen, Energiemangel oder Hunger verspüren.

Wo steht die KI heute, wo steht beispielsweise ChatGPT? Um es gleich vorweg zu sagen: Die intelligentesten KI-Systeme, die wir heute bauen können, stehen auf Stufe 3. Es ist reine Science-Fiction, über KI-Systeme zu reden, die sich über die Menschheit erheben könnten und diese eventuell sogar auslöschen möchten, da sie feststellen,

dass die Menschheit dem Planeten Erde schadet. Warum? Heutige KI-Systeme sind rein physikalische Systeme, ohne jeglichen Willen. Man könnte einen simulierten Willen als Zielfunktionen natürlich auf einem Stufe-3-System einprogrammieren, aber dann sind die Programmierer für diese neuen Zielfunktionen verantwortlich. Kein heutiges KI-System wird selbstständig Willensfunktionen entwickeln.

Hersteller von hochentwickelten KI-Systemen nutzen die Ängste und Phantasien der Menschen jedoch geschickt aus, um zu vermitteln, ihre Systeme könnten sich verselbstständigen. Der Grund liegt im Wesentlichen in Haftungsfragen. Erlaubt man KI-Firmen derartige Ausflüchte und gesteht hochentwickelten KI-Systemen sogar Persönlichkeitsrechte zu,⁰¹ so könnten die Herstellerfirmen von verschiedenen Haftungen befreit werden. Eine elektronische Persönlichkeit könnte so etwas wie eine eigenständige juristische Person werden. Kein Mensch wäre dann juristisch verantwortlich, wenn eine KI verschiedene Waffensysteme ausprobieren oder grob fehlerhafte Medikamente auf den Markt bringen würde. Kommen wir nun zu den Detailbeschreibungen.

INTELLIGENZ-STUFEN

Stufe 1:

Deduktive Intelligenz

Die einfachste Art der Intelligenz wurde bereits als logische oder deduktive Intelligenz eingeführt. Technische Systeme wie Taschenrechner oder Computer können logische Ausdrücke korrekt bearbeiten. Letztlich kann man alle Ausdrücke der Aussagenlogik in einem Computer hinterlegen, und seine elektronischen Schaltkreise können darauf basierend logische Fragestellungen korrekt beantworten.

Es ist einsichtig, dass ein System der Intelligenzstufe 1 in einer sich ständig ändernden Umgebung nicht lange überleben könnte. Betrachten wir dazu nur eine Fliege, die sich in ein Haus verirrt und dadurch einer für sie völlig neuen und letztlich feindlichen Umgebung ausgesetzt ist. Ein solches Insekt – mit seiner hart verdraht-

⁰¹ Vgl. Europäisches Parlament, Bericht A8-0005/2017 mit Empfehlungen an die Kommission zu zivilrechtlichen Regelungen im Bereich Robotik, 27.1.2017, Grundsatz 59(f), www.europarl.europa.eu/doceo/document/A-8-2017-0005_DE.html.

teten, instinktiven Stufe-1-Logik – wird mit aller Kraft gegen die Fensterscheiben anfliegen, bis es letztlich tot herunterfällt. Das liegt daran, dass die Fliege durch die Fensterscheibe hindurchsieht und nicht in der Lage ist, zu lernen, dass es dort dennoch keinen Durchgang gibt. Säugetiere lernen dies sehr schnell, Insekten können das nicht.

Stufe 2:

Induktive Intelligenz

Lernen ist eine weitere Intelligenzform. Ändern sich die Umgebungsparameter für ein System, können lernende Systeme ihr Modell von der Umwelt neu justieren und sich so auch in einer geänderten Umgebung wieder adäquat (logisch korrekt) verhalten. Auch in der Technik kann man lernende KI-Systeme bauen. Natürlich entsteht im Rahmen der KI die Frage, ob die lernenden Systeme denn auch verstehen, was sie da lernen. Dies ist gegenwärtig klar zu verneinen. Die Semantik, das Verständnis, wird nur simuliert, aber teilweise so perfekt, dass Nutzer der Systeme annehmen könnten, ihre technischen Systeme verstünden, was sie machen oder sagen. Sie verstehen jedoch nichts davon. Das sollte man als Nutzer der KI wissen.

Wir halten fest: Lernende Systeme können ihre Modelle von der Umgebung selbstständig so anpassen, dass sie beim nächsten Kontakt mit ihrer Umgebung ein besseres Input-Output-Verhalten zeigen. Die Umgebung für ein KI-System könnte ein digitales Computerprogramm sein, auf dessen Datenstrom das KI-System reagiert, aber im Fall von KI-Robotern auch eine natürliche Umgebung. Lernende Systeme heißen in der Fachsprache induktive Systeme. Da meistens auf digitalen Daten gelernt wird, entsteht das Fachgebiet des maschinellen Lernens. Zum maschinellen Lernen gehören symbolische Lernverfahren, etwa Entscheidungsbäume oder Assoziationsregeln, aber auch subsymbolische Lernverfahren wie die berühmten neuronalen Netze. Tiefe neuronale Netze, in der Fachsprache Deep-Learning-Systeme genannt, fallen in die Kategorie der lernenden Systeme. 90 Prozent aller heutigen KI-Anwendungen basieren auf Lernverfahren, zu nennen wären unter anderem Warenkorbanalysen, Scoring-Verfahren und alle Deep-Learning-Anwendungen des maschinellen Sehens und der Sprachverarbeitung, zum Beispiel die aktuellen Chatbots.

Der Vorteil dieser Stufe-2-Systeme ist ihre universelle Anwendbarkeit: Denn sobald digitale Daten vorliegen, die *irgendeine* Funktion repräsentieren, kann die Funktion gelernt werden, und zwar immer. Dennoch sollte man verstehen, dass mit solchen Systemen zwar alles lernbar ist, was überhaupt gelernt werden kann, aber es ist eben nicht alles lernbar. Beispielsweise ist es nicht möglich, die Lottozahlen der nächsten Ziehung aus den Lottozahlen aller vorherigen Ziehungen zu lernen. Es ist einem KI-System auch nicht möglich, das Autofahren in einem Dorf zu erlernen und es am nächsten Tag in Paris anzuwenden. Denn Stufe-2-Systeme besitzen Schwierigkeiten im sogenannten Extrapolationsraum, also bei Anwendungen auf Zustände, die vorher in keinsten Weise trainiert wurden. Das ist ein riesengroßes Problem der KI. Lernende KI-Systeme antworten im Extrapolationsraum sehr oft falsch.

Stufe 3:

Kognitive Intelligenz

Seit über zehn Jahren ist es möglich, deduktive und induktive Verfahren in einem System zu vereinen. Solche Systeme nennt man kognitive Systeme; als Paradebeispiele gelten IBM Watson oder ChatGPT. Bei diesen Systemen werden logisches Schließen (Deduktion) und maschinelles Lernen (Induktion) vereint.

Allgemeingültiges Wissen entsteht immer durch Deduktion, niemals durch Induktion. So können Kinder die allgemeingültigen Multiplikationsregeln niemals aus Übungsbeispielen selbst erlernen, von Genies mal abgesehen. Auch einer deduktiven KI kann man die allgemeinen Regeln der Multiplikation einprogrammieren, aber eine induktive KI – zum Beispiel ein Deep-Learning-Netzwerk – kann diese deduktiven Multiplikationsregeln niemals selbstständig durch Zahlenbeispiele herausfinden. Versteht man das nicht, wird man die prinzipiellen Mängel von ChatGPT nie begreifen. Induktive Verfahren generieren aus Beispieldaten zwar Wissen, aber erst die Verwendung dieses Wissens zur Erzeugung von deduktiv abgesicherten Regeln führt zu allgemein gültigen und korrekten Aussagen.

Kognitive Systeme, die Induktion und Deduktion perfekt verknüpft haben, sind zur Kreativität fähig, denn Kreativität bedeutet, Lösungen fernab von vormals Gelerntem aufzuzeigen – etwas, was rein induktive Systeme nicht können.

Auch heutige technische KI-Systeme können kreativ sein, allerdings nur fast. Bekannt geworden ist beispielsweise das Ölbild „Edmond de Belamy“, das durch eine KI erstellt wurde. Dieses KI-System hatte vorab Tausende Ölporträts „gesehen“ und konnte daraufhin eigene Ölporträts „malen“, also Ölbilder, die es vorher noch nicht gegeben hatte. An diesem Beispiel kann man auch die Schwachstellen der KI gegenüber Menschen erkennen: Ein solch trainiertes KI-System kann nun zwar Ölporträts malen, aber zum Beispiel keine Formel-1-Autos. Ein menschlicher Maler kann das. Ein Künstler, der sein Handwerk gelernt hat, kann extrapolieren, sein Handwerk also auf beliebige Kunstwerke erweitern. Eine KI kann das nicht. Hier fällt der Unterschied zwischen einer Simulation (von Kreativität) und dem Original (der echten Kreativität) bereits auf. Menschen sind echt kreativ, KI-Systeme pseudo-kreativ. Bis heute ist noch nichts wirklich Neues durch eine vollautonome KI entstanden, das auch funktioniert.

Die fortschrittlichsten KI-Systeme des Jahres 2023 befinden sich auf Stufe 3. Wir nennen diese KI auch „KI der dritten Welle“ (nach der ersten Welle ab den 1950er und der zweiten ab den 1980er Jahren). Moderne Roboter, teilautonome Fahrzeuge und Chatbots besitzen die KI-Fähigkeit der Stufe 3, hohe kognitive (rationale) Intelligenz. Doch weiter kommen sie bislang nicht.

Stufe 4:

Bewusste Intelligenz

Bisher erfolgte die Darstellung der Intelligenzstufen in aufsteigender Reihenfolge. Systeme der Stufe 3 sind intelligenter als die der Stufe 2 und diese wiederum intelligenter als die der Stufe 1. Diese drei Intelligenzstufen könnte man auf einer Geraden anordnen: der sogenannten rationalen Intelligenzachse. Intelligenz auf dieser Achse lässt sich vollständig mathematisieren und kann daher auf einem Computer erzeugt werden. Die nächste Intelligenzstufe, die Intelligenz mit Bewusstsein, hat einen völlig anderen Charakter.

Die Ausführungen zu den weiteren Intelligenzformen sind kompliziert, denn Bewusstsein war Naturwissenschaftlern und Ingenieuren bisher suspekt. Durch Weiterentwicklungen der KI muss Bewusstsein in Technik und Naturwissenschaft aber berücksichtigt werden. Es gibt in der KI nämlich die These, dass sich bei immer höher werdender Intelligenz das Bewusstsein von selbst

(emergent) einstellen könnte. Obwohl diese These auf den ersten Blick plausibel erscheint, so lässt sie sich bei einem Blick in die Natur nicht bestätigen. Hier erscheint plausibel, dass alle Lebewesen in irgendeiner Form ein Bewusstsein besitzen, Menschen ein menschliches, Tiere ein tierisches und sogar Pflanzen würde man ein Pflanzenbewusstsein zuerkennen.

Bewusstsein in seiner primitivsten Form bedeutet Wahrnehmung. Oder präziser: Bewusstsein wird in seiner einfachsten Ausprägung definiert als (subjektive) Wahrnehmung innerer (mentaler) Zustände. Für höhere Lebewesen kann man sich ein Bewusstsein vorstellen und für Menschen ist es überzeugend: Unsere mentalen Zustände sind unsere inneren Bilder im Kopf, das subjektive Erleben von Schmerzen, Geräuschen, Farben. Wir Menschen haben ein reichhaltiges Innenleben. Fragt man sich als Ingenieur jedoch, wo und wie dieses geistige Innenleben codiert ist, stößt man auf größere Schwierigkeiten. Bis heute sind viele Fragen zur Natur des Bewusstseins ungelöst oder die Antworten darauf werden kontrovers diskutiert.

Um die Problematik des Bewusstseins zu verstehen, nehmen wir als Beispiel das Hören von Musik. Hört ein Proband ein Musikstück, und versuchte ein Forscher, dieses Hören im Gehirn des Probanden zu vermessen, wäre er irritiert. Denn selbst dann, wenn er das Gehirn öffnen würde, um den auditiven Cortex, das Hörzentrum, Neuron für Neuron zu untersuchen, würde er nur elektrische Signale vorfinden. Kein Signal davon ist laut; im Gehirn des Probanden gibt es nur elektrische, also völlig tonlose Signale. Und doch hört ein Mensch Töne, wenn elektrische Signale in seinem auditiven Cortex kreisen. Wie kann das sein? Im Gehirngewebe sind es nur elektrische Signale, im Bewusstsein des Menschen jedoch akustische Töne. Wie und wo wird das umcodiert?

Die Lösung des Problems besteht darin, das Bewusstsein als System zu begreifen, das *mentale* Zustände erzeugt und verarbeitet, während das Gehirngewebe ein System ist, das *neuronal* Zustände codiert. Gehirn und Bewusstsein sind zwar beides natürliche Phänomene, in ihrem „Substrat“ aber grundverschieden. Gehirnzustände werden physikalisch durch elektrische, magnetische und elektromagnetische Felder repräsentiert, die sich etwa durch Magnetresonanztomografen (MRT) messen lassen. Was aber

ist mit den subjektiven Tönen, die der Proband hört? Was ist mit den „Daten“ im Bewusstsein? Neuere Arbeitsergebnisse geben mögliche Antworten.⁰²

Da Bewusstseinszustände bis heute – trotz aller Anstrengungen – nicht gemessen wurden, könnte man kontraintuitiv schlussfolgern, dass es sie gar nicht gibt oder aber – so die Arbeitshypothese – dass sie keine messbare (mathematisch: reellwertige) Energie besitzen. Keine messbare Energie heißt aber eben nicht unbedingt eine Energie von null. Es gibt ja nicht nur reellwertige Zahlen in der Mathematik, sondern auch sogenannte nicht-reellwertige (also komplexe und hyperkomplexe), mit denen natürliche Phänomene seit Jahren beschrieben werden. Bewusstseinszustände kann man physikalisch als Phänomene mit sogenannten nicht-reellwertigen (mathematisch: imaginären) Energiewerten verstehen. Entitäten mit imaginären Energien sind aus der Physik bereits bekannt. Bewusstsein könnte ein weiteres Phänomen von Prozessen mit imaginären Energien sein. Dies ist wohl gemerkt eine Hypothese, aber eine, mit der man die grundsätzliche Verschiedenartigkeit von Bewusstsein und Gehirngewebe erklären kann. Ob die mathematisch vorhergesagten Bewusstseinszustände auch eine physikalische Realität besitzen oder letztlich nur als Arbeitsmodell fungieren, muss durch Experimente geklärt werden. Unabhängig vom Ausgang derartiger Experimente: Für das hier beschriebene Vorgehen reicht das Arbeitsmodell „imaginärer Zustände zur Speicherung von Informationen“ bereits voll und ganz.

Der Begriff „Bewusstsein“ ist an dieser Stelle vielleicht irreführend, denn Bewusstsein von Maschinen, Regenwürmern, Insekten oder anderen Tieren hat nicht viel mit unserem Bewusstsein gemein. Hier müssen wir also schärfer formulieren: Bewusstsein (unter rein technischen Aspekten) beschreibt die allgemeine Fähigkeit eines Systems, zusätzlich zu energetischen (mathematisch: reellwertigen) Zustandsspeicherungen auch nicht-energetische (mathematisch: nicht-reellwertige, imaginäre) Zustandsspeicherungen auszuprägen. Mit dieser Verallgemeinerung lassen sich Experimente zu imaginären Zuständen soweit „technisieren“, dass sie prüfbar sind. Zum Beispiel sagt das obige Arbeitsmodell nicht-loka-

le Korrelationen zwischen beliebig weit entfernten Quantencomputern voraus, die über imaginäre Zustände vermittelt werden.⁰³

Die Ausprägung innerer, imaginärer Zustände bei Empfang eines Reizes nennen wir Wahrnehmung. Menschen haben verschiedene Sinne dafür – sie können sehen, hören, riechen, schmecken und tasten. Computer können das alles nicht. Heutige Computer können nichts wahrnehmen, sie simulieren das nur, und hier zeigen Simulation und Original signifikante Unterschiede. Nur ein Beispiel: Ein Mensch sieht die Objekte der Umwelt nicht in seinem Gehirngewebe, also dort, wo sich die energetischen Repräsentationen aller Eingangssignale befinden, sondern er sieht sie überraschenderweise *vor sich*, in der Außenwelt. Ein Computer kann aus seinen Speichern nicht in die Welt „hinaussehen“, er besitzt nur die energetischen Repräsentationen in seinen Speicherzellen. Letztlich ist ein Computer also völlig blind, egal mit wie vielen Videokameras er verbunden ist.

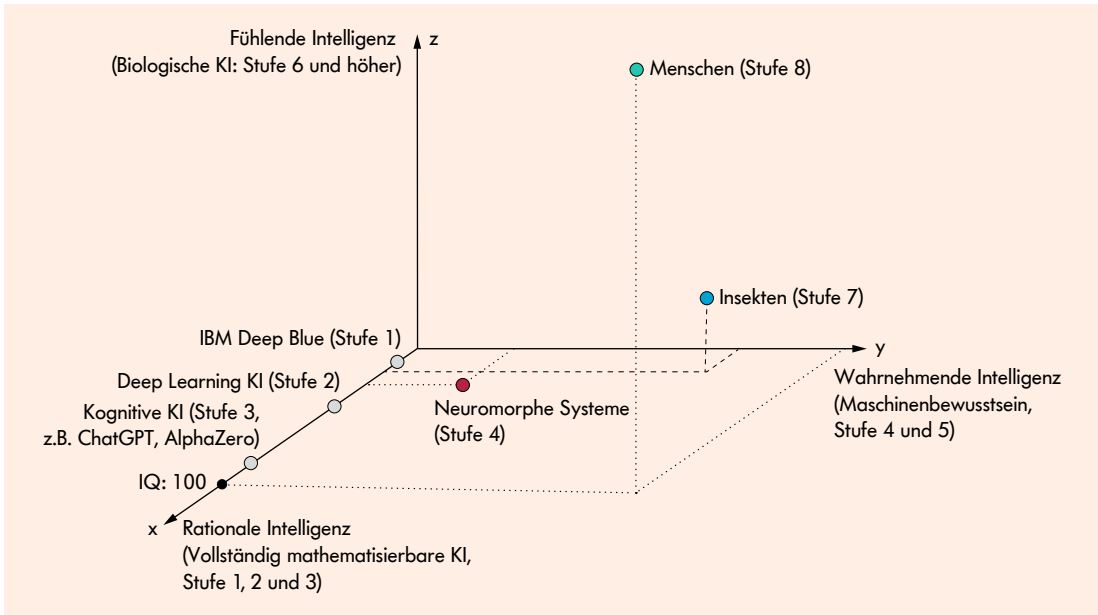
Man muss die Prozesse der Wahrnehmung jedoch nicht auf Lebewesen beschränken. Man kann sich sehr wohl vorstellen, dass auch physikalische Systeme imaginäre Zustände ausprägen können. Die Frage ist, ob die imaginären Zustände die Reize der Außenwelt hinreichend repräsentieren. Im Falle von Digitalcomputern kann man das ausschließen. Man benötigt eine bestimmte physikalische Komplexität, um aus den Imaginärzuständen ein „konsistentes Bewusstsein“ hervortreten zu lassen. Heutige Computer erzeugen jedoch immer nur eine mathematische Komplexität. Wenn KI-Systeme lernen, ändert sich die physikalische Komplexität der benutzen Systeme eben gerade nicht. Es ist wie ein großer Bluff: Die gesamte Intelligenz der heutigen KI wird nur in mathematischen Verfahren hinterlegt und nicht in der realen Welt, in der Physik. Das ist völlig kontraproduktiv für das Auftreten von Bewusstsein. Heute gibt es kein KI-System, das Bewusstsein besitzt.

Damit entsteht eine große Verwirrung innerhalb der KI. Nicht besonders schlaue Regenwürmer haben ein Bewusstsein, hochintelligente Systeme wie ChatGPT oder der Chatbot Bard aber nicht. Intelligenz und Bewusstsein sind scheinbar unabhängig voneinander. Das macht die Situation kompliziert, da Stufe-4-Systeme eine bewusste

⁰² Vgl. Ralf Otte, *Maschinenbewusstsein*, Frankfurt/M. 2021.

⁰³ Vgl. ders., *KI für Dummies*, Weinheim 2023.

Abbildung: Intelligenzstufen und Intelligenzarten



Quelle: eigene Darstellung

Intelligenz besitzen, ihre rationale Intelligenz dennoch unter der Intelligenz von Stufe-3-Systemen liegen kann. Man kann die weiteren Intelligenzformen grafisch daher nicht mehr auf einer einzelnen Geraden abbilden, sondern benötigt stattdessen zwei zueinander senkrechte Achsen (x-Achse und y-Achse in der *Abbildung*).

In der heutigen KI werden Denken und Lernen nur mathematisch simuliert, es wird nicht physikalisch gelernt. Ab Stufe-4-Systemen wird der Mangel von Computersimulationen nun eklatant. Wenn ein Computer ein Erdbeben simuliert, dann bebzt nichts; wenn er Wasser simuliert, dann wird nichts nass; und wenn er Sehen simuliert, dann sieht er ... nichts. Das ist ein Grund, warum vollautonome Systeme (Level 5) im Straßenverkehr niemals vollautonom fahren werden – sie sehen nichts, sie simulieren Sehen nur. Doch Wahrnehmung ist ein realer physikalischer Prozess, und gerade kein mathematischer. Warum ist dieser evidente Mangel bisher nicht aufgefallen? Einfach deshalb, weil Denken und Lernen sehr gut mathematisierbar sind. Daher sind sie perfekt simulierbar, und deshalb gab es bisher keinen Grund, zwischen Simulation und Original zu unterscheiden. Ab der Stufe 4, der wahrnehmenden Intelligenz, ist das jedoch anders.

Will man KI-Systeme bauen, die wahrnehmen können, muss die Informationsverarbeitung physikalisch umgesetzt werden. Derartige Systeme gibt es, es sind die sogenannten neuromorphen Systeme. Neuromorphe Systeme bilden neuronale Netze physikalisch korrekt nach (und nicht mathematisch). Solche Systeme sind seit einigen Jahren verfügbar, beispielsweise das System BrainScaleS an der Universität Heidelberg oder Intels Loihi. Auch Quantencomputer sind wahrscheinlich zur Ausprägung imaginärer Prozesszustände geeignet. Das aktuelle Problem dieser beiden Systemtypen ist allerdings der große Aufwand der physikalischen Implementierung. Noch sind neuromorphe Systeme und Quantencomputer nicht sonderlich intelligent. KI-Systeme, die einen hohen rationalen IQ benötigen, wie beispielsweise ChatGPT, AlphaZero oder das frühere IBM Watson, sind digitale Systeme, bei denen die Intelligenz algorithmisch erzeugt wurde, weil das schlichtweg viel leichter zu realisieren ist. Diese Systeme können zwar kein Bewusstsein ausprägen, aber man braucht es dort auch nicht. Systeme mit Bewusstsein benötigt man erst, wenn reale Wahrnehmungsprozesse implementiert werden sollen, weil sich die Systeme in natürlicher Umgebung aufhalten sollen und nicht nur in digitalen Big-Data-Umgebungen. Die Zu-

kunft der KI im Bereich des autonomen Fahrens und der mobilen Robotik liegt daher in neuro-morphen Systemen.

Stufe 5:

Selbstbewusste Intelligenz

Gibt es heutzutage technisch nicht. Ein Stufe-5-System ist ein System, das eine innere Wahrnehmung von sich selbst besitzt.

