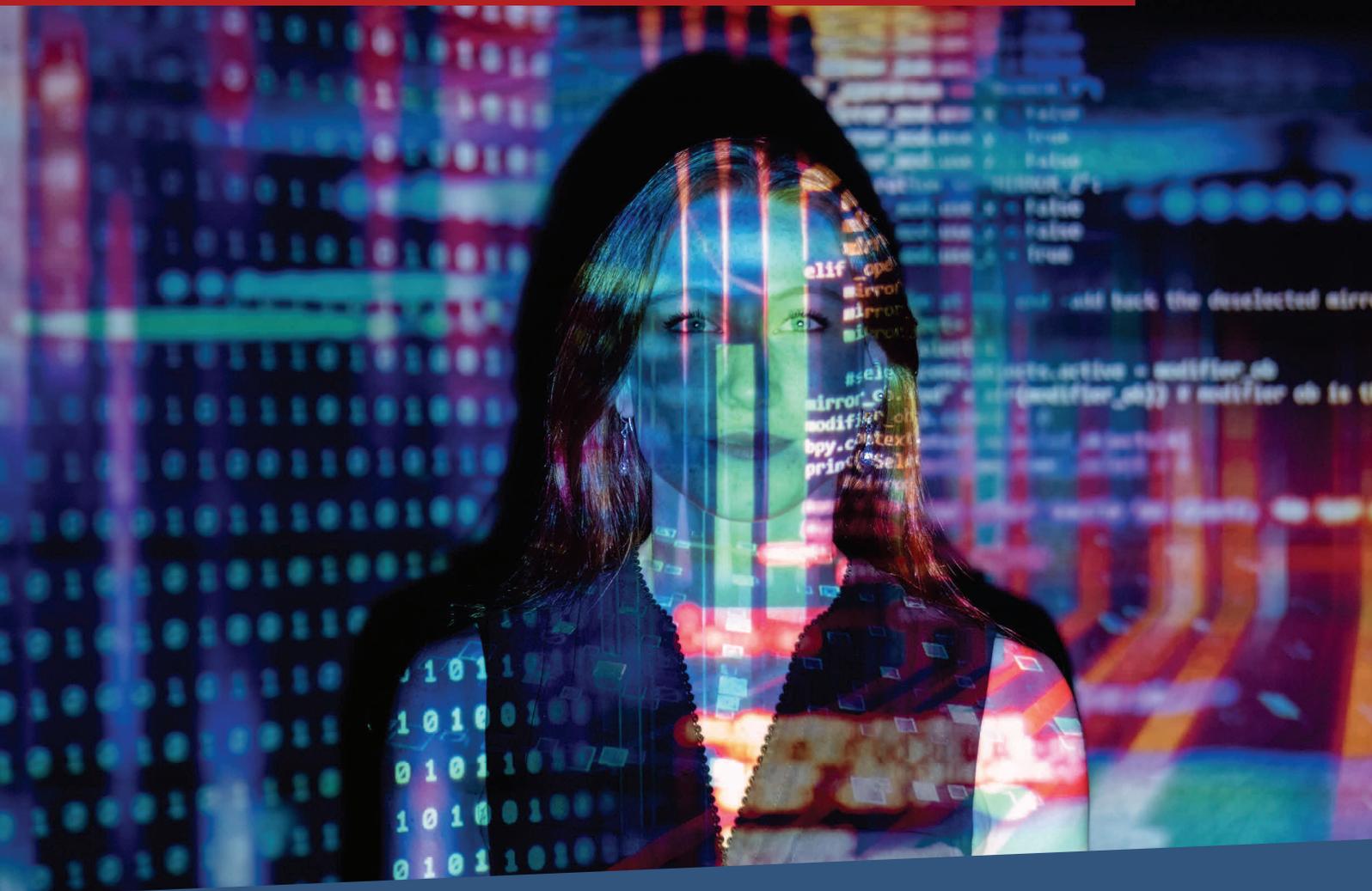




Stiftung für die Rechte zukünftiger Generationen

Generationengerechte Gestaltung von Künstlicher Intelligenz SRzG-Positionspapier



INHALT

1. Künstliche Intelligenz und Generationengerechtigkeit.....	3
1.1 Warum Generationengerechtigkeit bei der Entwicklung und dem Einsatz von KI wichtig ist	3
1.2 Menschliche Intelligenz und Arten der Künstlichen Intelligenz	5
2. Einfluss von KI auf die Lebensumstände zukünftiger Generationen	7
2.1 Auswirkungen von KI auf das Arbeitsumfeld.....	7
2.2 Digitales Eigentum	15
2.3 Transparenz von Algorithmen.....	18
2.4 Diskurs zu KI schaffen	21
2.5 Zukunftsinvestition KI.....	23
3. Starke KI als Gefahr für zukünftige Generationen	26
4. Fazit	30
Literatur	32
Über die Stiftung für die Rechte zukünftiger Generationen (SRzG)	39

"AI could be the greatest or the worst thing to happen to our society since the industrial revolution. Regulation and agreed ethical practices are essential to ensure we protect the rights of all in society because this is not the future, this is the now."

(Stephen Hawking 2017)

1. Künstliche Intelligenz und Generationengerechtigkeit

Künstliche Intelligenz (KI, englisch: Artificial Intelligence, AI) an sich ist nicht neu, aber deren weitreichende Anwendung und Verflechtung mit allen Aspekten unseres Lebens schon. Die Möglichkeiten, die sich mit dieser Anwendung ergeben, sind aufregend, aber gleichzeitig auch beängstigend. Die Fachwelt ist sich einig: Künstliche Intelligenz ist die nächste technische Revolution. Somit wird sie eines der bestimmenden Themen für das Leben junger und noch kommender Generationen sein. Trotz der Vielzahl der Veröffentlichungen zum Thema fehlt bisher der explizite Fokus der Auswirkungen von Künstlicher Intelligenz auf nachrückende Generationen. Diese Lücke möchte dieses Positionspapier schließen.