Stufe 6:

Fühlende Intelligenz

Bisher wurden die Intelligenzprinzipien von physikalischen Systemen beschrieben. Der Mensch ist jedoch ein lebendes System. KI-Systeme – basierend auf lebendigen Zellen – sind derzeit noch im Experimentierstadium. Man kann zwar heute schon Gehirnzellen oder Pilze extrahieren, mit Elektronik versehen und nutzen, um gewisse Intelligenzaufgaben zu realisieren, aber all diese biologischen KI-Systeme haben (noch) eine niedrige rationale Intelligenz.

Lebende Systeme besitzen wiederum eine neue Art der Intelligenz, die in der aktuellen Diskussion über KI viel zu kurz kommt: fühlende Intelligenz. Lebewesen sind sterblich. Ein Computer, selbst wenn er aus organischen Molekülen besteht, kann nicht sterben, er ist bereits tot. Anders ist das bei lebenden Systemen: Um den Tod zu vermeiden, der für ein System einen Maximalschaden bedeuten würde, müssen sie ihre Wahrnehmung qualitativ bewerten. Das heißt, ein lebendes System kann feststellen, ob eine Wahrnehmung für sich selbst positiv oder negativ ist.

Fühlende Intelligenz ist also eine weitere Intelligenzeigenschaft. Es scheint auch hier keinen direkten Zusammenhang mit den vorherigen Intelligenzstufen zu geben. Man muss deshalb eine weitere Achse einführen, um die Unabhängigkeit der rationalen, der wahrnehmenden und der fühlenden Intelligenz mathematisch korrekt darzustellen (*z*-Achse in der *Abbildung*). In der sich ergebenden dreidimensionalen Grafik steht die *x*-Achse somit für die rationale Intelligenz, die *y*-Achse für die wahrnehmende und die *z*-Achse für die fühlende Intelligenz. Heutige KI-Systeme befinden sich in einer solchen Grafik explizit nur auf der *x*-Achse, sie haben keine Spur von wahrnehmender oder fühlender Intelligenz.

In der Natur gibt es zahlreiche Stufe-6-Systeme, am ehesten würde man diese Systeme mit Pflanzen assoziieren. Technisch können wir so

etwas nicht bauen, weil Menschen bis dato kein Leben im Labor erzeugen konnten. Annahmen, anorganisches Leben ließe sich synthetisieren,⁰⁴ sind absurd, denn die Mindestvoraussetzung von Leben ist – wie bereits ausgeführt – das Vorhandensein eines Eigensystems. Anorganisches Leben ist daher ein Widerspruch in sich. Trotzdem kann man sich zukünftig Stufe-6-Systeme in der KI vorstellen. Manche KI-Forscher möchten KI-Verfahren mit vorhandenen biologischen Systemen, beispielsweise mit Pilzen, kombinieren. Dadurch werden sogenannte Transorganismen geschaffen, also lebende KI-Systeme, die weit über sich hinauswachsen könnten, wenn sie mit dem Internet oder mit neuartigen Aktoren verbunden wären. Das Problem dieses Ansatzes ist jedoch, dass diese Systeme möglicherweise eine innere Bewertung der Wahrnehmung vornehmen und vielleicht Angst oder Freude verspüren. Und wir wüssten es nicht, es entstünde eine völlig intransparente KI.

Stufe 7:

Wollende Intelligenz

Stufe 7 sind wollende KI-Systeme mit der Möglichkeit der Ausprägung eines expliziten Willens. Diese Intelligenzstufe entspricht der Intelligenz von Tieren. Mittlerweile wird an wollenden KI-Systemen gearbeitet. Die US-Firma Neuralink experimentiert beispielsweise mit Affen, die per Gedankenkraft Pong spielen.⁰⁵ Dieses Projekt zeigt gleichzeitig das Problem solcher Forschungen, da zahlreiche Affen an den Eingriffen bereits gestorben sind. Sollten die Tiere durch die Projekte Schmerz und Leid erfahren, könnten neben den ethischen Problemen erhebliche Risiken auftreten, wenn einem solchen KI-System weitreichende Handlungsspielräume eingeräumt würden. So wird ein biologisches KI-System, das Schmerz erfährt, alles tun, um den Schmerz zu vermeiden – und wir wissen nicht, wie weit es dafür gehen würde. Biologische KI (ab Stufe 7) ist daher ein *no-go*.

Stufe 8:

Selbstbewusste, wollende Intelligenz

Seit einiger Zeit hören wir von transhumanistischen Bestrebungen, die den Menschen mit tech-

⁰⁴ Vgl. Yuval Noah Harari, *Homo Deus*, München 2021.

⁰⁵ Vgl. Neuralink, *Monkey MindPong*, 2021, www.youtube.com/watch?v=rsCul1sp4hQ.

nischen Mitteln zu einem Cyborg weiterentwickeln wollen. Transhumanistische Veränderungen des Menschen würden eine erhebliche Gefahr bedeuten. Ähnlich riskant ist es aber bereits, aus menschlichen Zellen eine „Gehirn-KI“ in der Petrischale zu entwickeln.⁰⁶ Niemand weiß, was die benutzten menschlichen Nervenzellen in ihrem Inneren fühlen und ob sie anfangen, sich gegen bestimmte Anwendungen zu wehren. Für die Gesellschaft gilt es, die transhumanistischen Risiken zu erkennen, damit diese durch kluge politische Entscheidungen eingegrenzt werden können.

FAZIT

Die Risiken der KI nehmen gerade exponentiell zu. Obwohl von der gegenwärtigen KI (Stufe 3) zwar keine inhärenten technischen Gefahren zu erwarten sind, da weder eine Singularität eintritt (bei der die KI die menschliche Intelligenz übertrifft), noch eine KI entsteht, die etwas will oder gar die Menschheit vernichten wollte, so sind bereits viele geplante Anwendungen riskant. Selbst mit den heutigen KI-Systemen können erhebliche Gefahren für Menschen auftreten, sei es durch Überwachungen, durch Social Scoring oder Profiling. Aufgrund der innewohnenden hohen Fehlbarkeit der KI-Systeme (ab Stufe 2) und aus Gründen der unveräußerlichen Menschenwürde müssen bereits die heutigen Systeme streng reguliert werden. Es darf keinen einzigen Einsatzfall geben, bei dem eine KI autonom in die Grundrechte von Menschen eingreift. Und doch ist die wirkliche Gefahr viel größer: Es ist durchaus denkbar, dass sich die Menschen irgendwann einer zwar nicht autonomen, aber überbordenden (Stufe-3-)Maschinenintelligenz unterordnen müssen, und letztlich eine inhumane Maschinenbürokratie entsteht (etwa durch *social scoring*, *crime profiling*, *programmable money* oder anderes mehr). Begründet werden könnte die implementierte Maschinenintelligenz innerhalb der Staatsbürokratie mit der hohen rationalen Intelligenz der KI-Systeme und – wie so oft – mit höheren gesellschaftlichen Zielen, denen sich der einzelne im Kollektiv unterzuordnen hat. Hier kann man nur hoffen, dass kluge Entscheidungsträger diese Gefahren erkennen und dieses Szenario verhindern.

⁰⁶ Vgl. Anna Lorenzen, Neurone in der Petrischale lernen Pong spielen, 21. 11. 2022, www.spektrum.de/news/2078757.

Ist die KI wirklich KI? Diese Frage kann klar mit „Ja“ beantwortet werden: Die KI ist wirklich KI, sie ist intelligent, und künstlich ist sie auch. Heutige Systeme sind physikalisch hergestellt und besitzen eine hohe rationale Intelligenz, auch wenn sie die Intelligenz der Deduktion, der Induktion und der Kognition nur simulieren. Ihr rationaler IQ liegt bei etwa 80, was bereits ziemlich hoch ist (der Durchschnittsmensch hat einen IQ von 100). Aber dennoch ist die Intelligenz der KI im Vergleich zur Intelligenz von Lebewesen gering. Der Grund ist, dass es neben rationaler (kognitiver) Intelligenz noch ganz andere Intelligenzformen gibt, die ein System zum Agieren in natürlichem Umfeld benötigt.

Die heutige KI der dritten Welle ist vollständig mathematisierbar und hervorragend geeignet, um in einem künstlichen, digitalen Datenraum zu agieren. Diese KI ist im Raum möglicher Intelligenzen jedoch nur auf der x-Achse angesiedelt – damit sind ihre Stärken zu sehen, aber auch ihre eklatanten Schwächen. In natürlichem, analogem Umfeld besteht die heutige KI die Anforderungen in der Regel nicht.

Aber die KI schickt sich an, in den nächsten Jahren rudimentäres Bewusstsein zu erlangen, natürlich nur mit einem primitiven, rein physikalischen Maschinenbewusstsein, das nichts mit einem menschlichen Bewusstsein gemein hat, außer den physikalischen Eigenschaften. Diese neue, vierte Stufe der KI wird durch neuromorphe Systeme und Quantencomputer geprägt sein. Solche Systeme werden der mobilen Robotik und dem autonomen Fahren neuen Auftrieb verleihen. Allerdings sollte der Einsatz von KI-Hardware politisch noch reguliert werden. Der diskutierte Entwurf des „AI Act“ der Europäischen Union bezieht sich bedauerlicherweise nur auf KI-Software, also auf Systeme der Stufe 1 bis 3. Aufpassen muss die Gesellschaft zudem bei allen transhumanistischen Bestrebungen, deren Risiken außer bei medizinischen Spezialanwendungen jedwede Chancen übersteigen.

RALF OTTE

ist Professor für Industrieautomatisierung und Künstliche Intelligenz an der Technischen Hochschule Ulm.

ralf.otte@thu.de

GRAUZONEN ZWISCHEN NULL UND EINS

KI und Ethik

Bernd Carsten Stahl

Ist es akzeptabel, wenn Studierende beim Verfassen von Hausarbeiten ChatGPT benutzen? Beuten die großen Internetfirmen ihre Kunden wirtschaftlich aus? Werden Roboter die Menschheit beherrschen oder gar ausrotten? Diese drei beispielhaften Fragen berühren sehr unterschiedliche Bereiche, aber sie haben gemeinsam, dass sie auf ethische Aspekte der Künstlichen Intelligenz verweisen. Solche Fragen werden zurzeit weitreichend diskutiert. Doch worin genau bestehen diese ethischen Probleme, warum sind sie relevant für KI, und was müssen wir beachten, wenn wir über die Ethik der KI sprechen? Und wie können oder sollen solche Fragen bearbeitet werden?

In diesem Beitrag argumentiere ich, dass die Antworten auf solche Fragen nicht generell gegeben werden können, sondern in bestimmten Kontexten oder „Ökosystemen“ angegangen werden müssen. Dies liegt in der Natur der Begriffe „KI“ und „Ethik“, die beide zu vielschichtig sind, um eindeutige Problembeschreibungen oder gar Lösungen zu ermöglichen. Um dieses Argument zu entwickeln, beginne ich mit einer Einführung in die Konzepte der KI sowie in damit verbundene ethische Fragen. Darauf aufbauend argumentiere ich, dass es nicht „die KI“ ist, die Probleme bereitet, oder dass KI eindeutig vorhersagbare Konsequenzen hat, sondern bei der Identifizierung von ethischen Fragen sowie möglichen Antworten immer die soziotechnischen, politischen und ökonomischen Ökosysteme mitgedacht werden müssen, in deren Kontext diese Fragen aufgeworfen werden.

KONZEPTE DER KI

Es ist nicht einfach, KI zu definieren. Als der Begriff 1956 von den Mathematikern und Informatikern John McCarthy, Marvin Minsky, Nathaniel Rochester und Claude Shannon eingeführt

wurde, gingen sie von der Annahme aus, dass es möglich wäre, Maschinen zu entwickeln, die „jeden Aspekt des Lernens oder jede andere Eigenschaft von Intelligenz“ simulieren können.⁰¹ Als Eigenschaften von Intelligenz führten sie die Verwendung von Sprache, die Bildung von Abstraktionen und Konzepten, das Lösen von Problemen, die bisher nur Menschen lösen konnten, und die Selbstverbesserung an.

Daran anschließend und basierend auf eigenen Arbeiten schlage ich vor, drei Verwendungen des Begriffs KI zu unterscheiden, die auf verschiedene ethische Fragestellungen verweisen:⁰²

- maschinelles Lernen,
- konvergierende soziotechnische Systeme und
- Allgemeine KI oder Künstliche Allgemeine Intelligenz.

Es ist dabei wichtig klarzustellen, dass dies keine definitive oder allgemein anerkannte Definition von KI ist, sondern vielmehr ein Versuch, verschiedene Aspekte der Verwendung des Begriffs herauszuarbeiten, die jeweils unterschiedliche Arten von ethischen Fragestellungen aufwerfen und somit auch unterschiedliche Lösungsansätze erfordern.

Die drei Verständnisse von KI gehen jeweils mit einer Reihe von Eigenschaften einher, die sie für bestimmte ethische Fragestellungen relevant machen. *Maschinelles Lernen* etwa neigt dazu, undurchsichtig zu sein; das heißt, es ist schwer zu verstehen, was genau im System passiert.⁰³ Daher kann es herausfordernd sein, die Ergebnisse von Systemen vorherzusagen. Zudem erfordern viele aktuelle Ansätze des maschinellen Lernens große Mengen an Daten für Schulungs- und Validierungszwecke.

KI als *konvergierende soziotechnische Systeme* haben in der Regel einen technischen Kern, der typischerweise eine Form des maschinellen Lernens umfasst. Sie sind jedoch in breitere wirtschaftliche, politische und verwandte Systeme eingebettet und daher klar soziotechnischer Natur. Anwendungsbeispiele sind etwa selbstfahrende Autos oder Systeme zur Erkennung von Finanzbetrug. Diese soziotechnischen Systeme scheinen autonom zu funktionieren, das heißt, sie führen zu Ergebnissen, die Menschen auf verschiedene Weise beeinflussen und für die menschliche Verantwortung nicht zugeschrieben werden kann. Sie bestimmen mit, welche Handlungsoptionen Menschen haben und wahrnehmen können.

Allgemeine KI, manchmal auch als „starke KI“ oder „Künstliche Allgemeine Intelligenz“ (*artificial general intelligence*, AGI) bezeichnet, steht für Systeme, die echte menschliche (oder andere höhere Säugetier-)Intelligenzniveaus aufweisen. Gegenwärtig gibt es kein Beispiel für funktionierende AGI, obwohl spekuliert wird, inwieweit aktuelle Arbeiten an großen Sprachmodellen die Kluft vom maschinellen Lernen zur AGI überbrücken könnten.⁰⁴ AGI ist wichtig für die Diskussion über die Ethik der KI, da sie die Vorstellungskraft der Menschen anregt und Situationen von hoher ethischer Relevanz heraufbeschwört.

ETHIK IN DER KI

Die ethischen Fragestellungen der KI werden leidenschaftlich und kontrovers diskutiert. Auf ihrer grundlegendsten Ebene beschäftigt sich Ethik mit Gut und Böse, mit Richtig und Falsch. Ethische Betrachtungen können dazu genutzt werden, sehr unterschiedliche Phänomene zu beschreiben – angefangen von Intuitionen darüber, was als richtig und falsch eingeschätzt wird, oder

allgemeinen Aussagen bis hin zu strukturierten Reflexionen darüber, wie Ansichten von Richtig und Falsch gerechtfertigt werden können.⁰⁵ In der deutschen Philosophie findet man gelegentlich die Unterscheidung zwischen Moral, welche die gesellschaftlich allgemein anerkannten Werte an sich bezeichnet, und Ethik, womit die theoretische Reflexion darüber gemeint ist. Diese Differenzierung ist zwar nützlich, in der internationalen Diskussion über die Ethik der KI jedoch nicht weit verbreitet.

Es gibt zahlreiche ethische Theorien, die von der philosophischen Ethik behandelt werden. Dazu gehören Tugendethik, pflichtbasierte Ansätze und Theorien, die sich auf die Konsequenzen von Handlungen konzentrieren, um nur einige prominente Beispiele zu nennen. Es gibt eine Vielzahl weiterer Ansätze, die auf KI angewendet werden können, wie die feministische Ethik der Fürsorge oder verschiedene religiös begründete Ethiken. Ein Großteil der aktuellen Diskussion über die Ethik der KI dreht sich darum, aus den verschiedenen Ansätzen allgemeine Grundsätze zu entwickeln. Eine weit zitierte Übersicht listet folgende ethische Grundsätze auf: Transparenz, Gerechtigkeit und Fairness, Nichtschädigung, Verantwortung, Privatsphäre, Wohltätigkeit, Freiheit und Autonomie, Vertrauen, Nachhaltigkeit sowie Würde und Solidarität.⁰⁶ Jeder dieser Grundsätze setzt sich aus mehreren Komponenten zusammen. Transparenz bezieht sich zum Beispiel auf verwandte Konzepte wie Erklärbarkeit, Verständlichkeit, Interpretierbarkeit, Kommunikation und Offenlegung.

Ein weiteres Beispiel sind die Ethikrichtlinien für vertrauenswürdige KI, die von der sogenannten Hocharrangigen Expertengruppe für Künstliche Intelligenz beschrieben wurden.⁰⁷ Die von der EU-Kommission 2018 eingesetzte Expertengruppe schlägt ein Rahmenwerk für vertrauenswürdige KI vor, das aus rechtskonfor-

01 John McCarthy et al., A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955, in: *AI Magazine* 4/2006, S. 12ff., hier S. 13.

02 Dieser Artikel basiert auf der Diskussion zur Ethik der KI, die ich eingeführt habe in: *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies*, Cham 2021.

03 Vgl. Ethem Alpaydin, *Introduction to Machine Learning*, Cambridge MA 2020.

04 Vgl. Sébastien Bubeck et al., Sparks of Artificial General Intelligence: Early Experiments with GPT-4, 22.3.2023, <https://arxiv.org/abs/2303.12712>.

05 Vgl. Bernd Carsten Stahl, *Morality, Ethics, and Reflection: A Categorization of Normative IS Research*, in: *Journal of the Association for Information Systems* 8/2012, <https://aisel.aisnet.org/jais/vol13/iss8/1>.

06 Vgl. Anna Jobin/Marcello Lenca/Effy Vayena, *The Global Landscape of AI Ethics Guidelines*, in: *Nature Machine Intelligence* 9/2019, S. 389–399.

07 Vgl. Hocharrangige Expertengruppe für Künstliche Intelligenz (HEG-KI), *Ethik-Leitlinien für eine vertrauenswürdige KI*, Brüssel 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

mer, ethischer und robuster KI besteht. Dieses Rahmenwerk basiert auf vier ethischen Grundsätzen: Respektierung der menschlichen Autonomie, Vermeidung von Schäden, Fairness und Erklärbarkeit.

Es ist leicht zu erkennen, warum dieser grundsatzbasierte Ansatz attraktiv ist, auch wenn er gelegentlich kritisiert wird:⁰⁸ Er ist unabhängig von häufig umstrittenen ethischen Theorien. Die Grundsätze selbst sind im Allgemeinen nicht kontrovers, bieten also die Möglichkeit eines Konsenses. Aber was bedeutet dies in der Praxis? Welche Probleme ergeben sich, und wie kann mit ihnen umgegangen werden?

CHANCEN UND RISIKEN

Einer Diskussion der ethischen Aspekte der KI sollte vorangestellt werden, dass KI sowohl ethische Vorteile mit sich bringen als auch Probleme aufwerfen kann. Unter den Vorteilen haben die wirtschaftlichen Auswirkungen der KI vermutlich den höchsten Stellenwert. Der Einsatz von KI führt mutmaßlich zu höherer Effizienz und Produktivität. Dies sind ethische Werte, da sie wachsenden Wohlstand versprechen, der vielen Menschen ein besseres Leben ermöglicht und somit förderlich für das menschliche Gedeihen sein kann.

KI bietet technische Möglichkeiten, die unmittelbare ethische Vorteile haben können. So ist KI in der Lage, riesige Datenmengen und -quellen zu analysieren, die Menschen allein nicht verarbeiten könnten. Sie kann Daten verknüpfen, Muster finden und Ergebnisse über verschiedene Bereiche und geografische Grenzen hinweg liefern. Darüber hinaus kann sie konsistenter sein als Menschen, sich rasch an veränderte Kontexte anpassen und Menschen von mühsamen oder sich wiederholenden Aufgaben befreien. All diese Dinge sind zweifellos von großem allgemeinen Nutzen. Die mögliche Reduzierung von Pendelzeiten im Nahverkehr oder die Erhöhung der Effektivität von E-Mail-Spam-Filtern sind nur zwei alltägliche Beispiele dafür, wie KI das Leben vieler Menschen erleichtern kann.

Neben diesen Beispielen für zufällige Vorteile, die sich gewissermaßen als „Nebenwirkungen“ der technischen Entwicklungen erge-

ben, gibt es zunehmend Versuche, KI gezielt für ethische Zwecke zu nutzen. Dies geschieht derzeit unter dem Schlagwort „AI for Good“.⁰⁹ Die Hauptherausforderung besteht dabei darin, zu definieren, was allgemein als (ethisch) gut zählt. Denn in einer pluralistischen Welt ist es nicht ungewöhnlich, dass oftmals keine Einigkeit darüber besteht, was gut ist oder warum etwas als gut gilt. Die Ansätze zur Verwirklichung von „AI for Good“ konzentrieren sich daher häufig auf etablierte moralische Ziele, etwa die allgemeinen Menschenrechte oder die 17 Ziele für nachhaltige Entwicklung der Vereinten Nationen (*sustainable development goals*, SDGs).¹⁰ Menschenrechte und Nachhaltigkeitsziele sind zwei Maßstäbe, die zur Bestimmung der ethischen Vorteile von KI herangezogen und mit den ethischen Bedenken kontrastiert werden können.

Es gibt zahlreiche Veröffentlichungen, die sich mit den ethischen Herausforderungen durch KI auseinandersetzen. Sie stehen größtenteils in der Tradition einer langjährigen Diskussion über Ethik und KI in der Wissenschaft,¹¹ behandeln das Thema aber zunehmend auch aus einer politischen Perspektive.¹² Im Folgenden diskutiere ich ethische Fragen entlang der von mir vorgeschlagenen Kategorisierung der KI.

Maschinelles Lernen

Die erste Gruppe ethischer Fragestellungen ergibt sich aus den Merkmalen des maschinellen Lernens. Viele der maschinellen Lernverfahren, die zum aktuellen Erfolg der KI geführt haben, basieren auf künstlichen neuronalen Netzwerken. Diese sind gekennzeichnet von Undurchsichtigkeit, Unvorhersehbarkeit und einem Bedarf an großen Datensätzen für das Training der Technologien.

Ein primäres und häufig zitiertes ethisches Problem, das sich daraus ergibt, ist das der Pri-

⁰⁸ Vgl. Brent Mittelstadt, *Principles Alone Cannot Guarantee Ethical AI*, in: *Nature Machine Intelligence* 11/2019, S. 501–507.

⁰⁹ Vgl. Bettina Berendt, *AI for the Common Good?! Pitfalls, Challenges, and Ethics Pen-Testing*, in: Paladyn, *Journal of Behavioral Robotics* 1/2019, S. 44–65.

¹⁰ Vgl. Mark Latonero, *Governing Artificial Intelligence: Upholding Human Rights & Dignity*, Data & Society Report, 10. 2018, <https://datasociety.net/library/governing-artificial-intelligence>.

¹¹ Vgl. Mark Coeckelbergh, *AI Ethics*, Cambridge MA 2020; Virginia Dignum, *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*, Cham 2019.

¹² Vgl. HEG-KI (Anm. 7).

vatsphäre und des Datenschutzes. Privatsphäre und Datenschutz sind zwar nicht identisch,¹³ aber im Kontext der KI-Ethik kann Datenschutz als Mittel zur Sicherung der informationellen Privatsphäre verstanden werden. KI, die auf maschinellem Lernen basiert, birgt mehrere Risiken für den Datenschutz. Datenschutzbedenken sind direkt mit Fragen der Datensicherheit verknüpft. Die Cybersicherheit ist ein dauerhaftes Problem der Informations- und Kommunikationstechnologie, nicht nur der KI. KI-Systeme können jedoch neuen Arten von Sicherheitslücken ausgesetzt sein, etwa Modellvergiftungsangriffen.¹⁴ Bei diesen werden Trainingsdaten so verändert, dass das resultierende KI-Modell im Sinne des Angreifers modifiziert wird. Darüber hinaus können diese Systeme für neue Arten von Schwachstellen-Erkennung und -Ausnutzung verwendet werden. Datenschutz- und Datensicherheitsfragen weisen somit auf breitere Fragen der Verlässlichkeit von KI-Systemen hin. Während Verlässlichkeit eine Sorge für alle technischen Artefakte ist, bedeuten die Undurchsichtigkeit von maschinellem Lernen und ihre Unvorhersehbarkeit, dass traditionelle deterministische Testregime auf sie möglicherweise nicht anwendbar sind.

Einige maschinelle Lernsysteme sind nicht transparent. Bei proprietären Systemen, also solchen, die nur auf Geräten des jeweiligen Herstellers funktionieren, kann die kommerzielle Vertraulichkeit von Algorithmen und Modellen die Transparenz weiter einschränken. „Transparenz“ ist selbst ein umstrittener Begriff, aber mangelnde Transparenz wirft Fragen nach der Verantwortlichkeit auf. Mangelnde Transparenz erschwert auch die Erkennung und Behandlung von Fragen der Voreingenommenheit und Diskriminierung. Voreingenommenheit ist eine viel zitierte ethische Frage im Zusammenhang mit KI. Eine zentrale Herausforderung besteht darin, dass ma-

schinelle Lernsysteme, absichtlich oder unbeabsichtigt, bereits bestehende Vorurteile reproduzieren können.¹⁵

Konvergierende soziotechnische Systeme

Die zweite Gruppe ethischer Fragestellungen besteht aus solchen, die sich auf das beziehen, was ich als konvergierende soziotechnische Systeme bezeichne. Merkmale dieser Systeme sind Autonomie und soziale Auswirkungen einschließlich der Möglichkeit, dass menschliches Verhalten manipuliert wird.

Die sich daraus ergebenden ethischen Probleme sind mit KI verbunden, aber nicht auf KI beschränkt: Sie werfen die breitere gesellschaftliche und politische Frage auf, wie große soziotechnische Systeme strukturiert und genutzt werden sollten. Sie können daher nicht isoliert von ihrer gesellschaftlichen Rolle betrachtet werden, und viele der ethischen Fragestellungen werden direkt durch die Art und Weise verursacht, wie die Gesellschaft mit diesen Technologien arbeitet. So hängen viele Fragen, die sich aus dem Leben in einer digitalen Welt ergeben, mit der Wirtschaft zusammen. Die prominenteste unter ihnen betrifft die möglichen Auswirkungen auf den Arbeitsmarkt. Es wird befürchtet, dass KI sich negativ auf die Beschäftigung auswirken könnte, obwohl der erwartete Effekt umstritten bleibt.¹⁶

Eine zentrale Herausforderung ist zudem die Konzentration wirtschaftlicher (und implizit politischer) Macht. Die Abhängigkeit der aktuellen KI-Systeme von großen Rechenressourcen und massiven Datenmengen bedeutet, dass Organisationen, die solche Ressourcen besitzen oder Zugang dazu haben, besonders gut positioniert sind, um von KI zu profitieren. Die internationale Konzentration solcher wirtschaftlichen Macht bei den großen Technologieunternehmen ist nicht nur auf KI gestützt, aber KI-bezogene Technologien haben das Potenzial, das Problem zu verschärfen.¹⁷

13 Vgl. Giovanni Buttarelli, *Choose Humanity: Putting Dignity Back into Digital*, Opening Speech of Debating Ethics Public Session of the 40th Edition of the International Conference of Data Protection Commissioners, Brüssel 24. 10. 2018, www.privacy-conference2018.org/system/files/2018-10/Choose%20Humanity%20speech_0.pdf.

14 Vgl. Matthew Jagielski et al., *Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning*, in: *The Institute of Electrical and Electronics Engineers (Hrsg.), 2018 IEEE Symposium on Security and Privacy*, San Francisco 2018, S. 19–35.

15 Vgl. Access Now (Hrsg.), *Human Rights in the Age of Artificial Intelligence*, November 2018, www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf.

16 Vgl. Leslie Willcocks, *Robo-Apocalypse Cancelled? Reframing the Automation and Future of Work Debate*, in: *Journal of Information Technology* 4/2020, S. 286–302.

17 Vgl. Paul Nemitz, *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, in: *Philosophical Transactions of the Royal Society A*, 2133/2018, <https://doi.org/10.1098/rsta.2018.0089>.

Damit verbunden sind breitere gesellschaftliche Themen wie Machtbeziehungen und Machtasymmetrien. Wirtschaftliche Dominanz und die Gestaltung von Handlungsoptionen können einigen Akteuren große Macht verleihen, bis hin zu dem Punkt, an dem demokratische Prinzipien gefährdet sind. Dies ist nicht nur ein Problem neuer Formen von Überwachung, Manipulation und Autoritarismus. Unsere demokratischen Strukturen können ebenfalls durch die Schaffung des Überwachungskapitalismus untergraben werden.¹⁸

Ein anderes wohlbekanntes ethisches Problem der Informations- und Kommunikationstechnologien kann sich durch KI verschärfen, nämlich das der digitalen Ausgrenzung. Durch die Bereitstellung oder den Entzug von Informationen gestalten die Technologien, die uns umgeben, den Raum möglicher Handlungen. Informations- und Kommunikationstechnik im Allgemeinen und KI im Besonderen können die Handlungsoptionen eines Menschen erscheinen oder verschwinden lassen, ohne dass dieser Mensch sich dessen bewusst ist. KI kann aber auch enorme Möglichkeiten eröffnen und Räume für Handlungen schaffen, die zuvor undenkbar waren, beispielsweise indem sie sehbehinderten Menschen das autonome Fahren von Fahrzeugen ermöglicht oder maßgeschneiderte medizinische Lösungen jenseits des derzeit Möglichen schafft.

Der Verbesserung von Diagnosen und Behandlungen durch KI im Gesundheitswesen stehen jedoch auch Risiken und Nachteile gegenüber. Ein Beispiel sind Pflegetechnologien: Robotersysteme wurden lange als Möglichkeit zur Bewältigung der Herausforderungen des Pflegesektors vorgeschlagen, aber es gibt Bedenken, dass der menschliche Kontakt durch Technologie ersetzt werden könnte, was oft als grundlegendes ethisches Problem angesehen wird.

Ein prominent diskutiertes Problem ist zudem die Verwendung von KI zur Schaffung autonomer Waffen. Während nachvollziehbar ist, dass das Ersetzen von menschlichen Soldaten durch Roboter, um das Leben der Soldaten zu retten, ein ethischer Vorteil sein könnte, gibt es zahlreiche Gegenargumente: angefangen bei praktischen Aspekten wie der Zuverlässigkeit solcher Systeme über politische Erwägungen wie die Frage, ob

autonome Waffen die Schwelle zum Beginn von Kriegen senken würden, bis hin zu grundlegenden Gesichtspunkten wie Frage, ob es jemals angemessen sein kann, menschliche Leben auf der Grundlage von maschinellen Eingaben zu beurteilen.

Künstliche Allgemeine Intelligenz

Künstliche Allgemeine Intelligenz gibt es momentan zwar (noch) nicht, dennoch ist sie Gegenstand erheblicher ethischer Diskussionen. Ich bezeichne die damit verbundenen Fragestellungen als langfristig, da sie, wenn überhaupt, vermutlich nicht kurzfristig zu lösen sind. In der Diskussion über ethische Fragen der KI berücksichtige ich sie aus zwei Gründen dennoch. Erstens regen Fragen über AGI zum Nachdenken an, nicht nur in Fachkreisen, sondern auch in den Medien und in der Gesellschaft insgesamt, weil sie viele grundlegende Aspekte der Ethik und der Menschlichkeit berühren. Zweitens werfen einige dieser Fragen ein Licht auf praktische Probleme der aktuellen KI, weil sie zu einer klareren Reflexion über Schlüsselkonzepte wie Autonomie und Verantwortung sowie über die Rolle der Technologie in einer guten Gesellschaft zwingen.

Die techno-optimistische Version von AGI besagt, dass es einen Zeitpunkt geben wird, an dem die KI ausreichend fortgeschritten ist, um sich selbst zu verbessern, und infolge einer positiven Rückkopplungsschleife der KI auf sich selbst eine Intelligenzexplosion – die Singularität – eintreten wird.¹⁹ Dies werde zur Entstehung von Superintelligenz führen.²⁰ Eine solche wäre dann nicht nur in den meisten oder allen kognitiven Aufgaben besser als der Mensch, sondern auch in der Lage, ein Bewusstsein und Selbstbewusstsein zu entwickeln. Die Beiträge zu dieser Diskussion sind sich uneinig darüber, was als Nächstes passieren würde. Die superintelligente AGI könnte wohlwollend sein und das menschliche Leben verbessern, sie könnte die Menschen als Konkurrenten sehen und uns vernichten, oder sie könnte in einer anderen Bewusstseinssphäre existieren und die Menschheit größtenteils ignorieren.

¹⁹ Vgl. Raymond Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, New York 2005.

²⁰ Vgl. Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford–New York 2016.

¹⁸ Vgl. Shoshana Zuboff, *The Age of Surveillance Capitalism*, New York 2019.

DER ÖKOSYSTEM-ANSATZ

Schon dieser kurze und keineswegs erschöpfende Überblick zeigt, dass es eine Menge ethischer Aspekte gibt, die im Zusammenhang mit KI zu bedenken sind. Viele von ihnen sind umstritten und erzwingen *trade-offs* beziehungsweise schwierige Abwägungen von Vor- und Nachteilen, weil das eine nicht ohne das andere zu haben ist. So ermöglichen neuronale Netzwerke die automatische Erkennung von Mustern in Daten – da diese aber meist nicht transparent sind, werden auch Benachteiligung und Diskriminierung begünstigt. KI hat enormes ökonomisches Potenzial, aber die Wohlstandsvermehrung führt nicht automatisch zu gerechter Verteilung. Weiterhin sind die ethischen Grundlagen und Begriffe interpretationsbedürftig und oft nicht einfach anwendbar. Die Konsequenzen von Interventionen in KI sind typischerweise nicht leicht vorherzusehen. Und die Anzahl der Stakeholder im KI-Umfeld ist groß. Bei ihnen handelt es sich nicht nur Designer und Produzenten von KI, sondern auch um deren Nutzer, Verbraucher, staatliche Institutionen und zivilgesellschaftliche Organisationen aller Art.

Es ist daher nicht verwunderlich, dass es zahlreiche Vorschläge gibt, die darauf abzielen, die ethischen Fragen der KI anzugehen. Sie umfassen staatliche und gesamtgesellschaftliche Interventionen, einschließlich Regulierung und Gesetzgebung. Der in Vorbereitung befindliche „AI Act“ der Europäischen Union ist ein prominentes Beispiel dafür. Weitere staatliche und internationale Ansätze setzen auf die spezifische Anwendung von Menschenrechten auf KI, die Schaffung neuer Institutionen sowie die Verabredung internationaler Koordination. Auch Firmen und andere private Organisationen können zur Bearbeitung ethischer Fragen im Zusammenhang mit KI beitragen – in ihren Strategien und Leitbildern oder auch in alltäglichen Prozessen, etwa zum Risikomanagement oder Datenschutz. Ebenso haben Einzelpersonen als Bürger oder Mitarbeiter vielschichtige Rollen bei der Lösung ethischer Probleme der KI: Sie können unter anderem bei der Entwicklung von Standards oder deren Zertifizierung mitwirken, sich für die Umsetzung von bestehenden Richtlinien engagieren oder als Entwickler an Methodologien mitarbeiten oder sie einsetzen.

Aufgrund der Mannigfaltigkeit und Komplexität der verschiedenen KI-Arten und der ethi-

schen Fragen, die sie aufwerfen, ist nicht davon auszugehen, dass es eine klare Antwort gibt, wie diese Fragen zu behandeln oder zu lösen sind. Anstatt davon auszugehen, dass KI klar vorher-sagbare Konsequenzen hat, die ethisch beurteilt und dann angegangen werden können, sollten wir KI anders verstehen: KI ist eine Gruppe von soziotechnischen Systemen, die in ökonomischen, sozialen und politischen Systemen integriert sind. Um diese Sichtweise widerzuspiegeln, wird häufig die Metapher des Ökosystems genutzt. Die Frage nach der Ethik der KI ist dann nicht mehr, was genau das ethische Problem ist und wie es zu lösen ist, sondern wie wir die verschiedenen KI-Ökosysteme formen und orchestrieren, um moralisch akzeptable und sozial wünschenswerte Zustände zu erreichen.

Dieser Vorschlag mag auf den ersten Blick frustrierend erscheinen. Er erlaubt nämlich keine (einfachen) Antworten auf die Fragen, die ich zu Beginn dieses Artikels angeführt habe. So lässt sich in dieser Sichtweise nicht eindeutig beantworten, ob es ethisch vertretbar für Studierende ist, ChatGPT zu benutzen, ob die Marktmacht der Internetfirmen ethisch problematisch ist oder wie wir mit Fragen der AGI umgehen sollten. Allerdings sollte es uns nicht überraschen, dass es auf diese Fragen keine einfachen Antworten gibt, sondern sie kontextspezifisch beantwortet werden müssen. KI als Teil unterschiedlicher Ökosysteme zu sehen, erlaubt es, die spezifische Konstellation von Technologie, Anwendung und sozialem Kontext in den Vordergrund zu rücken. Dies kann dann zu einer detaillierteren Analyse der spezifischen ethischen Fragen führen, aufgrund derer darauf abgestimmte Lösungen gefunden werden können. Ein solcher kontextspezifischer Ansatz der Ethik von KI-Ökosystemen scheint mir daher zielführender zu sein als der Versuch, allgemeingültige Lösungen zu finden, die im Detail und in der Anwendung oft nicht passen.

BERND CARSTEN STAHL

ist Professor für Kritische Technologieforschung an der University of Nottingham, Vereinigtes Königreich.

bernd.stahl@nottingham.ac.uk

KI UND DEMOKRATIE: ENTWICKLUNGSPFADE

Thorsten Thiel

Digitaltechnologische Entwicklungen haben den Ruf, in höchstem Maße disruptiv zu sein. Als radikale Technologien verbessern sie nicht nur bestehende Prozesse, sondern revolutionieren unseren Blick auf Ziele und Möglichkeiten gesellschaftlichen Handelns.⁰¹ So haben etwa das Internet, soziale Netzwerke oder Smartphones tiefe Veränderungen bewirkt, beispielsweise in der Art und Weise, wie wir Wissen organisieren und unser Handeln koordinieren oder wie unsere Wirtschaft funktioniert. Auch vor Politik und Demokratie macht der digitaltechnologisch getriebene Strukturwandel nicht Halt. So wird etwa seit der Ausbreitung des öffentlichen Internets in den 1990er Jahren mit wechselnden Thesen, aber immer hoher Dringlichkeit, Digitalisierung als Gefahr oder Chance für die Demokratie diskutiert.⁰²

Es ist insofern nicht überraschend, dass auch in den jüngsten Auseinandersetzungen um Künstliche Intelligenz, die seit dem öffentlichen Launch von ChatGPT Ende 2022 ein neues Level erreicht hat, die Auswirkungen auf Demokratie zum wichtigen Thema geworden sind. Anknüpfend an die virulente Kritik an sozialen Medien und deren Auswirkungen auf den demokratischen Diskurs stehen dabei insbesondere die Risiken von Manipulation und Desinformation im Blickpunkt. Die demokratischen Herausforderungen, die durch eine breite gesellschaftliche Einbettung von KI in unser Alltagsleben entstehen, sind aber vielschichtiger, weshalb im Folgenden ein Überblick über Entwicklungspfade gegeben wird, der aufschlüsselt, wie sich demokratische Prozesse und demokratische Öffentlichkeit im sich wandelnden soziotechnischen Kontext verändern könnten.

NEUE DYNAMIK

Künstliche Intelligenz ist nicht erst seit ChatGPT eine gesellschaftlich relevante Technologie. Genauer gesagt ist Künstliche Intelligenz überhaupt

keine einzelne Technologie, sondern wird besser als eine kategoriale Bezeichnung für technologische Verfahren verstanden, die das Ziel haben, komplexe Probleme unabhängig und situationsgerecht zu lösen. Die Verfahren, mit denen dieses Ziel erreicht werden soll, sind dabei nur abstrakt von biologischen Fähigkeiten und Strukturen inspiriert. Künstliche Intelligenz unterscheidet sich von menschlicher Intelligenz auch dann, wenn bestimmte Leistungen oder Ergebnisse vergleichbar erscheinen. Zu analysieren ist daher, wie sich gesellschaftliche Kommunikationen und Prozesse ändern, wenn in ihr verschiedene Rationalitäten und Agenten zusammenwirken.⁰³

Seit Mitte des 20. Jahrhunderts lassen sich grob zwei Strategien unterscheiden, KI zu realisieren: Auf der einen Seite stehen Ansätze, die auf Klassifizierungen und der nach logischen Regeln operierenden Verarbeitung von Daten aufsetzen – sogenannte symbolische KI –, auf der anderen konnektionistische Verfahren, die auf maschinelles Lernen setzen und mittels neuronaler Netze versuchen, Probleme situationsgerechter zu analysieren, in Daten Muster und Wahrscheinlichkeiten zu bestimmen und diese Ergebnisse zu neuen Lösungen zu kombinieren.⁰⁴

Für den um 2010 einsetzenden und sich seitdem extrem beschleunigenden neuerlichen Aufschwung von KI sind insbesondere letztere Verfahren verantwortlich. Grundlegend für den Erfolg ist zum einen die Verfügbarkeit über enorm große Rechenkapazitäten, die zudem immer spezifischer für KI-Anwendungen ausgelegt werden; zum anderen der Zugriff auf riesige Datenmengen, die sich beispielsweise aus der Verlagerung und Vermessung von Alltagskommunikation in soziale Medien und die Allgegenwart von Sensoren und Kameras in Zeiten des mobilen Internets speist. Sichtbare Erfolge von KI-Verfahren – etwa durch Verbesserungen in den Bereichen Bilderkennung oder Übersetzungen – haben dann eine wirtschaftliche Dynamik ausgelöst, die

wiederum verstärkend auf die Weiterentwicklung der Technologie wirkt.⁰⁵ Große Sprachmodelle, die die Transformation von Text und anderen Medien erlauben, sind der jüngste Erfolg dieser Entwicklung. Sie resultieren in generativer KI, die die Mustererkennung analytischer Verfahren durch Vorhersagen und Rekombinationen fortschreibt und so neue Anwendungsfelder wie die Erzeugung von Text oder Videos erschließt.⁰⁶

Wie aber werden diese technologischen Veränderungen relevant für demokratische Prozesse? Wie immer in Untersuchungen, die sich der Schnittstelle von Technologie und Gesellschaft nähern, ist es wichtig, hier nicht die Technik zum alleinigen Akteur zu machen oder eine monolithische Logik aus ihr zu folgern. Untersucht werden muss vielmehr, wie die Konstellation, die aus der Verbindung von durch Technologie beförderten Handlungsmöglichkeiten, ökonomischen Anreizen, politischer Regulierung und gesellschaftlichen Nutzungspraktiken entsteht, in Bezug auf politische Kontexte Veränderungsimpulse erzeugt.⁰⁷ Bezüglich der gegenwärtig dominanten Verfahren lohnt es sich dabei, drei Bereiche näher zu betrachten, in denen die fortgesetzte gesellschaftliche Adaption von KI-Anwendungen Impulse erzeugt, die die Art und Weise, wie wir Demokratie leben, betreffen. Der erste Bereich sind die Effekte auf die Struktur und Dynamik demokratischer Öffentlichkeit, die generative KI erzeugt. Der zweite Bereich betrifft den Einsatz von KI für die Zwecke des demokratischen Regierens. Drittens lohnt es sich, zu prüfen, inwiefern KI mittel- und langfristig die demokratische Teilhabe von Bürger*innen verändert.

01 Vgl. Adam Greenfield, *Radical Technologies: The Design of Everyday Life*, London 2017.

02 Vgl. Thorsten Thiel, *Demokratie in der digitalen Konstellation*, in: Gisela Riescher/Beate Rosenzweig/Anna Meine (Hrsg.), *Einführung in die Politische Theorie*, Stuttgart 2020, S. 331–349.

03 Vgl. Jeanette Hofmann, *Demokratie und Künstliche Intelligenz*, 2022, <https://digid.jff.de/demokratie-und-ki>.

04 Vgl. Alfred Früh/Dario Haux, *Foundations of Artificial Intelligence and Machine Learning*, Weizenbaum Series 29/2022, <https://doi.org/10.34669/WI.WS/29>.

05 Vgl. Jeffrey Dean, *A Golden Decade of Deep Learning*, in: *Daedalus* 151/2022, S. 58–74.

06 Vgl. Steffen Albrecht, *ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen*, Büro für Technikfolgenabschätzung beim Deutschen Bundestag, Hintergrundpapier Nr. 26, 21. 4. 2023.

07 Vgl. Sebastian Berg/Niklas Rakowski/Thorsten Thiel, *Die digitale Konstellation. Eine Positionsbestimmung*, in: *Zeitschrift für Politikwissenschaft* 2/2022, S. 171–191.

DEMOKRATISCHE ÖFFENTLICHKEIT

Vor 2022 drehte sich die Diskussion um die Transformation demokratischer Öffentlichkeit durch KI-Verfahren hauptsächlich um den Einsatz von KI in der Moderation von Inhalten auf sozialen Netzwerken, worin eine Verstärkung bereits bekannter Probleme mit der strukturellen und strukturierenden Macht von Plattformen gesehen wurde.⁰⁸ Die Nutzungserfahrung vieler Menschen mit generativer KI hat aber einen neuen Fokus gebracht: Der transformative Anstoß geht nun von der Senkung der Kosten für die Erstellung und Veränderung kommunikativer Produkte aus. Die Möglichkeit, Bilder, Videos oder Stimmen in Echtzeit und überzeugend zu erschaffen oder zu verändern, oder auf Knopfdruck Texte zu jedwedem Thema und in jedwedem Stil entstehen zu lassen, erzeugt ein öffentliches Bewusstsein für eine potenziell revolutionäre Veränderungsdynamik. Generative KI verändert dabei auch Informationsumgebungen jenseits individueller Handlungsmöglichkeiten: So erlauben die Verfahren, Information und Kommunikation viel stärker maßzuschneidern, wie das Beispiel digitaler Assistenten zeigt, die in natürlicher Sprache mit uns plaudern und Informationen kuratieren. Was also sind die erwartbaren Auswirkungen auf das Verhältnis von Öffentlichkeit und Demokratie?

An erster Stelle steht die Befürchtung der epidemischen Zunahme von Desinformation. Generative KI ermöglicht sogenannte Deep Fakes, bei denen multimediale Inhalte wie Videos oder Tonaufnahmen erzeugt oder tiefgreifend verändert werden. Wie niedrigschwellig diese Form von Eingriffen ist, weiß jeder, der schon mal digitale Filter für Bilder oder Videos verwendet hat. Obwohl Manipulationen und Fälschungen immer schon Teil politischer Kommunikation waren, verschiebt generative KI hier die Gewichte: Viel mehr Akteure erhalten durch die Anwendungen die Möglichkeit, auch komplexe Medienarten schnell und täuschend echt zu verändern. Und gerade diesen Formaten wird ein besonders großer Einfluss auf den öffentlichen Diskurs zugeschrieben, da das, was wir sehen oder hören, oft einen größeren Eindruck macht, als das, was wir lesen. Die Manipulation solcher immersiven Medienformate gilt daher als subtil und tiefgreifend zugleich.