1.1 Warum Generationengerechtigkeit bei der Entwicklung und dem Einsatz von KI wichtig ist

Die Erschaffung und der Einsatz von Künstlicher Intelligenz (KI, englisch: Artificial Intelligence, AI) zielen darauf ab, intelligentes, menschliches Verhalten bzw. Entscheidungsfindungsprozesse durch Maschinen zu imitieren, abzubilden und, zumindest in Bezug auf bestimmte Aspekte, zu übertreffen (Bitkom 2017). Zukunftsprognosen für die Auswirkung von Künstlicher Intelligenz reichen von dystopischen Verhältnissen, in denen Maschinen den Menschen überlegen sind, bis hin zu nie dagewesenem, durch Technologie geschaffenen, Wohlstand (vgl. WBGU 2019: 76). Es besteht noch viel Forschungsbedarf¹ und wie bei vielen disruptiven Innovationen sind die Auswirkungen schwerer vorherzusagen², je weiter man in die Zukunft blickt. Somit steht unsere Gesellschaft in der besonderen Verantwortung KI jetzt so zu gestalten, dass zukünftige Generationen dadurch nicht benachteiligt werden.

In den letzten Jahren hat sich der gesellschaftliche und politische Diskurs zur Künstlichen Intelligenz stark intensiviert. Auf bundesdeutscher, europäischer, internationaler sowie privatwirtschaftlicher Ebene werden Strategien und werteorientierte Prinzipien formuliert³. Dennoch wird in der Diskussion oft die Stimme zukünftiger Generationen nicht mitbedacht oder erarbeitete Strategien zielen lediglich auf die Maximierung von kurzfristigen wirtschaftlich sowie politischen Erfolgen ab. Da alle Errungenschaften, Veränderungen und Auswirkungen von Künstlicher Intelligenz die Menschheit, also alle nachrückenden Generationen, von nun an irreversibel begleiten werden, soll dieses Positionspapier die Diskussion von KI um den Aspekt der Generationengerechtigkeit bereichern.