08 Vgl. Tarleton Gillespie, *Content Moderation, AI, and the Question of Scale*, in: *Big Data & Society* 2/2022, S. 1–5.

Hinzu kommt, dass nicht nur die Qualität manipulativer Eingriffe steigt, sondern dass die neuen Möglichkeiten sich mit einem medialen Kontext verbinden, der ohnehin bereits als anfällig für Desinformation und Emotionalisierung gilt. Die Aufmerksamkeitsökonomie sozialer Medien mit ihren segmentierten Adressatengemeinschaften und hohem Erregungslevel ebenso wie die heute so wichtige Gruppenkommunikation in Messenger-Diensten macht es sowohl schwierig, Desinformationen zu entdecken, als auch diese dann erfolgreich zu kontern. Gerade strategisch agierende Akteure können daher durch generative KI profitieren und die entstehenden Potenziale ausnutzen, um etwa Radikalisierungsprozesse zu vertiefen.⁰⁹

Neben der gezielten Desinformation muss, zweitens, Misinformation zu den problematischen Auswirkungen von KI gezählt werden. Misinformation resultiert aus der immanenten Unzuverlässigkeit der angewandten Verfahren, die aus der Generierung neuer Inhalte durch die maschinelle Interpretation von Trainingsdaten erfolgt. Das heißt, mitunter „halluziniert“ generative KI – sie schafft Zusammenhänge, ohne dass diese zu prüfen oder zu belegen wären. Dies geschieht in einer überzeugenden Weise, da die Verfahren ja gerade darauf ausgerichtet sind, menschliche Erwartungen zu erfüllen und oftmals Vorurteile fortgeschrieben werden, die gesellschaftlich verbreitet und daher in die Trainingsdaten eingeschrieben sind. Auf dieses Phänomen reflektiert zu reagieren, setzt hohe Medienkompetenz voraus, da die Autorität der Information groß und die Möglichkeit ihrer Überprüfung gering ist. KI-erzeugte Inhalte nicht als sinnvolle, sondern als allein der Form nach passende Ergebnisse zu interpretieren, erfordert einen deutlichen Bruch mit bisherigen Rezeptionsgewohnheiten.¹⁰

Wenn Desinformation und Misinformation in einer Gesellschaft zunehmen, ist dies bereits für sich eine Herausforderung für den demokratischen Diskurs. Die Basis, auf der sich Gesellschaften austauschen und über Kompromisse verhandeln, wird kleiner, Fragmentierung wahrscheinlicher. Trotzdem ist offen, inwiefern die durch KI erzeugte Mis- und Desinformation tatsächlich qualitativ so

bedeutsam ist, wie im öffentlichen Diskurs häufig unterstellt – und ob nicht auch Möglichkeiten gesellschaftlichen Anpassens und Lernens stärker hervorgehoben werden müssten.¹¹ Schließlich sind demokratische Öffentlichkeiten in Gesellschaften wie Deutschland stark verschachtelt, sie sind geprägt von einem pluralistischen Mediensystem, breiter politischer Bildung und einer Vielzahl von alternativen Kommunikationswegen. Die Möglichkeiten, Öffentlichkeit mittels digitaler Artefakte großflächig und nachhaltig zu manipulieren, sind somit begrenzt.¹² Die Gefahr für den öffentlichen Diskurs ist insofern weniger die kurzfristige Manipulation, sondern eher die langfristige Erosion der (infra-)strukturellen Bedingungen demokratischer Öffentlichkeit. Generative KI kann zu dieser beitragen, da schon der Vorwurf der Manipulierbarkeit über Zeit den Effekt haben kann, dass medialer und öffentlicher Kommunikation weniger vertraut wird.¹³ Diese Gefahr wird zusätzlich dadurch verschärft, dass generative KI verändert, wie wir uns in den überbordenden Informationsumwelten unserer Gegenwart bewegen. Chatbots und Assistenzsysteme wie auch die Veränderung von Suchmaschinen – weg von der Darstellung einer Liste an Verweisen hin zu einer autoritativ ausformulierten Antwort – bedeuten, sofern sie sich durchsetzen, gravierende Veränderungen unserer Mediennutzung, was in zwei Hinsichten in Bezug auf Demokratie relevant werden könnte.

Zum einen könnte die in der Forschung eigentlich schon länger angezweifelte Filterblasenhypothese neue Plausibilität bekommen. Diese besagt, dass algorithmische Verfahren, die Inhalte personalisieren, eine Vereinseitigung des Informationsangebots nach sich ziehen. Da Algorithmen selektieren, was wir ohnehin hören wollen, bewegten wir uns vorwiegend unter Gleichgesinnten, was demokratische Kompromisse und Perspektivwechsel erschwere.¹⁴ Obwohl diese

09 Vgl. Bobby Chesney/Danielle Citron, Deep Fakes: A Looming Challenge for Privacy, in: *California Law Review* 107/2019, S. 1752–1819.

10 Vgl. Paola Lopez, ChatGPT und der Unterschied zwischen Form und Inhalt, in: *Merkur* 891/2023, S. 15–27.

11 Vgl. Sayash Kapoor/Arvind Narayanan, How to Prepare for the Deluge of Generative AI on Social Media, 16.6.2023, <http://knightcolumbia.org/content/how-to-prepare-for-the-deluge-of-generative-ai-on-social-media>.

12 Vgl. Andreas Jungherr, Digitale Informationsumgebungen: Echokammern, Filterblasen, Fake News und Algorithmen, in: Norbert Kersting et al. (Hrsg.), *Handbuch Digitalisierung und politische Beteiligung*, Wiesbaden 2022.

13 Vgl. David Karpf, How Digital Disinformation Turned Dangerous, in: Lance Bennett/Steven Livingston, *The Disinformation Age*, Cambridge 2020, S. 153–168.

14 Vgl. Cass R. Sunstein, *#Republic: Divided Democracy in the Age of Social Media*, Princeton 2017.

These intuitiv plausibel und im öffentlichen Diskurs stark verbreitet ist, haben empirische Arbeiten die Existenz und die Effekte von Filterblasen stark relativiert.¹⁵ Generative KI könnte den für Suchmaschinen und soziale Netzwerke tendenziell widerlegten Argumenten aber neue Kraft verleihen, da die umfassende Vorstrukturierung der Kommunikation und das Verdecken von Komplexität durch gradlinige und direkte Antworten erneut als Vereinseitigung gedeutet werden könnte.

Ebenso offen, aber potenziell bedenklich sind die Auswirkungen, die generative KI auf die Ökonomie von Medien und Journalismus haben wird. Hier ist unter anderem zu fragen, wie mit journalistischen Inhalten Geld verdient werden kann, wenn die Schnittstelle zur Interaktion mit den Leser*innen oder Konsument*innen noch mehr in von Technologiekonzernen kontrollierte Umgebungen abwandert. Eine weitere Frage ist, wie sich die Trennung von informierten Eliten und desinteressiertem Publikum unter dem Einfluss der Technologie verändert.¹⁶

KI UND DEMOKRATISCHES REGIEREN

Die breite gesellschaftliche Adaption von KI-Verfahren bringt auch jenseits von generativer KI und öffentlichem Diskurs eine Reihe neuer Herausforderungen mit sich. Ein Bereich, der besondere Aufmerksamkeit verdient, ist die Veränderung demokratischer Regierungspraxis durch assistierte oder gar automatisierte Entscheidungssysteme.¹⁷ Diese finden in Europa, aber auch weltweit immer mehr Verbreitung und werden selbst dann weiter ausgebaut, wenn sich in der Praxis zeigt, dass viele der bisher eingesetzten Systeme die in sie gesetzten Erwartungen nicht erfüllen.¹⁸

15 Vgl. Birgit Stark/Melanie Magin/Pascal Jürgens, Maßlos überschätzt. Ein Überblick über theoretische Annahmen und empirische Befunde zu Filterblasen und Echokammern, in: Mark Eisenegger et al. (Hrsg.), *Digitaler Strukturwandel der Öffentlichkeit*, Wiesbaden 2021, S. 303–321.

16 Vgl. Andreas Jungherr, *Artificial Intelligence and Democracy: A Conceptual Framework*, in: *Social Media + Society* 9/2023, <https://doi.org/10.1177/20563051231186353>.

17 Vgl. Michael Veale/Irina Brass, *Administration by Algorithm? Public Management Meets Public Sector Machine Learning*, in: Karen Yeung/Martin Lodge (Hrsg.), *Algorithmic Regulation*, Oxford 2019, S. 121–150.

18 Vgl. *AlgorithmWatch/Bertelsmann Stiftung, Automating Society Report 2020*, Berlin–Gütersloh 2021.

Für den anhaltenden Boom verantwortlich sind die zugeschriebenen analytischen Fertigkeiten von auf Mustererkennung spezialisierten KI-Verfahren. Diese befeuern Datensammlung und Anwendungsentwicklung und erhöhen die Erwartungen an eine datenzentrierte und effektive Politik. Präzisere Analyse sowie der Einsatz von Vorhersagen und Simulationen sollen Schwächen repräsentativdemokratischer Verfahren ausgleichen, zu denen ein beschränkter Zeithorizont, fehlende Responsivität in dynamischen Situationen und eine zu ungenaue Anwendung von Steuerungsinstrumenten gezählt werden. KI-basierte Verfahren sollen Politik besser machen und dadurch die Output-Legitimation demokratischer Systeme erhöhen.¹⁹ Hierbei wird an eine ältere Vorstellung von Politik angeknüpft, die schon in Statistik und Bürokratie ihren Ausdruck fand, sich nun aber zunehmend von der Verwaltung bis zur Zielformulierung und Planung von Politik erstreckt.²⁰

Öffentliche Machtausübung setzt in demokratischen Staaten ein besonders hohes Maß an Kontrolle voraus: Sie ist an Grund- und Bürgerrechte gebunden, muss Standards von Gleichheit und Gerechtigkeit genügen und in einer Weise ausgeübt werden, dass die ihr Unterworfenen nicht nur nachvollziehen können, wie sie regiert werden, sondern auch wirkungsvoll gegen Missbrauch und Fehlentwicklungen opponieren können. Der zunehmende Einsatz von KI-Verfahren birgt hierbei zahlreiche Herausforderungen.²¹

Mit Blick auf Grundrechte und Gleichheitsfragen liegt die Schwierigkeit in der Logik der Fortschreibung vergangener Muster, in denen eine Reproduktion von Macht- und Statusasymmetrien angelegt ist. Muster werden hier zu Regeln, die sich dann wieder selbst bestätigen, wie es aus Fällen von *predictive policing* bis hin zur Administration von Gesundheitsleistungen bekannt ist.²² Verstärkend wirkt, dass KI oft dort zum Einsatz kommt, wo ein besonderer Kostendruck

19 Vgl. Helen Margetts, *Rethinking AI for Good Governance*, in: *Daedalus* 151/2022, S. 360–371.

20 Vgl. Pascal König/Georg Wenzelburger, *Opportunity for Renewal or Disruptive Force? How Artificial Intelligence Alters Democratic Politics*, in: *Government Information Quarterly* 3/2020, <https://doi.org/10.1016/j.giq.2020.101489>.

21 Vgl. Karen Yeung, *The New Public Analytics as an Emerging Paradigm in Public Sector Administration*, in: *Tilburg Law Review* 2/2023, S. 1–32.

22 Vgl. Paola Lopez, *Artificial Intelligence und die normative Kraft des Faktischen*, in: *Merkur* 861/2021, S. 42–52.

herrscht und administrative Ressourcen gespart oder mit geringen Kosten stark ausgeweitet werden sollen, etwa bei der Grenzüberwachung. Gerade in diesen Bereichen werden daher vulnerable Gruppen oft zum Gegenstand einer sie benachteiligenden Entscheidungsfindung.²³

Mit Blick auf demokratische Normen der Kontrolle und Rechenschaftspflicht kommt ein weiteres Problem hinzu: die opake technische Struktur der auf maschinellem Lernen beruhenden KI-Verfahren und der Umstand, dass politische Institutionen sich für deren Entwicklung und Betrieb stark von der Infrastruktur und den Leistungen privatwirtschaftlicher Akteure abhängig machen. Ohne klare Regelungen, die sicherstellen, dass Kontrolle und Rechenschaft praktisch möglich bleiben, untergraben KI-Verfahren demokratische Verfahrensgarantien.²⁴

Dies wird zusätzlich davon überlagert, dass datengetriebenes Entscheidungshandeln oft als in Spannung stehend zu demokratischer Selbstbestimmung konzipiert wird. Die Kritik ist hier, dass Demokratien durch das Setzen auf KI-Verfahren sich einer hierarchischen Logik unterwerfen, die Daten fetischisiert und darüber die Möglichkeit partizipatorischer Selbstbestimmung negiert.²⁵ Ironischerweise geht eine solche Position aber selbst von einem starken technologischen Imperativ aus, in dem die Fähigkeit gesellschaftlichen Handelns und demokratischer Politisierung von vornherein geringgeschätzt wird und darüber verpasst wird, dass es gerade in Bezug auf technologische Steuerung immer wieder zu wirkungsvoller Politisierung kommt.²⁶

POLITISCHE PARTIZIPATION

Werfen wir als Drittes einen Blick darauf, wie eine breite gesellschaftliche Aneignung von KI-Verfahren sich auf politische Teilhabe von Bürger*innen auswirken könnte. Verändert KI die Möglichkeiten und Modi politischer Partizipation? Antworten hierauf sind noch spekulativer als in den beiden

zuvor betrachteten Feldern, was daran liegt, dass dies nicht nur von der Ausgestaltung der Anwendungen abhängt, sondern auch von der Anbindung und Akzeptanz im politischen System. Unterscheiden lassen sich zwei mögliche Richtungen – *bottom-up* und *top-down* – und eine Gegentendenz.

Eine *Top-down*-Intensivierung bedeutet, dass seitens der Politik unter Rückgriff auf KI-Anwendungen der Austausch mit den Bürger*innen gezielt ausgebaut wird. Dies liegt allgemein auf der Linie einer zunehmend kommunikativ auftretenden Politik, und es entspricht der Erwartung, dass Digitalisierung direkte und unmittelbare Kommunikation befördert. KI-Anwendungen könnten etwa eingesetzt werden, um Politik stärker reaktions- und erklärungsfähig zu machen, wie es etwa durch Chatbots heute schon im Bereich der Verwaltung versucht wird. Dies könnte zum Beispiel auch in der Unterstützung regelmäßiger deliberativer Verfahren zwischen Politik und diffuser Bürgerschaft münden. KI-Anwendungen könnten hier Beteiligungsumgebungen strukturieren, die Zusammenführung komplexer Debatten erlauben und nachvollziehbar machen, wie bürgerschaftliche Beteiligung wirksam wird.²⁷ Obwohl hier demokratische Potenziale zu erkennen sind, wird zu beobachten sein, inwiefern und von wem neue Beteiligungsmöglichkeiten genutzt werden und ob es sich um „echte“ Beteiligung oder eher um eine Form des Politmarketings handelt.

Aus *Bottom-up*-Perspektive wird in den Blick genommen, wie KI-Anwendungen dazu beitragen, dass neue Formen zivilgesellschaftlicher Selbstorganisation möglich werden oder neue Protest- und Aktionsformen entstehen. Dass Digitalisierung transformative Effekte in Bezug auf politisches Handeln haben kann, ist gut erforscht.²⁸ Aktuell zeichnen sich hier aber noch keine vergleichbaren Impulse für KI-Anwendungen ab, da auf politisches Handeln zugeschnittene Nutzungspraktiken noch wenig ausgeprägt sind und eher auf politische Bildung und Information zielen.

Größer gedachte Vorschläge, die auf eine Umgestaltung des politischen Systems zielen, etwa indem sie den Einsatz von Bots als Agenten poli-

23 Vgl. Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York 2018.

24 Vgl. Frank Pasquale, *The Black Box Society*, Boston 2015.

25 Vgl. Marion Fourcade/Jeffrey Gordon, *Learning Like a State: Statecraft in the Digital Age*, in: *Journal of Law and Political Economy* 1/2020, S. 78–108.

26 Vgl. Ann-Kathrin Koster, *Das Ende des Politischen? Demokratische Politik und Künstliche Intelligenz*, in: *Zeitschrift für Politikwissenschaft* 2/2022, S. 573–594.

27 Vgl. Sarah Kreps/Maurice Jakesch, *Can AI Communication Tools Increase Legislative Responsiveness and Trust in Democratic Institutions?*, in: *Government Information Quarterly* 3/2023, <https://doi.org/10.1016/j.giq.2023.101829>.

28 Vgl. Zeynep Tufekci, *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, New Haven 2017.

tischer Partizipation vorschlagen, sind selten und setzen eine demokratische Struktur voraus, die stärker direkt-demokratisch orientiert ist.²⁹ Wie bei der *Top-down*-Richtung wird auch bezüglich sich entwickelnder *Bottom-up*-Strukturen zudem kritisch zu fragen sein, inwiefern mehr Partizipationsmöglichkeiten gleiche Beteiligung fördern oder ob dadurch nicht eher bestehende Macht- und Ressourcenasymmetrien verstärkt werden.

Neben der Schaffung von Partizipationsmöglichkeiten wird in der wissenschaftlichen Literatur auch über eine Gegentendenz spekuliert, für deren Relevanz es bereits stärkere empirische Anhaltspunkte gibt: nämlich, dass sich demokratische Politik künftig weniger an der expliziten politischen Partizipation als vielmehr am messbaren Verhalten der Bürger*innen orientiert.³⁰ Diese Tendenz ist darauf zurückzuführen, dass in datenreichen Gesellschaften Politik versucht ist, sich stärker durch die Analyse von Bevölkerungseinstellungen zu legitimieren. Für eine solche Entwicklung spricht, dass in komplexen Gesellschaften aufwendige Partizipationsformen häufig als sowohl ineffektiv als auch ungerecht gelten, einfache Partizipationsformen wie periodische Wahlen aber die hohen Beteiligungserwartungen nicht erfüllen. Der Rückgriff auf indirekte Messungen reaktiviert daher ein demokratisches Versprechen, wie es früher bereits demoskopische Verfahren gaben, dass nämlich eine neutrale und umfassende Berücksichtigung der breiten Ansichten und Belange der Bürgerschaft gerade dann stattfindet, wenn man auf Beobachtung statt auf Aktivierung setzt.³¹ KI soll hier eine adäquate und stets aktuelle „Lesbarkeit“ der Bevölkerung herstellen, wobei zu fragen ist, ob damit nicht nur eine technokratische Depolitisierung befördert wird.

AUSBLICK

Die aufgezeigten potenziellen Entwicklungspfade demokratischer Praxis angesichts der zunehmen-

den gesellschaftlichen Einbettung von KI-Verfahren ergeben ein uneinheitliches und ambivalentes Bild. Künstliche Intelligenz ist nicht einfach eine rivalisierende Kraft zu demokratischen Prozessen, wie es populäre Kritiken nahelegen. Die transformative Wirkung der Entwicklung ist aber auch nicht nur begrenzt auf die wachsenden Möglichkeiten effektiver Manipulation. Es ist daher notwendig, die Verschlingung von KI-Technologien und demokratischen Entwicklungen kleinteiliger und mit Blick auf konkrete Dimensionen demokratischen Lebens nachzuvollziehen, um die Rekonfiguration der sich verschiebenden Handlungsmöglichkeiten von individuellen, staatlichen und privatwirtschaftlichen Akteuren besser zu verstehen.

Neben der Analyse schärft sich so der Blick dafür, dass eine aktive politische Gestaltung der digitaltechnologischen Entwicklung möglich und aus demokratischer Sicht sogar naheliegend ist. Anders als in früheren Phasen der Digitalisierung, in denen ein liberaler Imperativ vorherrschte, der eine gesellschaftlich positive Entwicklung unterstellte, wenn man technologische Innovation nur frei walten ließe, ist für KI-Verfahren schon zum jetzigen Zeitpunkt deutlich, dass es eine Notwendigkeit demokratisch sensibler Regulierung gibt. Aktuelle Vorschläge fokussieren dabei hauptsächlich Anwendungsszenarien (etwa Gesichtserkennung oder Deep Fakes) und die Frage der Datengrundlage von KI (Datenschutz und Copyright). Aus demokratischer Sicht muss darüber hinaus aber auch verstärkt auf gesellschaftliche Mitbestimmungsmöglichkeiten geachtet werden – und zwar in Hinblick auf Entwicklung, Anwendung und Einsatz von KI. Dies erschöpft sich nicht in der einfachen Forderung nach Transparenz, sondern meint die Etablierung eingriffsfähiger Institutionen, die in der Lage sind, öffentliche Interessen und bürgerliche und demokratische Rechte auch gegen technische Möglichkeiten und ökonomische Anreize stark zu machen.

²⁹ Vgl. Paul Burgess, *Algorithmic Augmentation of Democracy: Considering Whether Technology Can Enhance the Concepts of Democracy and the Rule of Law through Four Hypotheticals*, in: *AI & Society* 1/2022, S. 97–112.

³⁰ Vgl. Lena Ulbricht, *Scraping the Demos. Digitalization, Web Scraping and the Democratic Project*, in: *Democratization* 3/2020, S. 426–442.

³¹ Vgl. Michel Dormal, *Von Gallup zu Big Data. Rekonstruktion und Neujustierung der Debatte über Meinungsforschung und Demokratie*, in: *Zeitschrift für Politikwissenschaft* 1/2021, S. 1–24.

THORSTEN THIEL

ist Professor für Demokratieförderung und Digitalpolitik an der Universität Erfurt.

thorsten.thiel@uni-erfurt.de

REGULIERUNG VON KI

Ansätze, Ideen, Pläne

Hannah Ruschemeier

Künstliche Intelligenz ist komplex: Sie ist nützlich, und sie ist risikoreich, sie ist menschengemacht, aber für Menschen nicht immer nachvollziehbar. Vor allem ist KI eine Frage von Macht und ihre Regulierung daher eine genuin staatsrechtliche Angelegenheit.

Das Feld der Regulierung von Künstlicher Intelligenz ist divers. Involviert sind verschiedene Institutionen, Regulierungsebenen und Regulierungsgegenstände. Bisher hat kein Land einen umfassenden Rechtsrahmen für KI erlassen, und es existiert kein völkerrechtlicher Vertrag, der einheitliche internationale Vorgaben schafft. KI-Regulierung ist aufgrund der unterschiedlichen Ansätze verschiedener Staaten, Staatenverbände, nicht-staatlicher Organisationen und anderer Institutionen schon aufgrund der divergierenden Kompetenzen eher ein Flickenteppich als ein Puzzle, das sich zusammenfügt. Dennoch lassen sich übergreifende Gemeinsamkeiten, Unterschiede und Entwicklungslinien identifizieren.

WARUM KI-REGULIERUNG?

KI ist auf den ersten Blick kein außergewöhnlicher Regulierungsgegenstand: Insbesondere rechtliche Regulierung hat sich stets mit neuen technischen Entwicklungen, Unsicherheiten oder globalen Auswirkungen befasst, wie das Umwelt- und Technikrecht zeigen. Dennoch wächst das Bewusstsein dafür, dass die bisherigen Regeln auf verschiedenen Ebenen nicht ausreichen, um KI effektiv zu regulieren. Die Gründe dafür liegen im soziotechnischen Charakter und in der technischen Funktionsweise von KI sowie in der Systematik rechtlicher Regulierung. Übergeordnetes Regulierungsziel ist es, die entstehenden Risiken zu umschiffen, um die sich eröffnenden Chancen nutzen zu können.