Kommende Generationen sind heute noch nicht geboren und damit nicht Teil des üblichen politischen Interessenausgleichs. Daher braucht es ein besonderes Engagement der

¹ Vgl. KI als Thema des Wissenschaftsjahres 2019 (BMBF 2019).

² Vgl. auch das schnelle Durchlaufen des Gartner-Hypezyklus, welcher als maßgeblicher Bewertungsmaßstab bei der Einführung neuer Technologien dient (Panetta 2018).

³ Vgl. „Strategie Künstliche Intelligenz der Bundesregierung“ (Bundesregierung Deutschland 2018), „EU Weißbuch Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen“ (Europäische Kommission 2020), sowie „The AI Policy Sourcebook 2020“ (Rotenberg 2020).

heutigen Generation für nachrückende. Auch junge Menschen sind in politischen Machtgefügen unterrepräsentiert und somit bedarf es auch hier besonderer Anstrengungen, um Gestaltungsmöglichkeiten zu schaffen bzw. zu bewahren. KI ist demnach eine Frage der Gerechtigkeit zwischen der Gesamtheit aller heute Lebenden und den zukünftig geborenen Generationen, aber auch eine Frage der Gerechtigkeit zwischen heute lebenden älteren und jungen Menschen (intertemporale bzw. temporale Generationengerechtigkeit, vgl. Tremmel 2012; 2019). Die Entwicklung und Ausgestaltung von KI muss schon heute generationengerecht sein. Dabei muss bedacht werden, dass wir heute weder die zukünftigen Lebensumstände noch die zukünftigen Präferenzen abschätzen oder gar vorhersagen können (vgl. Caney 2018: 481-482).⁴ Somit fordern wir in Bezug auf Künstliche Intelligenz unter anderem, dass politische und gesellschaftliche Gestaltungsspielräume für künftige Generationen mindestens genauso groß sind wie unsere heute.

Wie ist KI in Einklang mit den Bedürfnissen und Rechten nachrückender Generationen zu bringen? Diese Bedürfnisse können durch verschiedene Formen von Kapitalien operationalisiert werden (Tremmel 2012: 119-212). Nachrückende Generationen benötigen

- ökologisches Kapital in Form eines gesunden Planeten (vgl. Renn und Knaus 1998: 45),
- künstliches Kapital in Form von Infrastruktur, Institutionen und finanziellem Vermögen,
- soziales Kapital in Form intakten Beziehungen, Solidarität, sowie gefestigten gesellschaftlichen Normen,
- kulturelles Kapital wie die sprachliche Vielfalt und
- sie haben, als Individuen, auch einen Anspruch auf menschliches Kapital wie Gesundheit, Bildung, Fähigkeiten und Wissen.

Diese Aspekte finden sich in der Bewertung von KI in den einzelnen Kapiteln wieder. Dieses Papier behandelt zunächst die Auswirkungen von KI auf die Arbeitswelt sowie auf die Herausforderungen, vor denen Menschen in diesem Zusammenhang stehen, inklusive digitale Kenntnisse. Anschließend werden KI-Algorithmen unter dem Aspekt des geistigen Eigentums sowie Transparenz und Verständlichkeit betrachtet. Danach liegt der Fokus auf einem gesellschaftlichen Diskurs zu KI und langfristigen Investition in die Zukunft einer generationengerechten KI.