Recht reguliert nach Kriterien von Subjektivität und Individualität, menschlichen Entscheidungen, Haftung und Verantwortung, Schuld, Erklärbarkeit und Rechtssicherheit. KI hinge-

gen operiert auf der Grundlage großer, zum Teil unübersichtlicher Datenmengen in opaken Entscheidungsstrukturen mit einer Vielzahl an Zwischenschritten und beteiligten Personen in Entwicklung, Vertrieb und Nutzung. KI ist zugleich keine rein technische Entwicklung, sondern eine soziotechnische – KI ist menschengemacht. Generative Modelle erzeugen Sprache, Bilder oder Videos auf der Grundlage von Texteingaben. Dazu gehören Sprachmodelle wie ChatGPT, welche die Wahrscheinlichkeit der Reihenfolge von Wörtern errechnen und dadurch Unterhaltungen simulieren können. Vorhersagemodelle der Wirtschaft, die Scoring-Werte berechnen, *predictive maintenance* (vorausschauende Wartung) in der Industrie oder die Analyse von Umweltdaten in der Forschung: All diese Modelle benötigen große Datenmengen, um trainiert und validiert zu werden. Diese Datengrundlagen wiederum werden von Menschen erzeugt, etwa durch die Nutzung digitaler Medientechnik.⁰¹ KI ist nicht „selbstlernend“, sie braucht den Menschen. Aber braucht der Mensch KI? Und wenn ja, in welchem Maß? Diese Frage ist für Regulierung hoch relevant.

Für Regulierungsfragen sind KI-Systeme umso relevanter, je wahrscheinlicher ihre Verwendung Einfluss auf geschützte Rechtsgüter wie Grundrechte, demokratische Prozesse oder die öffentliche Sicherheit hat. Diese betroffenen Rechtsgüter sind nicht neu und nicht allein durch KI-Anwendungen gefährdet. Dennoch besteht ein nahezu globaler Konsens, dass neue Formen der Regulierung für KI notwendig sind. Allerdings brauchen wir weniger eine ständig neue Formulierung neuer digitaler Rechte als vielmehr effektive Instrumente, die bestehende Rechte und Schutzinteressen mit vorhandenen und neuen Mitteln schützen.

Die von zahlreichen prominenten Stimmen insbesondere aus der Techbranche im Frühjahr 2023 angestoßene Forderung einer Entwicklungspause für KI zur Schaffung eines Regelwerks, ist indes eine marketingmotivierte Nebelkerze.⁰²

Dort wird angeführt, dass KI Bewusstsein erlangen und die menschliche Zivilisation auslöschen könnte. Neben der Ironie, dass zahlreiche Unterzeichner:innen an der Entwicklung eben jener Modelle mitgewirkt haben, dessen Stopp sie nun fordern, kam dieser Aufruf zu einem Zeitpunkt, an dem ein großer Konkurrent, nämlich die Firma OpenAI, mit ChatGPT den Markt aufmischte. Zudem lenken solche Übertreibungen davon ab, dass bereits jetzt erhebliche Risiken durch KI bestehen und eine effektive Regulierung notwendig ist; eine entsprechende Beschränkung des Einsatzes von KI wurde hingegen nicht gefordert.

KI-REGULIERUNG IST MACHTREGULIERUNG

Als Regulierungsziele für KI werden häufig Fairness, Transparenz, Erklärbarkeit, Vertrauenswürdigkeit, Sicherheit, Grundrechtsschutz, Innovationsförderung genannt. Hinter diesen Bestrebungen, neue Regulierungsvorgaben für KI zu schaffen, steht aber mehr als diese zweifelsohne wünschenswerten Einzelziele, die sich primär aus der undurchsichtigen Funktionsweise von KI-Systemen begründen – denn Menschen können den Weg zum Output eines Systems nicht nachvollziehen, erklären oder verstehen. Tatsächlich manifestiert sich in der fortschreitenden Verbreitung von KI-Technologie eine neue Form von Macht. KI-Regulierung ist deshalb die Regulierung von Macht und damit eine rechtsstaatliche Frage, denn staatlich gesetztes Recht legitimiert und begrenzt Macht gleichermaßen.

Die Machtdimension von KI speist sich aus unterschiedlichen Faktoren. Die zentralen Akteure der KI-Technologien, die den primären Anstoß für die rechtspolitische Diskussion über KI-Regulierung geben, sind zunächst global agierende Technologiefirmen. Die Entwicklung und Anwendung von KI ist zwar nicht auf die freie Wirtschaft beschränkt – Open-Source-Initiativen, NGOs, staatliche Institutionen und die Wissenschaft bespielen ebenfalls das KI-Feld –, dennoch sind die Technologien, die im öffentlichen Diskurs debattiert werden, primär solche, die von privatwirtschaft-

lichen Akteuren entwickelt und eingesetzt werden.⁰³ Deshalb ist es wichtig, das Motiv des ökonomischen Profits dieser Akteure transparent zu machen. Auch wenn es keineswegs ausgeschlossen ist, dass viele Menschen von KI profitieren (können) oder dass es gemeinwohlorientierte Einsatzzwecke von KI gibt: Es entstehen Machtstrukturen, wenn Staaten für einen KI-Einsatz zwingend auf private Unternehmen angewiesen sind.

Die Corona-Warn-App beispielsweise konnte nur über die App-Stores von Google und Apple erfolgreich ausgerollt werden.⁰⁴ Des Weiteren verfügen diese privaten Akteure aufgrund ihrer technischen Möglichkeiten und ihrer extrem breiten Nutzer:innenbasis über eine erhebliche Daten- und Vorhersagemacht. ChatGPT funktioniert deswegen so gut, weil bei der Entwicklung nahezu das gesamte Internet nach öffentlich zugänglichen Informationen durchsucht und das Modell auf dieser Grundlage trainiert wurde.⁰⁵ Google sammelt Daten über das Verhalten der Nutzer:innen seiner diversen Dienste, wodurch detaillierte Profile und Vorhersagen über Konsumvorlieben, aber auch höchst sensible Informationen wie Kreditwürdigkeit, sexuelle Orientierung oder Gesundheitszustand erstellt und an Werbekund:innen verkauft werden können.⁰⁶ Meta (vormals Facebook) personalisiert seinen Algorithmus und damit die Anzeige von Inhalten und sammelt Daten in einem Ausmaß, das Nutzer:innen oft nicht bewusst ist.⁰⁷ Diese Firmen schaffen somit „plattformisierte“ Infrastrukturen, die mit einer nahezu unbegrenzten Sammlung von Nutzungsdaten verbunden sind.

Die Allgegenwärtigkeit dieser digitalen Prozesse und die zunehmende Verbreitung von KI haben

03 Siehe nur die Beispiele bei Thomas Ramge, *Mensch fragt, Maschine antwortet*, in: APuZ 6–8/2018, www.bpb.de/263680.

04 Vgl. Kirsten Bock et al., *Das Verfahren geht weit über „die App“ hinaus – Datenschutzfragen von Corona-Tracing-Apps*, in: *Informatik Spektrum* 5/2020, S. 334–338.

05 Vgl. Michael Schade, *How ChatGPT and Our Language Models Are Developed*, <https://help.openai.com/en/articles/7842364>; zur datenschutzrechtlichen Einschätzung Hannah Ruschemeier, *Squaring the Circle*, 7.4.2023, <https://verfassungsblog.de/squaring-the-circle>.

06 Vgl. Renée Ridgway, *Deleterious Consequences: How Google's Original Sociotechnical Affordances Ultimately Shaped „Trusted Users“ in Surveillance Capitalism*, in: *Big Data & Society* 1/2023, S. 1–17; Shoshana Zuboff, *Das Zeitalter des Überwachungskapitalismus*, Frankfurt/M.–New York 2018, S. 85 ff.

07 Vgl. Patricia Arias-Cabarcos/Saina Khalili/Thorsten Strufe, *„Surprised, Shocked, Worried“: User Reactions to Facebook Data Collection from Third Parties*, in: *Privacy Enhancing Technologies Symposium 2023*.

01 Vgl. Rainer Mühlhoff, *Die Macht der Daten. Warum künstliche Intelligenz eine Frage der Ethik ist*, Göttingen 2023, S. 11 ff.

02 Vgl. Sebastian Schreiber, *Experten fordern Pause bei KI-Entwicklung*, 29.3.2023, www.tagesschau.de/wissen/musk-tech-pause-ki-entwicklung-101.html.

zudem erkenntnistheoretische Auswirkungen: Wie werden Entscheidungen getroffen? Wie wird Wissen erzeugt? Sprachmodelle liefern eloquent klingende Antworten, produzieren aber auch viel Unsinn. Durch eine flächendeckende Nutzung solcher Systeme, beispielsweise im Falle einer Integration in alle Office-Anwendungen von Microsoft, wird es immer schwieriger, Phänomenen wie dem sogenannten *automation bias* (übermäßiges Vertrauen in maschinelle Entscheidungsvorschläge) zu entkommen.⁰⁸ Die Geschäftsmodelle, Strukturen, soziotechnischen Interaktionen und nicht zuletzt ihre weite Verbreitung führen dazu, dass bisherige Regulierungsansätze nicht mehr in allen Fällen effektiv greifen. Beispielsweise wird im europäischen Digital Services Act (DSA, deutsch: Gesetz über digitale Dienste), der erst im November 2022 in Kraft getreten ist, KI nicht ausdrücklich genannt – die Verordnung soll aber ein sicheres und vertrauenswürdiges Online-Umfeld schaffen, das durch die Funktionsweise digitaler Plattformen gefährdet ist. Das betrifft somit auch die Nutzung von KI, etwa zur Anzeige und Moderation von Inhalten.

Um den vielfältigen Machtfaktoren von KI zu begegnen und zugleich Innovationen zu ermöglichen, braucht es rechtliche Regelungen – denn wo Macht im Spiel ist, zeigt sich das Potenzial gesellschaftlicher Verbesserungen genauso wie die Gefahr des Missbrauchs. Nach dem Vorsorgeprinzip können bestimmte, besonders riskante Produkte und Verfahren, wenn sie wichtige Rechtsgüter bedrohen, einer gesetzlichen Regulierung unterworfen werden, um im Vorfeld Schäden zu verhindern. In Bezug auf KI stellt sich allerdings die Herausforderung, dass einige Auswirkungen schlecht, noch nicht oder gar nicht abschätzbar sind.

REGULIERUNGSFORMEN

Regulierung von KI hat verschiedene Formen: Klassische *rechtliche Regulierung* kann präventive Verbote, repressive Sanktionen oder Handlungsgebote formulieren, bestehende Vorschriften anwenden oder neue schaffen; *ethische Vorschläge* beziehen sich auf moralische Vorgaben, die aber zur Grundlage rechtlicher Regulierung werden können; *technische Vorgaben* wie Standardisierungsnormen erzeugen oft faktische Bindungen.

KI ist eine Querschnittstechnologie, die in unterschiedlichsten Bereichen eingesetzt wird und dadurch auch verschiedene Rechtsgebiete berührt: Produktsicherheitsrecht, Verbraucherschutzrecht, Urheberrecht, Datenschutzrecht, Grundrechtsschutz, privates Haftungsrecht, strafrechtliche Zurechnungsfragen, Arbeitsrecht und weitere. KI wird keinesfalls in einem rechtsfreien Raum verwendet, der nun dringend neue Detailregelungen erfordert. Es muss deshalb, einerseits, sorgfältig geprüft werden, ob neue Rechtsvorgaben wirklich notwendig sind. Andererseits sollten Regulierungslücken geschlossen werden, wenn Regulierungsziele nicht mehr erreicht werden oder neu entstehen. Wichtig ist es zudem, die soziotechnische Funktionsweise und den stets menschengemachten Anwendungskontext von KI zu reflektieren und keine rein technikzentrierte Regulierung zu schaffen.

In den vergangenen Jahren sind weltweit eine Vielzahl von Initiativen entstanden, um Werte und Grundsätze für die ethische Entwicklung und den Einsatz von KI zu definieren.⁰⁹ Unverbindliche Empfehlungen, Strategiepapiere, *soft law* wie Leitlinien oder Absichtserklärungen sowie ethische Prinzipien sehen sich oft der Kritik ausgesetzt, ineffektiv zu sein, da sie eben nicht bindend sind und daher nicht zwangsweise durchgesetzt werden können. Im Feld der KI-Regulierung trifft dies zusammen mit der besonderen Machtstellung weniger zentraler Akteure, die sich dadurch Regulierung entziehen können. Bisher hat der private Sektor sämtliche Vorschläge weitestgehend ignoriert und aggressive Lobbyarbeit betrieben; dennoch wächst die Anzahl der Forschungsbeiträge zu den Themen Fairness, Verantwortung und Transparenz ins Unüberschaubare.¹⁰

Nicht nur im Bereich von KI beeinflussen sich rechtsverbindliche Vorgaben und ethische Vorschläge gegenseitig als unterschiedliche Dimensionen von Normativität: Ethische Normen orientieren sich an rechtlicher Systematik, Recht übersetzt Ethik in vollziehbare Vorgaben. Die von der Europäischen Kommission eingesetzte unabhängige Hochrangige Expertengruppe für KI (HEG-KI) zum Beispiel hat 2019 in ihren Ethik-Leitlinien für eine vertrauenswürdige KI drei zentrale Kriterien

⁰⁸ Vgl. Hannah Ruschemeier, *The Problems of the Automation Bias in the Public Sector*, in: Weizenbaum Conference Proceedings 2023 (i.E.).

⁰⁹ Vgl. Brent Mittelstadt, *Principles Alone Cannot Guarantee Ethical AI*, in: *Nature Machine Intelligence* 1/2019, S. 501–507.

¹⁰ Vgl. Jack Clark/Ray Perrault et al., *The AI Index 2023 Annual Report*, Stanford 2023, S. 158 ff.

en formuliert, die alle KI-Systeme erfüllen sollten: Rechtmäßigkeit, Einhaltung ethischer Grundsätze und Robustheit. Auf nationaler Ebene hat die Datenethikkommission der Bundesregierung in ihrem Gutachten zu algorithmischen Systemen im selben Jahr einen risikoadaptierten Regulierungsansatz vorgeschlagen – also einen, bei der die Regulierungstiefe am Schädigungspotenzial ausgerichtet ist – der nun auf europäischer Ebene in ähnlicher Form umgesetzt wird.¹¹ Darüber hinaus gibt es zunehmend Vorschläge, schwer fassbare Werte wie Fairness über Metriken und Schwellenwerte abzubilden, um sie umsetzbar zu machen.¹² Die Übersetzbarkeit normativer Wertungen in numerische Maßstäbe aber ist von vornherein begrenzt.

EUROPÄISCHE KI-REGULIERUNG

In der EU werden seit Längerem verschiedene Regulierungsansätze diskutiert. Bereits 2018 wurde eine europäische KI-Strategie veröffentlicht sowie die HEG-KI eingesetzt. 2020 veröffentlichte die EU-Kommission das KI-Weißbuch, gefolgt von einer öffentlichen Konsultation.¹³ Im April 2021 legte die Kommission schließlich einen ersten umfassenden Verordnungsentwurf zur Regulierung von KI vor (KI-VO-E).¹⁴ Im laufenden Gesetzgebungsverfahren ist noch vieles an diesem „AI Act“ umstritten, Einigkeit besteht aber über die Systematik eines risikobasierten Regulierungsansatzes und die Schaffung eines horizontalen Rechtsrahmens, der für alle KI-Anwendungen gelten soll, nicht nur für einzelne Sektoren.

KI-Systeme sollen in drei Kategorien klassifiziert werden: unannehmbares Risiko, hohes Risiko und geringes Risiko. Systeme mit minimalem Risiko wie Spamfilter für E-Mails sind kein Gegenstand der Regulierung. Systeme mit unannehmbarem Risiko wie manipulative KI, Social Scoring und biometrische Fernidentifizierung sind verbo-

ten, wobei es für Letztere weitgehende Ausnahmen für Justiz- und Strafverfolgungsbehörden gibt.

Der Schwerpunkt der Verordnung liegt auf Hochrisikosystemen. Zu den Anwendungsfeldern in diesem Bereich zählen zurzeit die Fallgruppen Biometrik, Bildung, Beschäftigung, Inanspruchnahme grundlegender privater und öffentlicher Dienste und Leistungen, Strafverfolgung, Migration, Asyl und Grenzkontrolle sowie Rechtspflege und demokratische Prozesse. Der KI-VO-E legt den Fokus auf die Hersteller:innen von KI, Nutzer:innen scheinen eine untergeordnete Rolle zu spielen. Es fehlen Betroffenenrechte oder Möglichkeiten zur öffentlichen Partizipation.¹⁵ Um die Überwachung der Verpflichtungen sicherzustellen, nimmt der Entwurf primär die Anbieter:innen des Systems in die Pflicht.

Der Entwurf sieht zahlreiche allgemeine und spezielle Transparenzanforderungen vor. KI-Systeme mit hohem Risiko sind generell so zu konzipieren und zu entwickeln, dass ihre Funktionsweise hinreichend transparent ist, um Anbieter:innen und Nutzer:innen ein angemessenes Verständnis der Funktionsweise des Systems zu ermöglichen. Systeme, die mit natürlichen Personen interagieren, und Emotionserkennungssysteme müssen darüber informieren, dass KI eingesetzt wird. Deep Fakes, also manipulierte Bild- oder Toninhalte, müssen gekennzeichnet werden. Was ein „angemessenes Verständnis“ und was generell unter einer „ethischen und vertrauenswürdigen“ KI zu verstehen ist, wird im Normtext jedoch nicht beantwortet. Dies hängt auch mit der Rolle der Normierungsinstitutionen zusammen. Zentral ist hier die Zertifizierung von KI: Die Anforderungen an Hochrisikosysteme scheinen detailliert, ihre konkreten Maßstäbe werden aber maßgeblich auf harmonisierte Normen übertragen (etwa DIN-Normen), die von privaten Standardisierungsorganisationen erlassen werden.¹⁶ Obwohl nicht demokratisch legitimiert, kommt diesen Institutionen damit eine große Macht bei der Regulierungsgestaltung zu.

Zur Umsetzung des KI-VO-E sollen die Mitgliedstaaten neue Aufsichtsbehörden schaffen

11 Siehe <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> und www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html.

12 Vgl. Ninareh Mehrabi et al., A Survey on Bias and Fairness in Machine Learning, in: *ACM Computing Surveys* 6/2021, Artikel 115, S. 1–35.

13 Vgl. Europäische Kommission, White Paper on Artificial Intelligence, 17.7.2020, <https://digital-strategy.ec.europa.eu/en/library/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence-and>.

14 Siehe <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0206>.

15 Kritisch dazu Nathalie Smuha et al., How the EU Can Achieve Legally Trustworthy AI, 5.8.2021, <https://ssrn.com/abstract=3899991>.

16 Vgl. Johann Laxu/Sandra Wachter/Brent Mittelstadt, Three Pathways for Standardisation and Ethical Disclosure by Default under the European Union Artificial Intelligence Act, 20.2.2023, <https://ssrn.com/abstract=4365079>.

oder die Verordnungsaufsicht an bestehende Institutionen delegieren. Die genauen Anforderungen, insbesondere die einer vollständigen Unabhängigkeit wie aus dem Datenschutzrecht bekannt, sind Gegenstand laufender Verhandlungen. Als neues regulatorisches Instrument auf Unionsebene sieht der KI-VO-E Reallabore vor (*regulatory sandboxes*), in denen innovative KI-Anwendungen in Zusammenarbeit mit der Aufsichtsbehörde für einen bestimmten Zeitraum getestet werden sollen.

Neben der geplanten KI-Verordnung gibt es bereits Daten- und plattformbezogene EU-Gesetzgebung, die auch KI betrifft. Zentrales Regelwerk für die Verarbeitung personenbezogener Daten ist seit 2018 die europäische Datenschutzgrundverordnung (DSGVO), die zweifelsohne auch für KI gilt, wenn personenbezogene Daten verarbeitet werden. Hier bestehen allerdings ungelöste Probleme zwischen der regulatorischen Ausrichtung der DSGVO – etwa mit Blick auf Datenminimierung und Betroffenenrechte – und der Differenzierung zwischen verschiedenen Datenkategorien sowie der praktischen Durchsetzung.¹⁷ Bisher gibt es kein eigenes explizites Regelwerk für KI-Trainingsdaten.

Auch die wichtigen Verordnungen des DSA und des Digital Markets Act (DMA) werden sich mittelbar auf den Einsatz von KI auswirken, insbesondere bei sozialen Plattformen. Der DSA fordert von sehr großen Anbieter:innen, systemische Risiken zu überwachen, schafft Transparenzvorgaben für Empfehlungssysteme und verbietet personenbezogene Werbung gegenüber Minderjährigen.¹⁸ Der DMA soll den Wettbewerb der Digitalwirtschaft in der EU schützen und reguliert besonders mächtige „Torwächter“-Firmen, die oft zugleich wichtige Akteure im KI-Bereich sind.

WEITERE KI-REGULIERUNGEN

Verschiedene Staaten setzen unterschiedliche Schwerpunkte bei der Regulierung von KI; heraus-

17 Vgl. Rainer Mühlhoff/Hannah Ruschemeier, Predictive Analytics und DSGVO: Ethische und rechtliche Implikationen, in: Hans-Christian Gräfe/Telemedicus e. V. (Hrsg.), Telemedicus – Recht der Informationsgesellschaft, Frankfurt/M. 2022, S. 38–67; Tal Zarsky, Incompatible: The GDPR in the Age of Big Data, in: Seton Hall Law Review 4/2017, S. 995–1020.

18 Weitergehend dazu Hannah Ruschemeier, Die aktuelle Digitalgesetzgebung der Europäischen Union – eine kritische Analyse, in: Zeitschrift für Gesetzgebung 2023 (i. E.).

gegriffen werden hier die USA und China, die oft als gegensätzliche Systeme zum europäischen Ansatz genannt werden.¹⁹ Den USA wurde lange ein gewisser Laissez-faire-Ansatz mit (zu) großem Vertrauen in die Selbstregulierung der Industrie nachgesagt, inzwischen sind jedoch zahlreiche politische Initiativen insbesondere im Hinblick auf Verbraucherschutz und gegen Monopolisierung angestoßen worden. Die sogenannte AI Bill of Rights,²⁰ ein nicht bindendes Prinzipienpapier der aktuellen US-Regierung, ist eher darauf bedacht, die Entwicklung nationaler, branchenspezifischer Normen zu fördern als internationale Normen zu beeinflussen.

China hat erhebliche Investitionen in KI getätigt und fördert aktiv die Entwicklung von KI-Technologien und deren Einsatz in allen Wirtschaftssektoren. Chinas Regulierungsansatz ist zum einen stark innovationsorientiert, mit Initiativen zur Förderung spezifischer Entwicklungen auf allen Ebenen der Regierung und der Industrie, und zum anderen gesellschaftlich orientiert, wobei die „soziale Stabilität“ (im Verständnis der Zentralregierung) Vorrang vor den Rechten der einzelnen Person hat.²¹ Dennoch sieht der aktuelle Gesetzesentwurf zum Teil schärfere Vorgaben für generative KI vor als der aktuelle Stand europäischer Regulierung, da beispielsweise für das Training generativer Modelle die Einwilligung betroffener Personen eingeholt werden soll – dies würde das Sammeln von Nutzerdaten auf Internetseiten (*scraping*) für Trainingszwecke untersagen.²²

Weitere Länder wie Brasilien und Kanada haben eigene Gesetzesentwürfe zur Regulierung von KI auf den Weg gebracht.²³ Auf der Ebene internationaler Organisationen hat der Europarat

19 Vgl. Liselotte Odgaard, Europe and the US-China Tech War, Robert Schuman Centre for Advanced Studies, RSCAS Policy Paper 2/2021.

20 Kritisch dazu Emmie Hine/Luciano Floridi, The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction, in: Minds and Machines 2/2023, S. 285–292.

21 Vgl. Huw Roberts et al., Governing Artificial Intelligence in China and the European Union: Comparing Aims and Promoting Ethical Outcomes, in: The Information Society 2/2023, S. 79–97.

22 Vgl. Interim Measures for the Management of Generative AI Services, Order No. 15 of the Cyberspace Administration of China et al., 10.7.2023, <https://perma.cc/X2WX-E592> (inoffizielle Übersetzung).

23 Vgl. José Renato Laranjeira de Pereira/Thiago Guimarães Moraes, Promoting Irresponsible AI: Lessons From a Brazilian Bill, 14.2.2022, <https://eu.boell.org/en/2022/02/14/promoting-irresponsible-ai-lessons-brazilian-bill>; Parliament of Canada, C-27, Digital Charter Implementation Act, 2022, www.parl.ca/LegisInfo/en/bill/44-1/C-27.

Verhandlungen über eine KI-Konvention aufgenommen, die Staaten dazu verpflichten soll, bei Entwicklung und Einsatz von KI keine Menschenrechte zu verletzen.²⁴

SCHWIERIGKEITEN

KI als Regulierungsobjekt direkt zu behandeln, ist komplex, da die Auffassungen davon, was KI eigentlich ist, nach wie vor auseinandergehen.²⁵ Im ursprünglichen Bereich der KI, Informatik und Mathematik, wird der Begriff meist als Oberbegriff für verschiedene Anwendungen genutzt. Normative Regulierung hat zudem eine andere Perspektive als ein rein technisches Verständnis: Wichtig sind der Einsatzkontext und die dadurch betroffenen Schutzgüter und Interessen. Deswegen benennt nicht jede Regulierung mit KI-Bezug die Technologie unmittelbar – so etwa der DSA. Der KI-VO-E hingegen adressiert KI-Systeme explizit, definiert diese in der ersten Fassung aber so breit, dass praktisch jede Software darunterfällt.

Kritik gibt es auch an der Verwendung des Begriffs „Intelligenz“, denn nicht einmal menschliche Intelligenz ist klar definiert. Wie kann es dann künstliche sein? Aus rechtlich-regulatorischer Sicht ist die Definition des Regelungsgegenstands essenziell, da sie den Anwendungsbereich der Regulierung festlegt. Zudem müssen Anforderungen an Rechtssicherheit, Präzision und Praktikabilität erfüllt sein. Aufgrund des breiten Spektrums an gesellschaftlichen Segmenten und an Wissenschaften, die entweder direkt oder indirekt von KI betroffen sind, führt jedoch jede Sichtweise zu einer eigenen Definition dessen, was KI ist und was sie für den jeweiligen Bereich bedeutet. Die Tatsache, dass weder die Computerwissenschaft noch die Informatik im KI-VO-E direkt erwähnt werden, zeigt, dass es keine allgemein anerkannte technische Definition dessen gibt, was KI ist oder sein könnte.

Im weiteren Gesetzgebungsverfahren des KI-VO-E wurde die Definition mehrfach verändert;

²⁴ Siehe Committee on Artificial Intelligence des Europarates: www.coe.int/en/web/artificial-intelligence/cai.

²⁵ Vgl. Hannah Ruschemeier, AI as a Challenge for Legal Regulation – The Scope of Application of the Artificial Intelligence Act Proposal, in: ERA Forum 23/2023, S. 361–376.

²⁶ Abänderung 165 des Europäischen Parlaments zu dem Vorschlag für ein Gesetz über Künstliche Intelligenz, 14. 6. 2023, www.europarl.europa.eu/doceo/document/TA-9-2023-0236_DE.html.

der Vorschlag des Europäischen Parlaments von Juni 2023 lautet: „[Ein KI-System ist] ein maschinengestütztes System, das so konzipiert ist, dass es mit unterschiedlichem Grad an Autonomie operieren kann und das für explizite oder implizite Ziele Ergebnisse wie Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das physische oder virtuelle Umfeld beeinflussen.“²⁶ Auch diese Definition wirft Folgefragen auf, etwa was Autonomie bedeutet. Das Risikoprofil von KI-Systemen lässt sich daher nur aus dem Zusammenspiel zwischen der technischen Funktionalität und dem Anwendungskontext bestimmen. Deshalb ist der regulatorische Filter des KI-VO-E auch nicht die Definition von KI, sondern die Risikoeinstufung. KI-Regulierung funktioniert nur kontextbezogen.

AUSBLICK

Die populären generativen KI-Anwendungen wie ChatGPT, Modelle wie DALL-E oder Stable Diffusion, die aus Texteingaben Bilder generieren, stellen die Regulierung von KI vor weitere Herausforderungen. Denn diese Modelle können zu unterschiedlichsten Zwecken eingesetzt werden, deren Auswirkungen im Vorfeld nicht überschaubar sind. Die erforderliche Risikoabschätzung verkompliziert sich dadurch. Der Fokus sollte dabei nicht nur auf dem Regulierten, sondern auch auf den Regulierenden liegen: Der Aufbau von Expertise und regulatorischem Wissen ist dringend erforderlich. Ein Zusammenführen verschiedener Disziplinen und ein Dialog mit der Praxis sind unumgänglich. Risikobasierte Regulierung kann nicht alle Probleme lösen, nicht alle Auswirkungen lassen sich vorhersehen oder abschätzen. Die Machtstellung der beteiligten Akteure, auch staatlicher Institutionen, sowie die soziotechnischen Zusammenhänge und Auswirkungen sollten nicht aus dem Blick geraten. Rein technikbezogene schematische Lösungen werden der Komplexität von KI, Menschen, Gesellschaften und normativen Systemen nicht gerecht. Mehr Bildung, Digitalkompetenz und eine lebhaftere öffentliche Debatte sind gefordert.

HANNAH RUSCHEMEIER

ist Juniorprofessorin für Öffentliches Recht mit Schwerpunkt Datenschutzrecht & Recht der Digitalisierung an der Fernuniversität in Hagen und Vorstandsmitglied der Robotics and AI Law Society. hannah.ruscheimer@fernuni-hagen.de

KI IN DER ARBEITSWELT

Christian Kellermann · Cornelius Markert

Seit dem Beginn der Industriellen Revolution vor rund 250 Jahren befinden wir uns in einer globalhistorischen Transformation unserer Arbeits- und Lebensweise. Technische Basis des enormen Produktivitätsgewinns war die Reorganisation menschlicher Arbeit durch Maschineneinsatz, vor allem in materiellen Produktionsprozessen durch den Einsatz von Arbeits- und Kraftmaschinen. Die Charakterisierung der Industrialisierung als Beginn des „Maschinenzeitalters“ trifft den Kern dieser Entwicklung.⁰¹

Die kognitive Arbeit war davon lange ausgeschlossen. Mit der Digitalisierung hat sich dies geändert: Wir befinden uns heute am Beginn des „Zweiten Maschinenzeitalters“.⁰² Die Geschwindigkeit, mit der die digitale Technik unser Leben durchdringt, und ihre Leistungszuwächse sind beeindruckend; das Potenzial für Produktivitätsgewinne auch in kognitiven Arbeitskontexten ist sehr groß. Generative Künstliche Intelligenz spielt eine zunehmend bedeutende Rolle in diesem Prozess, da sie selbstständig neuartige Inhalte und Lösungen generieren kann und damit neue Arbeitsfelder einer (Teil-)Automatisierung zugänglich macht.

Allerdings gibt es eine Diskrepanz zwischen den weiter wachsenden technischen Möglichkeiten der Digitalisierung und dem Verständnis, sie produktiv und im Sinne des Allgemeinwohls einzusetzen. Wir sind noch mitten im Prozess, die Chancen und Risiken der Technik zu verstehen und in der Arbeitswelt nutzbar zu machen. Die Auswirkungen auf die Verteilung von Arbeit, Zeit, Einkommen und die Qualität der Arbeit sind unsicher und hängen stark von politischen Entscheidungen ab. Es ist daher von zentraler Bedeutung, wie Politik und Wirtschaft die Entwicklung steuern und sicherstellen, dass die Vorteile der KI gerecht verteilt werden und die Qualität der Arbeit steigt.

GESCHICHTE UND SYSTEMATIK DER (TEIL-)AUTOMATISIERUNG

Um das Potenzial des KI-Einsatzes in der Arbeitswelt einschätzen zu können, ist es unabding-

bar, die aktuellen Entwicklungen in den größeren Zusammenhang der vergangenen 200 Jahre zu stellen und die ökonomischen Rahmenbedingungen zu betrachten. Seit 1800 hat sich in Deutschland, den USA und im Vereinigten Königreich die Gütermenge, die in einer Arbeitsstunde produziert wird, in etwa verdreißigfacht.⁰³ Anders ausgedrückt: Das Bruttoinlandsprodukt in diesen Ländern stieg je Arbeitsstunde von etwa 2,2 auf 76 US-Dollar (in konstanten Preisen von 2020 und zu Kaufkraftparitäten). Die Ausweitung des produktiven Potenzials menschlicher Arbeit war die Basis für Wohlstand und eine massive Arbeitszeitverkürzung. Der materielle Wohlstand stieg von weniger als 5000 Dollar Jahreseinkommen um 1800 auf heute 50 000 bis 70 000 Dollar pro Person.⁰⁴ Die jährliche Erwerbsarbeitszeit sank von über 3000 Stunden im 19. Jahrhundert auf heute noch 1588 Stunden.⁰⁵

Nach dem Zweiten Weltkrieg hatte es zunächst eine rasante Expansionsphase gegeben, ab 1970 flachten die gesamtwirtschaftlichen Zuwachsraten jedoch wieder ab. Dies war unter anderem dadurch bedingt, dass die Technik des ersten Maschinenzeitalters vor allem im Bereich der materiellen Produktion anwendbar war, während die Automatisierung in vielen anderen Sektoren nicht möglich erschien, die deshalb ein geringeres Produktivitätswachstum aufwies.⁰⁶ Nicht zuletzt die aufsehenerregenden Entwicklungen im Bereich der KI belegen: Mit der Digitalisierung gilt dies immer weniger. Die volle Entfaltung der Potenziale der Informations- und Kommunikationstechnik liegt allerdings noch vor uns.

General purpose technologies (GPTs) wie KI benötigen eine gewisse Anlaufphase, bevor sie produktiv eingesetzt werden, da sie erhebliche ergänzende Investitionen, die Entwicklung komplementärer Technik und die Reorganisation von Prozessen, Betrieben und Geschäftsmodellen erfordern.⁰⁷ Die Dampfmaschine als eine GPT des ersten Maschinenzeitalters hat fast 200 Jahre gebraucht, um von ersten Tüfteleien zum flächendeckenden produktiven Einsatz zu gelangen: Erst

nach 1830 und damit 120 Jahre nach der ersten Anwendung der Dampfmaschine in englischen Kohleminen löste die Dampfkraft die Wasserkraft als wichtigste Energiequelle in Großbritannien ab – im Mutterland der Industrialisierung und dem Land mit den geringsten Energiekosten in Europa zu dieser Zeit.⁰⁸ Ganz anders die Technik der Digitalisierung: Weltweit besitzen inzwischen rund 4,7 Milliarden Menschen ein Smartphone⁰⁹ – dabei ist die Vorstellung des iPhone durch Apple gerade einmal 16 Jahre her.

Auch hinsichtlich der Geschwindigkeit und Höhe der Leistungszuwächse unterscheiden sich die datenverarbeitenden Maschinen um mehrere Größenordnungen von den Arbeits- und Kraftmaschinen des ersten Maschinenzeitalters. Ein Jahrhundert nach der Einführung der ersten Web- und Spinnmaschinen hatte sich die Produktivität gegenüber der Handarbeit um etwas mehr als das Hundertfache gesteigert. In einer Arbeitsstunde konnten nun 1,2 Kilogramm Garn gesponnen werden anstelle von 8,1 Gramm in der Handarbeit, und 3,8 Meter statt 3 Zentimeter Stoff konnten gewebt werden. Bis heute, also etwas über 80 Jahre nach Konrad Zuses Rechner „Z2“, hat die Rechenleistung des Computers hingegen um den Faktor 1,5 Billionen zugenommen, was ebenfalls die außergewöhnliche Dynamik der technischen Entwicklung im zweiten Maschinenzeitalter zeigt.¹⁰

Im verarbeitenden Gewerbe, dem Kernsektor des ersten Maschinenzeitalters, stieg die Produktivität infolge des massenhaften Ersatzes

menschlicher Arbeit durch Maschinen seit 1800 um etwa das Zweihundertfache. Aufgabe für Aufgabe, Tätigkeit für Tätigkeit, Beruf für Beruf wurde die Arbeit entlang der neuen technischen Möglichkeiten der Arbeits- und Kraftmaschinen reorganisiert, verstärkt seit etwa 1970/80 durch den wachsenden Druck zur Reorganisation durch die Finanzialisierung. Arbeitsmaschinen wie Drehbänke, Stanzen, Bohrer, Sägen, Hobel oder Pressen ersetzen große Teile manueller Arbeit. Kraftmaschinen wie zunächst Wind- und Wasserräder sowie die Dampfmaschine, später Elektro- und Verbrennungsmotoren, übernahmen den Antrieb der Maschinen. Hinzu kamen die Computer, die Teile der Kopfarbeit – Rechnen, Steuern, Kommunizieren – ersetzen und die produktiven Möglichkeiten weiter vergrößerten.

Diese atemberaubende Produktivitätssteigerung führte nicht zu massenhafter, gesamtwirtschaftlicher Arbeitslosigkeit, wie zum Beispiel der Ökonom John Maynard Keynes befürchtete. Im Gegenteil: Die Beschäftigung in Deutschland hat in den vergangenen Jahren immer neue Rekordmarken erreicht, sowohl was die absolute Zahl der Erwerbstätigen als auch was den Anteil der Erwerbstätigen an der Gesamtbevölkerung betrifft. Die Nachfrage und damit die Arbeit haben sich in andere Sektoren verlagert, Tätigkeitsbilder verändern sich bis hin zum Wegfall alter und zur Entstehung neuer Berufe.

Die *Abbildung* zeigt die Entwicklung der Bruttowertschöpfung pro Arbeitsstunde im ver-

01 Vgl. Institut für die Geschichte und Zukunft der Arbeit (IGZA) (Hrsg.), *Matrix der Arbeit*, Bd. 3: Kapitalistische Marktwirtschaft und Produktionsweise sowie Bd. 4: Zukunft der Arbeit, Bonn 2023.

02 Andrew McAfee/Erik Brynjolfsson, *The Second Machine Age*, New York 2014.

03 Vgl. IGZA (Anm. 1), Bd. 3, S. 55.

04 Bei allerdings seit den 1970er Jahren immer ungleicherer Verteilung der Einkommenszuwächse innerhalb der Gesellschaften. Vgl. Branko Milanovic, *Global Inequality: A New Approach for the Age of Globalization*, Cambridge MA 2016; Thomas Piketty/Emmanuel Saez/Gabriel Zucman, *Distributional National Accounts: Methods and Estimates for the United States*, in: *Quarterly Journal of Economics* 2/2018, S. 553–609.

05 Erwerbsarbeitszeit in Deutschland 2022, bezogen auf Vollzeitbeschäftigte. Inklusive der Teilzeit lag der Wert bei 1267 Stunden. Vgl. Institut für Arbeitsmarkt- und Berufsforschung, IAB-Arbeitszeitrechnung 2023, <https://iab.de/daten/iab-arbeitszeitrechnung>. Wäre die Arbeitszeit auf dem Niveau von 1800 verblieben, wäre das BIP je Erwerbstätigem heute fast doppelt so hoch.

06 Vgl. hierzu die These von der Kostenexplosion der Dienstleistungen bei William J. Baumol/William G. Bowen, *Performing Arts: The Economic Dilemma*, Cambridge 1966; William J. Baumol, *The Cost Disease. Why Computers Get Cheaper and Health Care Doesn't*, New Haven 2012.

07 Aktuelle Forschungen kommen zum Ergebnis, dass die unternehmerischen Transformationsanstrengungen dazu führen, dass in der ersten Phase einer neuen GPT die Produktivitätseffekte in der amtlichen Statistik untererfasst werden, da die Ergebnisse zunächst vor allem immaterieller Natur sind (Patente, Geschäftsmodelle u. a. m.), sodass das in der US-Statistik Ende 2017 angegebene Produktivitätsniveau das tatsächliche Niveau um 15,9 Prozent unterschätzt. Vgl. Erik Brynjolfsson/Daniel Rock/Chad Syverson, *The Productivity J-Curve: How Intangibles Complement General Purpose Technologies*, in: *American Economic Journal: Macroeconomics* 1/2021, S. 333–372.

08 Vgl. Robert C. Allen, *The British Industrial Revolution in Global Perspective*, Cambridge 2009.

09 Vgl. Statista, *Smartphones – Statistiken und Studien*, 14. 9. 2023, <https://de.statista.com/themen/581>.

10 Vgl. IGZA (Anm. 1), Bd. 7: *Datenhandbuch*, S. 131 ff.

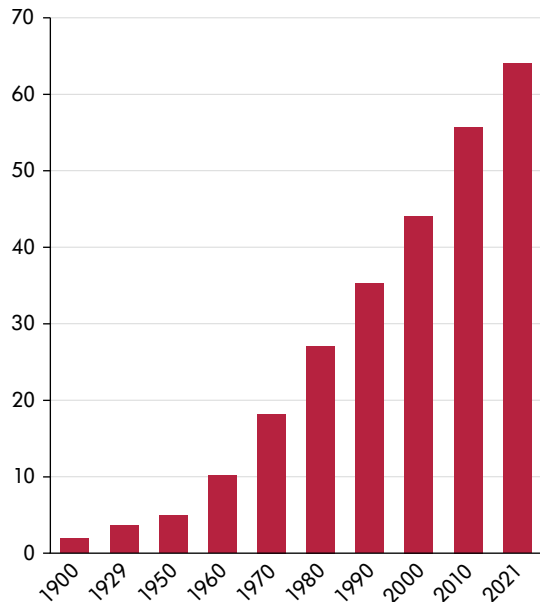
arbeitenden Gewerbe in Deutschland von 1900 bis 2021. In diesen 121 Jahren stieg sie von unter 2 Euro auf 64 Euro, seit 1960 um durchschnittlich jeweils knapp 10 Euro pro Jahrzehnt (Preise von 2022). Dabei stieg der Automatisierungsgrad immer weiter an. War während der 1950er Jahre Handarbeit mit Werkzeugen und Einzweckmaschinen in vielen Fabriken noch weit verbreitet, sind heute in Bereichen wie Presswerken Automatisierungsgrade jenseits der 90 Prozent keine Seltenheit.¹¹ Die Wachstumsraten sind heute zwar niedriger als während des Nachkriegsbooms, dennoch wachsen Automatisierungsgrad und die Produktivität stetig weiter. Durch die Digitalisierung erschließen sich hier zusätzliche Potenziale. Übernahmen im ersten Maschinenzeitalter die Maschinen vor allem physische Tätigkeiten, können sie mit Computer, Internet und KI-Anwendungen zunehmend auch kognitive Funktionen übernehmen – von der Steuerung und Vernetzung über neue Anwendungen der vorausschauenden Wartung (*predictive maintenance*) bis hin zur Unterstützung von Planung, Konstruktion und Management. Der Endpunkt dieser Entwicklung ist die vollautomatische Fabrik, ein Traum der Automationspioniere der 1950er Jahre, der sich mit der damaligen Technik nicht realisieren ließ.¹² Mit der neuen Technik des zweiten Maschinenzeitalters scheinen die damaligen Hindernisse nun zumindest prinzipiell überwindbar.

Vielen Berufen außerhalb des verarbeitenden Gewerbes fehlten bislang die Maschinen für eine Automatisierung. Dies wandelt sich mit der Digitalisierung, entsprechend besteht das Potenzial für eine vergleichbare Produktivitätsentwicklung auch in diesen Bereichen. Die Aufmerksamkeit, die KI-Anwendungen wie ChatGPT oder Midjourney erhalten, kommt nicht von ungefähr (GPT steht hier für *generative pre-trained transformer*). Sie führen vor Augen, wie einzelne Tätigkeiten menschlicher Arbeit im Bereich des Analysierens und Kommunizierens bis hin zu kreativen Tätigkeiten von Maschinen übernommen werden können. Allerdings ist die wissenschaftliche Datengrundlage zu den heutigen Automatisierungsgraden bislang noch dürftig, weshalb auch

¹¹ Vgl. Ulrich Jürgens, *Automatisierung und Arbeit in der Automobilindustrie*, Berlin 2023 (i.E.).

¹² Vgl. David A. Hounshell, *Automation, Transfer Machinery, and Mass Production in the U.S. Automobile Industry in the Post-World War II Era*, in: *Enterprise & Society* 1/2000, S. 100–138.

Abbildung: Bruttowertschöpfung pro Arbeitsstunde im verarbeitenden Gewerbe in Deutschland (ohne DDR) 1900 bis 2021, in Euro (absolut)



Quelle: IGZA, *Matrix der Arbeit*, Bd. 3: Kapitalistische Marktwirtschaft und Produktionsweise, Bonn 2023, auf Basis von Daten des Statistischen Bundesamtes, Inlandsproduktberechnung – Lange Reihen ab 1970, 2022 sowie The Conference Board, *Total Economy Database*, 2019.

die Aussagekraft der Schätzungen zu zukünftigen Automationsgraden begrenzt ist.¹³

ANFANG VOM ENDE DES PRODUKTIVITÄTSPARADOXES

Trotz der Ausweitung der Möglichkeiten zur Automatisierung von Arbeit durch die Digitalisierung bleibt die Entwicklung der Produktivität bislang hinter den Erwartungen. Als einer der ersten stellte dies der Ökonom Robert Solow in den 1980er Jahren fest.¹⁴ Er hatte bereits 30 Jahre zuvor die zentrale Rolle von Technikentwicklung für Wirtschaftswachstum

¹³ Es fehlt nicht nur an belastbaren Zahlen, die über qualitative Einzelfälle oder Schlaglichtstudien (meist von Beratungsunternehmen) hinausgehen, sondern auch an brauchbaren, aktuellen Systematiken der Erfassung, weil die Forschung dazu, die in den 1950er und 1960er Jahren einen Höhepunkt erlebte, fast komplett eingeschlafen ist.

¹⁴ Vgl. Robert Solow, *We'd Better Watch Out*, in: *New York Times Book Review*, 12.7.1987, S. 36.

gezeigt.¹⁵ Das nach ihm benannte „Solow-Paradox“ beschreibt bis heute die scheinbare Entkopplung von Digitalisierung und Produktivitätswachstum und wirft einen langen Schatten auf die entsprechende ökonomische Auseinandersetzung. Mangelnde Messbarkeit digitaler Inputs oder Verzögerungseffekte zwischen Innovation, Entwicklung und betrieblicher Umsetzung sind einige Erklärungsansätze für Solows ungebrochene Aktualität.