⁴ Generationengerechtigkeit nach Caney ist dann verwirklicht, zukünftige Generationen mindestens die gleichen Chancen haben wie heutige Generationen, um ihre eigenen Bedürfnisse zu erfüllen (vgl. Caney 2018: 485-486). Nach der Definition des so genannten Brundtland Reports ist Generationengerechtigkeit jedoch bereits dann erreicht wenn heutige Generationen ihre eigenen Bedürfnisse nur in dem Rahmen befriedigen, der gewährleistet, dass auch zukünftige Generationen ihre eigenen Bedürfnisse in Zukunft befriedigen können: „Humanity has the ability to make development sustainable to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs“ (United Nations 1987: 8). Diese Definition setzt allerdings einen Suffizienzstandard bei der Befriedigung zukünftiger Bedürfnisse voraus (vgl. Loske 2007). Wir fordern, wie Caney (2018: 490), dass zukünftige Generationen mindestens genauso gut leben können wie wir.

Bis zu diesem Punkt geht es vornehmlich um heute schon angewandte oder in ähnlicher Form absehbare Anwendungen von KI als Unterstützung in bestimmten Aufgabenfeldern.⁵ Abschließend widmet sich dieses Papier der potenziellen Entwicklung einer künftigen „Superintelligenz“ (vgl. Bostrom 2014) durch KI.

1.2 Menschliche Intelligenz und Arten der Künstlichen Intelligenz

Schon eine generelle Definition von (menschlicher) Intelligenz zu geben ist schwierig und immer abhängig vom Zusammenhang (Gabriel 2018). Somit ist auch die Definition von Künstlicher Intelligenz nicht einfach und noch nicht standardisiert. Allgemein geht es zunächst um „die selbstständige Lösung von Problemen durch technische Systeme“ (WBGU: 2019).⁶

Eine weitverbreitete Definition von KI stammt von Gartner (2019a):

„Künstliche Intelligenz ist eine Technologie, die die menschliche Leistung zu emulieren scheint (typischerweise durch Lernen), zu ihren eigenen Schlussfolgerungen kommt, komplexe Inhalte zu verstehen scheint, in natürliche Dialoge mit Menschen eintritt, die menschliche kognitive Leistung verbessert oder Menschen bei der Ausführung von nicht routinemäßigen Aufgaben ersetzt.“

Es wird oft zwischen schwacher und starker KI unterschieden, wobei starke KI Systeme beschreibt, die dieselben Eigenschaften haben wie Menschen und diese sogar übertreffen können. Schwache KI konzentriert sich auf die Lösung konkreter Problem- und Anwendungsfälle mit Hilfe der Stochastik und Informatik (Bundesregierung Deutschland

⁵ Der Fokus des Positionspapiers liegt auf dem generellen Verständnis der Auswirkungen von KI auf junge und zukünftige Generationen. Das Thema Datensicherheit wird vor allem im Kapitel zur Transparenz der Algorithmen adressiert. Technische Aspekte der Datensicherheit werden in diesem Papier nicht behandelt. Auf eine ethische Bewertung einzelner KI Anwendungen, wie sie unter anderem vom Bundesverband Digitale Wirtschaft (2019) vorgenommen wurde, wird Bezug genommen, sie ist jedoch nicht Kernthema dieses Positionspapiers. Allerdings fordern wir, dass solche Bewertungen gesamtgesellschaftlich und unter Einbezug der Bedürfnisse zukünftiger Generationen vorgenommen werden.