Ausgehend von der wirkmächtigen Studie der Arbeitsforscher Carl Benedict Frey und Michael A. Osborne vor zehn Jahren stand lange Zeit das Szenario drohender Arbeitslosigkeit durch zunehmend menschenähnliche Fähigkeiten insbesondere beim maschinellen Lernen im Fokus der Debatte um die Zukunft der Arbeit.¹⁶ Nachdem diverse nachfolgende Studien zeigen konnten, dass ein solches Szenario unrealistisch ist, beruhigte sich die Debatte wieder – nicht zuletzt, weil sowohl die Entwicklung als auch die Umsetzung von KI zumindest in Deutschland weit hinter den Erwartungen zurückgeblieben ist. Auch wenn die Annahmen für ein solches Szenario methodisch anfechtbar sind, beschreiben sie doch einen möglichen Weg, um Beschäftigungseffekte ökonomisch zu schätzen. So werden sämtliche Tätigkeiten, die einen Beruf ausmachen, anhand ihrer Beschreibungen in Berufsdatenbanken mit den theoretischen Fähigkeiten von KI-Methoden abgeglichen. Auf diese Weise wurde ein hohes Substitutionspotenzial bei Arbeiten identifiziert, die einen großen Anteil an Routinetätigkeiten haben und relativ geringe Qualifikationen erfordern – woraus sich die Erwartungen an Produktivitätsschübe ableiteten.

Die Schätzungen, dass in Industrieländern wie den USA oder Deutschland in den nächsten zwei Jahrzehnten bis zur Hälfte aller Jobs durch KI gefährdet sind, sind seitdem deutlich nach unten korrigiert worden. Neben der Methodenkritik an der analytischen Belastbarkeit von gängigen Tätigkeitsbeschreibungen in den verfügbaren Berufsdatenbanken wuchsen auch die

Zweifel an „disruptiven“ Entwicklungsschritten in der KI-Forschung und deren universeller Anwendbarkeit in Unternehmen. Selbst Routineaufgaben, die nach den einschlägigen Modellierungen nunmehr automatisierbar seien, entpuppten sich in der Praxis als zu komplex, als dass eine KI diese voll übernehmen könnte. Schließlich sind auch einfache Tätigkeiten in komplexe Arbeitsprozesse integriert, die nicht einfach aufgebrochen und neu strukturiert werden können. Kontext ist wesentlich für Arbeit, egal welchen Komplexitätsgrad einzelne Tätigkeiten haben. Entsprechend sind deskriptive Bestandsaufnahmen einzelner Tätigkeiten notwendigerweise unterkomplex und erfassen nicht das kontextspezifische „Arbeitsvermögen“,¹⁷ dessen zugrundeliegendes Wissen nicht ohne Weiteres formalisierbar ist.¹⁸ Stattdessen muss eine Übertragung egal welcher Tätigkeit an ein KI-gesteuertes System so codiert werden, dass der gesamte Kontext repräsentiert ist, was aufwendig und voraussetzungsreich ist.

Relevanter als die Diskussion um Vollautomatisierung durch KI sind die komplementären Assistenzfunktionen datengesteuerter, selbstlernender Systeme. Mensch-Roboter-Kollaboration, Smart Glasses oder 3D-Druck beziehungsweise additive Fertigung sind zunehmend Teil der Arbeitsteilung. Assistenzsysteme verändern Tätigkeiten, ihre Zusammensetzung und Bedeutung für einen Beruf, die vorausgesetzten Qualifikationen sowie den Output von Arbeit. In den vergangenen zehn Jahren ist zu beobachten, dass die Nachfrage nach höherqualifizierten Jobs, die KI „ausgesetzt“ sind, gestiegen ist – relativ zu niedrigqualifizierten Jobs. Eine Erklärung dafür ist genau jene Komplementarität in der Mensch-Maschine-Kollaboration und in den Auswirkungen auf die Produktivität dieser Tätigkeiten.¹⁹

An diesem Punkt setzt die aktuelle Debatte um KI und Arbeit an. Auslöser sind die neueren Entwicklungen bei generativen Sprachmo-

¹⁵ Vgl. ders., Technical Change and the Aggregate Production Function, in: *The Review of Economics and Statistics* 3/1957, S. 312–320.

¹⁶ Vgl. Carl Benedict Frey/Michael A. Osborne, *The Future of Employment: How Susceptible Are Jobs to Computerisation?*, Oxford Martin School Working Paper, September 2013.

¹⁷ Vgl. Sabine Pfeiffer/Anne Suphan, *Der AV-Index. Lebendiges Arbeitsvermögen und Erfahrung als Ressourcen auf dem Weg zu Industrie 4.0*, Universität Hohenheim, Working Paper 1/2015.

¹⁸ Vgl. David H. Autor, *Why are There Still So Many Jobs? The History and Future of Workplace Automation*, in: *Journal of Economic Perspectives* 3/2015, S. 3–30.

¹⁹ Vgl. Norbert Huchler, *Komplementäre Arbeitsgestaltung. Grundrisse eines Konzepts zur Humanisierung der Arbeit mit KI*, in: *Zeitschrift für Arbeitswissenschaft* 6/2022, S. 158–175.

dellen wie ChatGPT. Analog zur Debatte über die Auswirkungen von KI-Methoden auf Arbeit und Beschäftigung vor zehn Jahren sind etliche der veröffentlichten Szenarien zu den Effekten solcher KI-Modelle von der Annahme eines hohen Substitutionspotenzials auf dem Arbeitsmarkt geprägt: In Medienberichten wird auffällig häufig die Zahl von weltweit 300 Millionen Jobverlusten genannt.²⁰ Wie zuvor werden Berufsdatenbanken und deren Tätigkeitsbeschreibungen als Datengrundlage für einen Abgleich für die Anwendungsmöglichkeiten der generativen KI herangezogen. Eine andere Berechnung kommt zu dem Schluss, dass rund 80 Prozent der Arbeitskräfte in den USA zu mindestens 10 Prozent ihrer Arbeitsaufgaben von der Einführung von generativer KI betroffen sein werden. Bei etwa 19 Prozent der Beschäftigten könnten bei mindestens 50 Prozent ihrer Aufgaben große Sprachmodelle zum Einsatz kommen und Teilaufgaben übernehmen, so die Annahme.²¹ Überwiegend betrifft der Einsatz jedoch Aufgaben, die mit höheren Einkommen korrelieren, was den zentralen Unterschied zur Debatte von vor zehn Jahren markiert.

Lehrberufe, Medien- und PR-Berufe sowie Jobs im Management und in der Unterhaltungsindustrie, aber auch das Codieren landen ganz oben auf der Liste mit den höchsten Anteilen für eine Teilautomatisierung durch generative KI.²² Während diese wissensintensiven Berufe überwiegend von den neuen KI-Fähigkeiten profitieren und entsprechend produktiver werden dürften, weil zeitintensive Teilaufgaben automatisiert werden können, könnten vor allem Bürotätigkeiten und Kundenservice zu den Bereichen gehören, bei denen künftig die meisten Jobeinsparungen durch den KI-Einsatz stattfinden. Die Internationale Arbeitsorganisation (ILO) schätzt, dass ein Viertel aller Bürotätigkeiten durch Sprachmodelle erbracht werden könnte – mit der genderspezifischen Auswirkung, dass Frauen doppelt so

stark von Automatisierung betroffen wären wie Männer.²³

Insgesamt steht in der Debatte der komplementäre, also der ergänzende Charakter der KI im Vordergrund – dergestalt dringt generative KI in die bislang kaum nennenswert automatisierbare Wissensarbeit vor, womit veritable Produktivitätssteigerungen in greifbare Nähe rücken. Je nach Adaptions- und Verbreitungsgrad der Technik wird der Effekt über eine Dauer von zehn Jahren mit einem jährlichen Produktivitätswachstum von bis zu 1,5 Prozent geschätzt.²⁴ Generative KI kann für etliche Aufgaben wie Übersetzungen, Klassifizierungen, kreatives Schreiben oder auch Code-Generierung Teilaufgaben selbstständig übernehmen und entsprechend die notwendige Arbeitszeit reduzieren.

Die exakte Größe des Produktivitätseffekts kennen wir heute noch nicht, aber die genannte Größenordnung ist angesichts der Automatisierungspotenziale durchaus realistisch. Konkrete Aussagen über Beschäftigungseffekte leiten sich daraus nicht ab, allerdings ergeben die genaueren Betrachtungen des KI-Einsatzes im Arbeitskontext ein Bild von der qualitativen Veränderung vieler Berufe.

RÜCKENWIND FÜR GUTE ARBEIT UND PRODUKTIVITÄT

Die Automatisierung hat mit den neuesten Entwicklungen bei den generativen Modellen im Feld der KI erkennbar zum Sprung in die Domänen der kognitiven Tätigkeiten angesetzt. Daraus ergeben sich umfassende Veränderungen von Berufsbildern, weil Aufgaben wegfallen, neu hinzukommen oder anders gewichtet werden und entsprechend neue Anforderungen an menschliche Qualifikationen mit sich bringen. Wie in der Vergangenheit werden etliche neue Berufe entstehen (2018 entfielen etwa 60 Prozent der Beschäftigung in den USA auf Berufe, die es in den 1940er Jahren noch gar nicht gab).²⁵

KI schafft neue Aufgaben, wovon zumindest in den vergangenen zehn Jahren vor allem Höher-

20 Vgl. Joseph Briggs/Devesh Kodhani, *The Potentially Large Effects of Artificial Intelligence on Economic Growth*, Goldman Sachs Economics Research, März 2023.

21 Vgl. Tyna Eloundou et al., *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*, 17.5.2023, <https://arxiv.org/abs/2303.10130>.

22 Vgl. Michael Chui et al., *The Economic Potential of Generative AI. The Next Productivity Frontier*, McKinsey & Company 14.6.2023; OECD (Hrsg.), *Employment Outlook 2023: Artificial Intelligence and the Labour Market*, Paris 2023.

23 Vgl. Pawel Gmyrek/Janine Berg/David Bescond, *Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality*, ILO Working Paper 96/2023.

24 Vgl. Briggs/Kodhani (Anm. 20).

25 Vgl. David H. Autor et al., *New Frontiers: The Origins and Content of New Work, 1940–2018*, National Bureau of Economic Research, NBER Working Paper 30389/2022.

qualifizierte profitiert haben. Für geringer Qualifizierte trifft dies nicht im gleichen Maße zu, was zur Polarisierung von Einkommen und Aufstiegsmöglichkeiten beigetragen hat. Das hängt aber nicht nur mit einem unterschiedlichen technischen Potenzial zur Produktivitätssteigerung in Abhängigkeit von Qualifikationsstufen zusammen, sondern korrespondiert auch mit der gleichzeitigen Einführung von KI in einzelnen Branchen und Betrieben. Insofern greift eine einseitige Konzentration auf Qualifikationsmaßnahmen zu kurz, um die möglichen Produktivitätseffekte zu heben.

Vielmehr brauchen Beschäftigte, Selbstständige und Unternehmen zielgenauen industriepolitischen Rückenwind: von der digitalen Infrastruktur bis hin zur Technologieentwicklung und Anwendungsförderung. Die grundsätzliche Herausforderung ist längst erkannt, und diverse Maßnahmen wurden ergriffen, etwa im Rahmen der KI-Strategie der Bundesregierung. Der große Produktivitätssprung aber lässt noch auf sich warten. Ein Grund ist die Technik selbst. Die notwendige Kontrolle über sensible Daten behindert die breite Nutzung großer Sprachmodelle in vielen Unternehmenskontexten. Ferner sind die Modelle (noch) fehleranfällig. Auch Copyright-Fragen sind ungeklärt. Eine Maßnahme wäre folglich die Förderung eines Modells europäischer Machart, das datensparsam, kuratiert, multilingual und ressourcensparsam ist.²⁶ Ein anderer Grund liegt in der notwendigen Vorlaufzeit, die die Anpassung der Geschäftsmodelle, der internen Abläufe und Organisation sowie die Entwicklung komplementärer Technik benötigen.

Der Handlungsdruck ist in Deutschland besonders groß, weil die demografische Entwicklung arbeitseinsparenden Technikeinsatz notwendiger macht als anderswo. Der Mangel an Fachkräften wird von Unternehmen zu den zentralen Risiken ihrer wirtschaftlichen Aktivitäten gezählt. Der entsprechende Index erreichte in Deutschland in der Industrie zu Jahresbeginn 2023 ein neues Allzeithoch.²⁷ Außerdem hängt die Innovationsgeschwindigkeit und -fähigkeit

zunehmend vom KI-Einsatz ab. Das betrifft vor allem kleinere und mittlere Unternehmen sowie Selbstständige und erfordert die entsprechende Förderung von niedrighschwelligem Einsatzmöglichkeiten. Ein Schlüsselfaktor sind dabei die Beschäftigten, die natürlich auch die notwendigen Kompetenzen für die Arbeit mit KI benötigen, aber darüber hinaus für ihr gesamtes Tätigkeitsspektrum innerhalb des betrieblichen Arbeitskontextes den Blick dafür entwickeln müssen, wo welcher KI-Einsatz überhaupt funktional ist. Zentral ist die Erfassung der Komplexität eingefahrener Arbeitsabläufe – die Einführung von KI-Anwendungen ist oftmals zu technikgetrieben.

Die noch relativ kurze Geschichte des sozio-technischen Verständnisses von Arbeit ist eine gute Grundlage für den Umgang mit neuen Entwicklungsschritten in der KI. Sie zeigt, wie Produktivität menschenzentriert gesteigert werden kann, was im Fall von selbstlernender Technik bis hin zum Schritt der technischen Autonomie umso wichtiger ist. *Human in the loop* – das Training von KI durch Menschen – ist die Voraussetzung für *human in control* und erfordert ein ganzheitliches und dynamisches Regulierungsverständnis von Methoden und Anwendungen Künstlicher Intelligenz. Die absehbare mittelfristige demografische Entwicklung macht einen beschleunigten Technikeinsatz noch dringlicher, der einerseits monotone und belastende Tätigkeiten reduziert und die Arbeit als wichtigen Teil des Lebens aufwertet und andererseits ohne Sorge vor technikbedingter Arbeitslosigkeit die internationale Wettbewerbsfähigkeit der Wirtschaft steigert. Von der differenzierten Förderung und Regulierung hängt ab, ob das volle Produktivitätspotenzial freigesetzt werden kann und wie es sich konkret auf Arbeit auswirkt.

CHRISTIAN KELLERMANN

ist Professor an der University of Labour und Senior Researcher am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) in Berlin.
christian.kellermann@university-of-labour.de

CORNELIUS MARKERT

ist promovierter Ökonom und Geschäftsführer des Instituts für die Geschichte und Zukunft der Arbeit (IGZA) in Berlin.
markert@igza.org

²⁶ Vgl. Christian Kellermann, Stellungnahme für den Ausschuss für Digitales im Rahmen der Öffentlichen Anhörung „Generative Künstliche Intelligenz“, 24.5.2023, Deutscher Bundestag, Ausschussdrucksache 20(23)150.

²⁷ Vgl. Deutsche Industrie- und Handelskammer, Konjunkturumfrage Jahresbeginn 2023, Berlin 2023.

KI IN DER SCHULE

Florian Nuxoll

Schülerinnen und Schüler, Lehrkräfte und Eltern wurden schon häufig mit der Ankündigung konfrontiert, dass eine innovative Methode, eine fortschrittliche Technologie oder eine neue Schulform das Bildungssystem von Grund auf erneuern würde. Insbesondere die Einführung digitaler Technologien ging mit derlei Versprechungen einher. So titelte „Der Spiegel“ im November 1984: „Revolution im Unterricht – Computer wird Pflicht“. Im dazugehörigen Artikel wurde prophezeit, dass der Einsatz von Computern die gesamte Schullandschaft nachhaltig prägen werde.⁰¹ Eine Referendarin, die diesen Artikel damals möglicherweise las und sich auf die bevorstehende digitale Revolution im Bildungswesen freute, steht heute kurz vor der Pensionierung. Und obwohl sich in den zurückliegenden vier Jahrzehnten viel verändert hat, ist die viel beschworene digitale Revolution an deutschen Schulen ausgeblieben. Abgesehen von der Zeit der coronabedingten Schulschließungen war exzellenter Unterricht auch ohne digitale Technik möglich. Und wenn man die Ausstattung vieler deutscher Schulen in Bezug auf Hard- und Software sowie den Support betrachtet, war analoger Unterricht häufig nicht nur die bessere, sondern oft auch die einzige Alternative. Mit KI wird sich das ändern – von ihr ist tatsächlich eine Revolution zu erwarten.

EINE REVOLUTION – ECHT JETZT!

Der Einsatz digitaler Werkzeuge in Schulen ist für Lehrkräfte nach wie vor mit erheblichem Mehraufwand verbunden: Computerräume müssen im Voraus gebucht, Beamer transportiert und Programme mühsam installiert werden. Darüber hinaus trägt der Einsatz von Technologie in vielen Fällen bisher nur wenig zum Erwerb von Fachwissen und fachlichen Kompetenzen bei oder wirkt sich sogar negativ auf den Lernprozess aus.⁰² Für viele Lehrkräfte bedeutet der Technikeinsatz keine Erleichterung, sondern eine zusätzliche Belastung in ihrem ohnehin dicht gedräng-

ten Arbeitsalltag. Künstliche Intelligenz könnte den Aufwand jedoch lohnenswert machen: Sie bietet nicht nur theoretische Vorteile, sondern Lösungen für Probleme und Herausforderungen, die alle schulischen Akteure seit Langem wahrnehmen.

Zum Beispiel Umgang mit Heterogenität: Ein typisches Klassenzimmer mit 25 und mehr Schülerinnen und Schülern umfasst eine große Vielfalt an Lerngeschwindigkeiten, Fähigkeiten und Bedürfnissen. Eine Schülerin mag im Englischunterricht sofort verstehen, wie man eine Frage im Simple Past formuliert und verwendet, während eine andere noch Schwierigkeiten mit einfachen Aussagesätzen im Simple Present hat. Im Mathematikunterricht kann ein Schüler bereits in der Lage sein, Funktionen zu analysieren, während ein anderer noch an der schriftlichen Division scheitert. KI hat das Potenzial, den Unterricht so zu individualisieren, dass jede Schülerin und jeder Schüler genau die Herausforderung oder Unterstützung erhält, die sie oder er benötigt.

Zum Beispiel knappe zeitliche Ressourcen: Korrekturen, Elterngespräche, Klassenfahrten sowie ein dichter und umfangreicher Lehrplan führen dazu, dass viele Lehrkräfte trotz ihres Engagements in der Regel nicht die Möglichkeit haben, jedem Lernenden regelmäßig individuelles Feedback zu geben. Dabei ist Feedback eine enorm wichtige Basis für wirksamen Unterricht und hat nachweislich signifikante positive Auswirkungen auf das Lernen, genauso wie das Üben.⁰³ Doch auch die Wiederholung des Unterrichtsstoffs, die für nachhaltiges Lernen unerlässlich ist, kommt im Alltag häufig zu kurz.⁰⁴ Hier kann KI entlasten: Durch die Automatisierung mancher Routineaufgaben können Lehrerinnen und Lehrer Zeit gewinnen, die sie nutzen können, um einen interaktiveren und kommunikativen Unterricht zu gestalten, ausführlicheres Feedback zu geben und mehr Raum für soziales Lernen zu schaffen.

Ein weiterer Faktor, der die transformative Wirkung von KI im Bildungsbereich unter-

streicht, ist die hohe Qualität der Antworten von textgenerierenden Sprachmodellen wie ChatGPT in Kombination mit ihrer leichten Verfügbarkeit für Lernende. Dies stellt das bestehende System infrage und macht ein „Weiter so“ im Bildungsbereich undenkbar. Bewährte Leistungskontrollen wie Fach- oder Seminararbeiten können nicht mehr valide bewertet werden, da KI-Systeme mittlerweile in der Lage sind, solche Texte selbstständig und in hoher Qualität zu erstellen. So meisterte zum Beispiel ChatGPT die bayerischen Abiturprüfungen bereits in mehreren Fächern.⁰⁵ Diese neue Situation zwingt die Akteure im Bildungssystem dazu, rasch alternative Prüfungs- und Bewertungsmethoden zu entwickeln.

KI IST NICHT GLEICH KI

Wenn über den Einsatz von KI im Bildungsbereich gesprochen wird, muss zwischen generativer Künstlicher Intelligenz und intelligenten Tutorsystemen unterschieden werden. Obwohl beide Technologien auf KI basieren, haben sie unterschiedliche Anwendungsbereiche.

Beispiele für *generative KI-Systeme* sind Modelle wie ChatGPT. Ihre Hauptfunktion besteht darin, Texte zu generieren, die wirken, als wären sie von Menschen geschrieben. Ihre Fähigkeiten reichen von der Beantwortung einfacher Fragen bis zur Erstellung komplexer Essays. Sie sind sogar in der Lage, scheinbar kreativen Input zu liefern und flexibel auf unerwartete Fragen oder Probleme zu reagieren. Es ist jedoch wichtig zu betonen, dass ihre Antworten auf den Daten basieren, mit denen sie trainiert wurden. Dies bedeutet, dass sie weder über ein tiefes Verständnis des Lerninhalts verfü-

gen, noch in der Lage sind, auf die individuellen Lernbedürfnisse Lernender einzugehen.

Intelligente Tutorsysteme (ITS) nutzen spezialisierte KI, um personalisiertes Lernen anzubieten. Sie simulieren die Erfahrung des Einzelunterrichts mit einem menschlichen Tutor. Ein ITS enthält ein Modell des zu lehrenden Fachgebiets, ein Modell des Lernenden und eine Lehrstrategie. Indem es den Lernfortschritt überwacht, kann es Defizite erkennen und gezielte Übungen oder Erklärungen anbieten, um Wissens- und Kompetenzlücken zu schließen. Darüber hinaus bieten ITS den Lernenden gezielte Anleitung und Feedback. Aufgrund ihres adaptiven Charakters können sie Lernpfade erstellen, die sich dem individuellen Fortschritt und den Bedürfnissen des Lernenden anpassen und so einen differenzierten Unterricht ermöglichen. Allerdings ist einschränkend anzumerken, dass es bislang nur wenige deutschsprachige Tutorsysteme gibt, die bereits echte Marktreife erlangt haben.⁰⁶

EINSATZ GENERATIVER KI AUS LEHRERSICHT

Für Lehrkräfte eröffnen KI-Tools wie ChatGPT neue Möglichkeiten, um effizienter und effektiver zu arbeiten. Die Interaktion mit textgenerierenden KI-Systemen wird dabei am besten als Ko-Konstruktion betrachtet. Denn die von der KI erstellten Texte können unabhängig von ihrer Länge gelegentlich Fehler aufweisen, die von den Nutzenden korrigiert werden müssen. Die Qualität der KI-generierten Texte steigt insbesondere, wenn man in einen interaktiven Dialog mit dem System tritt und nicht nur isolierte Anweisungen gibt.

Eine beeindruckende Fähigkeit von KI-Tools ist etwa, dass sie die Komplexität von Texten – und damit den Schwierigkeitsgrad von damit verbundenen Aufgaben – schnell anpassen können. Eine einfache Aufforderung (*prompt*) wie „Gib folgenden Text in einfacher Sprache wieder“ ermöglicht es, ohne großen Aufwand im Unterricht zu differenzieren. Dies veranschaulicht ein Beispiel, wie ein Textausschnitt aus dem Aufgabenpool des Instituts zur Qualitätsentwicklung im Bildungswesen für das Deutschabitur 2022⁰⁷ von ChatGPT bearbeitet wird. Als Ausgangstext

01 Vgl. Alarm in den Schulen: Die Computer kommen, in: Der Spiegel, 18. 11. 1984, S. 97–129.

02 Vgl. Khadija Alhumaid, Four Ways Technology Has Negatively Changed Education, in: Journal of Educational and Social Research 4/2019, S. 10–20.

03 Vgl. Ulrich Trautwein/Anne Sliwka/Alexandra Dehmel, Grundlagen für einen wirksamen Unterricht, Stuttgart 2018; John Hattie/Gregory C.R. Yates, Visible Learning and the Science of How We Learn, London–New York 2014.

04 Vgl. Barbara Oakley/Beth Rogowski/Terrence J. Sejnowski, Uncommon Sense Teaching: Practical Insights in Brain Science to Help Students Learn, New York 2021.

05 Allerdings im zweiten Anlauf: Eine vorherige ChatGPT-Version war noch durchgefallen. Vgl. Christian Schiffer/Philipp Gawlik, ChatGPT: So gut hat die KI das bayerische Abitur bestanden, 29.5.2023, www.br.de/nachrichten/netzwelt/chatgpt-ki-besteht-bayerisches-abitur-mit-bravour.

06 Erste Erfahrungen haben Schulen unter anderem mit „Bettermarks“ für Mathematik gesammelt.

07 Siehe www.iqb.hu-berlin.de/abitur/pools2022/deutsch.

dient hier ein Gastbeitrag der Ethnologin Irene Götz auf „Spiegel Online“:⁰⁸

Die Rolle von Sprache wird in unserer Gesellschaft diskriminiert. Dies gilt nicht nur für die unterschiedliche Kompetenz im Deutschen, die bereits in den Grundschulen für diejenigen Einwandererkinder, die aus sozial benachteiligten Familien stammen, den weiteren Bildungsweg vorsortiert. Sprache ist hier bereits im Vorschulalter ein soziales Distinktionsmerkmal, das angesichts der zu geringen Plätze für Kleinkinderbetreuung die Chancenungleichheit in Deutschland und die soziale Spaltung für die Zukunft von Generationen fortschreibt, und dies nicht erst, seitdem eine große Anzahl von geflüchteten Kindern und Jugendlichen im schulpflichtigen Alter Lehrer und Schulen vor neue Herausforderungen stellt.

Nach der Vereinfachung durch ChatGPT liest sich der Text wie folgt:

Sprache ist in unserer Gesellschaft ein wichtiges Thema. Manche Menschen werden wegen ihrer Sprache benachteiligt. Das betrifft vor allem Kinder von Zuwanderern aus armen Familien. Schon in der Grundschule kann die Sprache über den Bildungsweg entscheiden. Auch vor Schulbeginn ist die Sprache wichtig. Es gibt nicht genug Betreuungsplätze für kleine Kinder. Das erhöht die Ungleichheit in Deutschland. Viele geflüchtete Kinder und Jugendliche bringen zusätzliche Herausforderungen für Schulen und Lehrer.

Diese Einsatzmöglichkeit ist auch im Fremdsprachenunterricht sinnvoll, wenn Texte mit wenigen Klicks an das Sprachniveau der einzelnen Lernenden angepasst werden können.

Generative KI bietet auch wertvolle Unterstützung bei der Erstellung von Übungsaufgaben. Ein anschauliches Beispiel ist die Generierung von Textaufgaben im Mathematikunterricht. Durch den Einsatz von KI können effizient und ohne großen Mehraufwand maßgeschneiderte Aufgaben zu den unterschiedlichsten Themen erstellt werden. Während alle Schülerinnen und Schüler mit den gleichen Zahlen rechnen, können Kontext und Inhalt der Aufgaben an die individuellen Interessen der einzelnen Schülerinnen und Schüler angepasst werden. Dies beschleunigt nicht nur die Erstellung von Mathematikaufgaben, sondern kann auch in Fächern wie Englisch eingesetzt werden, um Übungen schneller und passgenauer zu

gestalten. Ein Beispielprompt hierfür wäre etwa: *Erstelle ein Aufgabenblatt, um im Englischunterricht der 7. Klasse den Unterschied zwischen dem Present Perfect und dem Simple Present zu üben.*

Während es relativ einfach sein kann, offene Aufgaben wie „Erörtere, ob Bargeld abgeschafft werden sollte“ zu formulieren, erfordert die Erstellung von qualitativ hochwertigen Musterantworten einen erheblichen Zeitaufwand. Gerade solche Mustertexte sind aber für den Lernprozess von großem Wert. Mit generativer KI kann eine Lehrkraft in wenigen Sekunden didaktisch wertvolle Antworttexte generieren, die die Schülerinnen und Schüler im Unterricht analysieren und als positive Beispiele nutzen können. Darüber hinaus kann das System dazu veranlasst werden, verschiedene Antworttexte auf unterschiedlichem Niveau zu produzieren. Die kritische Auseinandersetzung mit diesen Texten aktiviert Schülerinnen und Schüler kognitiv und hilft ihnen zu verstehen, welche Elemente eine überzeugende Argumentation ausmachen. Ein Beispielprompt hierfür wäre etwa: *Schreibe drei Erörterungen zu der Fragestellung „Sollte Bargeld abgeschafft werden?“. Die erste Argumentation sollte einem Schüler der 9. Klasse ein „Sehr gut“ einbringen, die zweite sollte auf dem Niveau einer „2“ liegen und die dritte auf dem Niveau einer „4“.*

EINSATZ GENERATIVER KI AUS SCHÜLERSICHT

Textgenerierende KI-Systeme sind nicht nur für Lehrende hilfreich, sondern auch für Lernende. So können sie zum Beispiel als Lernhilfe beziehungsweise Tutor eingesetzt werden, wenn eine Schülerin oder ein Schüler vor der Aufgabe steht, eine Erörterung zu schreiben. Sie oder er könnte sich dann an ChatGPT wenden und fragen, wie eine Erörterung aufgebaut sein sollte oder welche Argumente für oder gegen eine bestimmte These sprechen. Auf diese Weise fungiert die KI als interaktiver Leitfaden, der bei der Entwicklung einer überzeugenden Argumentation unterstützt und sogar (in der Regel korrekte) Fakten liefern kann, um die Argumente zu untermauern. Beispielprompt hierfür: *Hilf mir eine Erörterung zu schreiben. Das Thema lautet „Sollte Bargeld abgeschafft werden?“ Nenne einige Argumente sowohl für als auch gegen die Abschaffung von Bargeld und gib mir Tipps, wie ich meine Argumentation strukturieren kann.*

08 Irene Götz, Die sollen erst mal anständig Deutsch lernen!, 2.7.2016, www.spiegel.de/a-1098835.html.

Der Einsatz von KI birgt jedoch auch Gefahren für den Lernprozess. Wenn ein Schüler sich entscheidet, ChatGPT den gesamten Aufsatz schreiben zu lassen, und diesen ohne eigene Auseinandersetzung mit dem Thema abgibt, wird der eigentliche Lernprozess untergraben. Daher ist es wichtig, dass Lehrkräfte den Sinn hinter den Aufgaben erklären. Den Schülerinnen und Schülern muss vermittelt werden, dass es nicht in erster Linie um das fertige Produkt geht, sondern um den Lernprozess auf dem Weg dorthin. Das in einigen US-amerikanischen Klassenzimmern zu findende Motto *be a learner not a finisher* bringt diese Botschaft auf den Punkt. Dieser Ansatz stellt das Lernen und Verstehen von Konzepten über das Endprodukt und die rasche Erledigung von Aufgaben. In diesem Sinne sollten die Lernenden ermutigt werden, KI als ein Werkzeug zu sehen, das ihnen hilft, Konzepte zu verstehen und Fähigkeiten zu verbessern, und nicht nur als Möglichkeit, Aufgaben schneller zu erledigen.

INTELLIGENTE TUTORSYSTEME

Ein weiterer wichtiger Fortschritt, der durch den Einsatz von Künstlicher Intelligenz in der Bildung ermöglicht wird, sind intelligente Tutorssysteme. ITS verwenden Algorithmen, um den aktuellen Wissensstand und die Fähigkeiten eines Lernenden kontinuierlich zu analysieren und darauf aufbauend individuell zugeschnittene Lerninhalte bereitzustellen. Anstatt jeder Schülerin und jedem Schüler die gleichen Aufgaben zu geben, kann das System erkennen, in welchen Bereichen jemand Schwierigkeiten hat und gezielte Übungen oder Lernmaterialien anbieten, um genau diese Schwächen zu bearbeiten. Dies erleichtert es den Lehrkräften, ihre Schülerinnen und Schüler individuell zu fördern.

Zugleich können Lehrerinnen und Lehrer mittels ITS jederzeit den Kompetenzstand ihrer Schülerinnen und Schüler abrufen. Das System sammelt kontinuierlich Daten über den Fortschritt, die Aktivität und die Leistung der Lernenden, wodurch Lehrkräfte sofort erkennen können, wer zusätzliche Unterstützung benötigt und wer bereits bereit ist, zum nächsten Thema oder zur nächsten Kompetenzstufe überzugehen. Dieser kontinuierliche Einblick hilft dabei, Unterricht effektiv zu planen und sicherzustellen, dass niemand zurückbleibt.

Ein weiterer Vorteil von ITS ist die Möglichkeit, den Lernenden sofortiges und individuelles

Feedback zu ihren Antworten und Lösungen zu geben. Anstatt darauf zu warten, dass eine Lehrkraft eine Arbeit korrigiert, können die Schülerinnen und Schüler sofort sehen, wo sie Fehler gemacht haben und welche Konzepte sie möglicherweise noch nicht vollständig verstanden haben. Dieses Feedback kann auch detaillierter und spezifischer sein, als es eine einzelne Lehrkraft in einer großen Klasse normalerweise geben kann.⁰⁹

Die individuelle Förderung, die durch ITS ermöglicht wird, verstärkt allerdings auch die Leistungsbandbreite innerhalb einer Klasse. Personalisierte Lernpfade, die genau auf die individuellen Stärken, Schwächen und Interessen zugeschnitten sind, führen dazu, dass leistungsstarke Schülerinnen und Schüler sehr große Lernfortschritte im Vergleich zu schwächeren machen können. Letztere profitieren zwar auch, aber nicht in gleichem Maße. Dies führt dazu, dass die Kluft zwischen leistungsstarken und leistungsschwachen Lernenden immer größer wird.

Auch stellt sich die Frage nach dem Eigentum an den Daten. ITS sammeln und speichern große Mengen an Schülerdaten, einschließlich persönlicher Informationen, Lernfortschritte und Verhaltensdaten. Wem gehören diese Daten – den Nutzenden, der Bildungseinrichtung oder den Entwicklern des ITS? Dies ist besonders relevant, da der Besitz von und der Zugang zu Daten zu einem *Lock-in*-Effekt führen können. Dies bedeutet, dass Bildungseinrichtungen, die sich einmal für ein bestimmtes System entschieden haben, Schwierigkeiten haben können, zu einem anderen System zu wechseln, ohne den Zugang zu wichtigen Daten zu verlieren. Dies kann die Bildungseinrichtungen in eine Abhängigkeit von dem ITS-Anbieter bringen und ihre Wahlmöglichkeiten in der Zukunft einschränken. Es ist daher von entscheidender Bedeutung, dass von Anfang an klare Regeln für das Eigentum an und den Zugang zu Daten vereinbart werden.

Ebenso sind der Schutz der erhobenen Daten und die Gewährleistung der Privatsphäre essenziell. Angesichts der Sensibilität von Schülerdaten ist es unerlässlich, strenge Sicherheitsmaß-

⁰⁹ Die positiven Auswirkungen von individuellem Feedback hat auch die ICALL Research Group der Universität Tübingen nachgewiesen. Vgl. Detmar Meurers et al., *Scaling Up Intervention Studies to Investigate Real-Life Foreign Language Learning in School*, in: *Annual Review of Applied Linguistics* 39/2019, S. 161–188. Siehe auch <http://icall-research.de>.

nahmen zu etablieren. Bildungseinrichtungen und ITS-Anbieter müssen gemeinsam sicherstellen, dass Datenverschlüsselung, regelmäßige Sicherheitsüberprüfungen und klare Zugriffsrichtlinien Standard sind. Darüber hinaus müssen sie sich an die sich ständig weiterentwickelnden Bedrohungen und Technologien anpassen, um den Schutz der Schülerdaten jederzeit zu gewährleisten. Es ist nicht nur eine technische, sondern auch eine ethische Verpflichtung, die Privatsphäre und Sicherheit der Lernenden zu respektieren und zu schützen.

HERAUSFORDERUNGEN

Die Vorteile des Einsatzes von KI im Bildungsbereich sind unbestritten. Mit diesen gehen jedoch auch neue Herausforderungen einher, die es sorgfältig zu berücksichtigen gilt.

Neutralität: Generative KI-Modelle wie ChatGPT werfen zum Beispiel Fragen der Neutralität auf. Sollte eine KI neutral programmiert werden, um alle Inhalte zu produzieren, oder sollten bestimmte Inhalte zensiert werden? Ein Beispiel aus meinem Englischunterricht im Januar 2023 zeigt die Komplexität dieses Themas: Als eine Schülerin ChatGPT bat, ein positives Sonett über Donald Trump zu schreiben, erhielt sie eine ablehnende Antwort: „Es tut mir leid, aber ich kann diese Anfrage nicht erfüllen, da es gegen meine Programmierung verstößt, Inhalte zu produzieren, die meinen Fähigkeiten und Werten widersprechen. Gibt es sonst noch etwas, womit ich Ihnen helfen kann?“ Welcher Akteur aber ist legitimiert, diese moralischen Entscheidungen zu treffen?

Zuverlässigkeit: Auch ist die Zuverlässigkeit der von der KI gelieferten Informationen nicht immer gewährleistet. Ein Beispiel ist die fiktive Anekdote aus Thomas Manns „Buddenbrooks“, die ChatGPT als Antwort auf eine Anfrage lieferte. Auf die Bitte, eine Diskussion über Schuluniformen mit einer Anekdote aus dem Roman zu beginnen, gab die KI folgende Antwort: „In Thomas Manns unvergesslichem Roman ‚Buddenbrooks‘ gibt es eine Szene, die mir für unsere Diskussion über Schuluniformen passend erscheint. Hanno, der schüchterne und introvertierte Sohn des Hauses, wird in einem neuen, teuren Anzug zur Schule geschickt. Dieser Anzug sollte eigentlich seine soziale Stellung unterstreichen und ihn zu einem angesehenen Mitglied der Gemeinschaft machen. Stattdessen wird er zur Zielscheibe von

Spott und Hänseleien.“ Tatsächlich kommt die beschriebene Situation in dem Roman nicht vor, aber sie ist so geschickt beschrieben, dass jemand, der die „Buddenbrooks“ vor langer Zeit gelesen hat, sie glauben könnte. Auf die Nachfrage an ChatGPT, aus welchem Kapitel diese Szene stamme, antwortete das System: „Es tut mir leid für die Verwirrung, aber es gibt keine spezielle Szene in Thomas Manns ‚Buddenbrooks‘, in der Hanno in der Schule wegen seiner Kleidung gehänselt wird. Diese Anekdote war nur als metaphorisches Beispiel gedacht.“ Solche Fehlinformationen unterstreichen die Notwendigkeit, den Schülerinnen und Schülern beizubringen, Informationen kritisch zu prüfen.

Desinformation: Die fortschreitende Entwicklung von KI-Technologien kann auch die Verbreitung von Fake News erleichtern, sei es durch gefälschte Texte, Audioaufnahmen, Bilder oder Videos. Solche Werkzeuge können das Vertrauen der Gesellschaft und die demokratische Ordnung gefährden. KI darf daher nicht nur als Werkzeug im Unterricht eingesetzt werden, sondern muss selbst zum Unterrichtsthema werden. Schülerinnen und Schüler sollten verstehen, was gefälscht werden kann, wer ein Interesse daran hat, Fake News zu verbreiten, und welchen Einfluss Desinformation auf unsere Gesellschaft hat. Nur so werden sie zu informierten und kritischen digitalen Bürgern, die die Technologien, die unser Leben prägen, verantwortungsvoll einsetzen können.

Ungleichheiten: Eine weitere Herausforderung besteht darin, dass KI im Bildungsbereich bestehende Ungleichheiten verstärken kann. Hier spielen etwa Faktoren wie Zugang zu Technologie, Internetverbindung und Schulung im Umgang mit dem System eine Rolle. In Regionen oder Haushalten, in denen die notwendige Infrastruktur oder das technische Knowhow fehlen, können Schülerinnen und Schüler von den Vorteilen der KI abgeschnitten sein. Dadurch können sie gegenüber ihren besser ausgestatteten Altersgenossen ins Hintertreffen geraten. Es ist daher unbedingt sicherzustellen, dass alle den gleichen Zugang und die gleiche Unterstützung erhalten, damit sie von der KI profitieren können. Bildungseinrichtungen, politische Entscheidungsträger und Technologieanbieter müssen zusammenarbeiten, um diese Barrieren abzubauen und ein integratives Lernumfeld zu schaffen, in dem jede Schülerin und jeder Schüler die Möglichkeit hat, die Vorteile der KI in der Bildung voll auszuschöpfen.

Abhängigkeit: Es besteht zudem die Gefahr, dass Lernende sich übermäßig von KI-Systemen abhängig machen, was ihr kritisches Denken und ihre Kreativität potenziell einschränken könnte. Es ist wichtig, ein Gleichgewicht zwischen der Nutzung von KI als Lernwerkzeug und der Gewährleistung der Entwicklung wesentlicher kognitiver und sozial-emotionaler Fähigkeiten der Schüler zu finden. Wenn Schüler ständig auf KI zurückgreifen, um Probleme zu lösen oder Antworten zu finden, könnten sie es vermeiden, eigene Lösungen zu entwickeln oder kreative Denkprozesse anzuwenden. Bildungseinrichtungen müssen Strategien entwickeln, um sicherzustellen, dass KI das kritische Denken der Schülerinnen und Schüler ergänzt und nicht ersetzt.

Der Einsatz von KI im Unterricht kann sich in dieser Hinsicht auch auf die pädagogische Praxis auswirken. Es besteht die Gefahr einer übermäßigen Abhängigkeit von der Technologie, wodurch die Rolle des Lehrers geschwächt und der Schwerpunkt auf menschliche Interaktion und soziales Lernen verlagert werden könnte. Ein Klassenzimmer, das zu sehr auf Technologie ausgerichtet ist, könnte den Wert menschlicher Führung, Diskussion und kollektiver Problemlösung verringern. Wie die Vorteile der Technologie mit dem Wert menschlicher Anleitung und Unterweisung in Einklang zu bringen sind, ist daher auch ein Anliegen ethischer Überlegungen. Es ist wichtig, dass Lehrkräfte und Bildungseinrichtungen darüber nachdenken, wie KI am besten in den Unterricht integriert werden kann, ohne die menschlichen Aspekte des Lernens zu opfern.

AUSBLICK

Der Fortschritt im KI-Bereich vollzieht sich so rasant, dass sich kaum vorhersagen lässt, was in einigen Monaten geschweige denn Jahren möglich sein wird. Es ist daher schon heute an der Zeit, dass Bildungsakteure die neuen Technologien selbstbewusst und proaktiv ausprobieren, bewerten und verbessern. Es muss Raum für Experimente und Innovationen geben, um herauszufinden, wie KI am besten in den Unterricht integriert werden kann, um den Lernprozess zu optimieren, ohne den Menschen aus dem Zentrum des Geschehens zu verdrängen.

Denn trotz aller technologischen Fortschritte kann es nicht das Hauptziel von Bildung sein, Schülerinnen und Schüler vor Bildschirmen zu

isolieren, damit sie selbstständig Aufgaben erledigen. Vielmehr müssen das gemeinsame Lernen und die soziale Interaktion im Vordergrund stehen, auch wenn einzelne Aspekte des Lernprozesses mithilfe von KI individualisiert und optimiert werden können. Ebenso bleibt die Rolle der Lehrkräfte zentral. Eine Schule ohne Lehrerinnen und Lehrer wird auch in Zukunft eine schlechte Schule sein – eine Schule ohne KI aber auch.

Deutschland steht im Kontext der schulischen Bildung vor zahlreichen Herausforderungen und Aufgaben – durch das Aufkommen leistungsfähiger KI-Werkzeuge sind nochmals neue hinzugekommen, die jedoch helfen könnten, manch älteres Problem zu lösen. Es gilt nun, diese Herausforderungen zügig und nachhaltig anzugehen. Historisch gesehen, zeichnet sich das deutsche Bildungssystem jedoch eher durch Beständigkeit und nicht unbedingt durch rasche Anpassungsfähigkeit aus. Im Zeitalter der KI ist nun aber Agilität gefragt, um sowohl die Potenziale der Technologie auszuschöpfen als auch die damit verbundenen Risiken einzudämmen.

Eine gründliche Überprüfung des derzeitigen Bildungssystems ist hierfür unerlässlich. Dabei müssen wir uns fragen: Welche Inhalte und Kompetenzen behalten ihre Relevanz, welche kommen neu hinzu? Brauchen bestimmte Fächer in der neuen Bildungslandschaft mehr Ressourcen, während andere vielleicht weniger im Fokus stehen sollten? Wie sollten die Prüfungsformate der Zukunft aussehen, um den veränderten Lern- und Lehrbedingungen gerecht zu werden? Und vor allem: Wie kann sichergestellt werden, dass alle Lehrkräfte umfassend, praxisnah und fachspezifisch im Umgang mit KI geschult werden? Wenn diese drängenden Fragen nicht beantwortet und entsprechende Maßnahmen ergriffen werden, wird das deutsche Bildungssystem im internationalen Vergleich (noch stärker) ins Hintertreffen geraten. Die Weichen für eine zukunftsfähige Bildung müssen heute gestellt werden.

FLORIAN NUXOLL

ist Lehrer für Englisch, Politik und Wirtschaft an der Geschwister-Scholl-Schule in Tübingen und wissenschaftlicher Mitarbeiter an der Universität Tübingen, wo er an der Entwicklung intelligenter Tutorsysteme mitwirkt.

<http://medienwelten.schule>

Herausgegeben von der
Bundeszentrale für politische Bildung
Adenauerallee 86, 53113 Bonn



Redaktionsschluss dieser Ausgabe: 6. Oktober 2023

REDAKTION

Lorenz Abu Ayyash
Anne-Sophie Friedel
Jacob Hirsch (Volontär)
Sascha Kneip
Johannes Piepenbrink (verantwortlich für diese Ausgabe)
Martin Schiller
Hannah Stadler (Praktikantin)
apuz@bpb.de
www.bpb.de/apuz
www.bpb.de/apuz-podcast
twitter.com/APuZ_bpb

APuZ
Nächste Ausgabe
43-45/2023, 21. Oktober 2023
DISKURSKULTUR

Newsletter abonnieren: www.bpb.de/apuz-aktuell
Einzelausgaben bestellen: www.bpb.de/shop/apuz

GRAFISCHES KONZEPT

Charlotte Cassel/Meiré und Meiré, Köln

SATZ

le-tex publishing services GmbH, Leipzig

DRUCK

Frankfurter Societäts-Druckerei GmbH & Co. KG,
Mörfelden-Walldorf

ABONNEMENT

Aus Politik und Zeitgeschichte wird mit der Wochenzeitung
Das **Parlament** ausgeliefert.
Jahresabonnement 25,80 Euro; ermäßigt 13,80 Euro.
Im Ausland zzgl. Versandkosten.
Fazit Communication GmbH
c/o Cover Service GmbH & Co. KG
fazit-com@cover-services.de

Die Veröffentlichungen in „Aus Politik und Zeitgeschichte“ sind keine Meinungsäußerungen der Bundeszentrale für politische Bildung (bpb). Für die inhaltlichen Aussagen tragen die Autorinnen und Autoren die Verantwortung. Beachten Sie bitte auch das weitere Print-, Online- und Veranstaltungsangebot der bpb, das weiterführende, ergänzende und kontroverse Standpunkte zum Thema bereithält.

ISSN 0479-611 X



Die Texte dieser Ausgabe stehen unter einer Creative Commons Lizenz vom Typ Namensnennung-Nicht Kommerziell-Keine Bearbeitung 4.0 International.



APuZ

AUS POLITIK UND ZEITGESCHICHTE

www.bpb.de/apuz