⁶ Der Wissenschaftliche Beirat der Bundesregierung für Globale Umweltveränderungen WBGU (2019: 75) schreibt dazu: „Der Begriff der Künstlichen Intelligenz (KI) wurde 1956 im Rahmen des Forschungsprojektes Dartmouth Summer Research Project on Artificial Intelligence geprägt, die von John McCarthy organisiert wurde. Die grundlegende These war, „dass jeder Aspekt von Lernen oder andere Eigenschaften der Intelligenz so präzise beschrieben werden können, dass man sie mit Hilfe von Maschinen simulieren kann“ (Burgard 2018: 2), [...] Ebenso wie bei „Intelligenz“ gibt es bis heute keine einheitliche Definition von KI, sondern plurale Verständnisse. Allgemein geht es dabei zunächst um „eine Forschungsfrage, die auf die selbstständige Lösung von Problemen durch technische Systeme abzielt“ (Djeffal 2018). Grundsätzlich zeichnen sich intelligente Systeme also durch die Fähigkeit aus, selbstständig und effizient Probleme zu lösen (Mainzer 2016: 2). KI bezeichnet entsprechend aus ihrer begriffsgeschichtlichen Genese „eine Disziplin innerhalb der Informatik, die sich mit der Entwicklung von Softwaresystemen befasst, welche Funktionen bereitstellen, deren Ausführung das erfordert, was typischerweise mit dem Wort Intelligenz bezeichnet wird“ (Burgard 2018). Solche algorithmenbasierten, mit Daten ausgeprägten bzw. maschinell trainierten Funktionen können rein softwareseitig, aber auch gekoppelt mit Hardware z. B. in der Robotik realisiert werden. KI erweist sich somit als „Sammelbegriff für diejenigen Technologien und ihre Anwendungen, die durch digitale Methoden auf der Grundlage potenziell sehr großer und heterogener Datensätze in einem komplexen und die menschliche Intelligenz gleichsam nachahmenden maschinellen Verarbeitungsprozess ein Ergebnis ermitteln, das gegebenenfalls automatisiert zur Anwendung gebracht wird“ (Datenethikkommission, 2018: 1).“

2018, Bitkom 2017). Die Übergänge sind allerdings oft fließend. Dieses Positionspapier bezieht sich auf beide Arten; die Möglichkeit einer starken KI wird in Kapitel 3 behandelt.

In diesem Positionspapier fordert die Stiftung für die Rechte zukünftiger Generationen, dass:

- die mittelfristig absehbaren Auswirkungen und Folgen von Künstlicher Intelligenz so gesteuert und gefördert werden, dass die Gesamtgesellschaft davon profitieren kann,
- dass einzelne Gruppen nicht aufgrund ihrer demografischen Zugehörigkeit (Geschlecht, ethnischer Hintergrund, Einkommen, Bildungsgrad, Alter oder Generationszugehörigkeit) benachteiligt werden, und
- dass heutige Entscheidungen nicht zu Entwicklungen führen dürfen, durch die politische und gesellschaftliche Gestaltungsspielräume für zukünftige Generationen unwiderruflich eingeschränkt werden.

Für die langfristigen Folgen gilt es den zukünftigen Handlungsspielraum nicht zu verkleinern. Demnach sollen die Entwicklung und der Einsatz von Künstlicher Intelligenz:

- zum Vorteil künftiger Generationen stattfinden,
- nach gemeinsam definierten ethischen Grundsätzen erfolgen,
- zur nötigen Transparenz verpflichtet sein,
- frei von Diskriminierung und inklusiv sein,
- nicht entkoppelt von ihrem Ressourcenverbrauch betrachtet werden und zu einer ökologisch nachhaltigen Entwicklung beitragen, und
- darauf abzielen zu jeder Zeit vom Menschen beherrschbar zu bleiben.

Das von der Bundesregierung Deutschland (2018) angestrebte Qualitätsmerkmal „AI made in Germany“ befürworten wir. Es strebt an, Merkmale für eine verantwortungsvolle und gemeinwohlorientierte Entwicklung und Nutzung von KI in Deutschland und für die internationale Zusammenarbeit zu definieren. Unser Positionspapier soll diese vertiefen und insbesondere um den Aspekt der Generationengerechtigkeit im Forschungs- und Anwendungsfeld Künstliche Intelligenz ergänzen.

Die Forderungen sind am Ende der jeweiligen Kapitel formuliert. Sie reichen von thematischen Forderungen (z.B. nach bestimmten gesetzlichen Rahmenbedingungen) bis hin zu prozeduralen Forderungen (z.B. bei der Forschung und Entwicklung von KI). Sie können auf lokale, nationale, europäische oder globale Ebenen bezogen werden, und sind deshalb allgemein und flexibel formuliert.

2. Einfluss von KI auf die Lebensumstände zukünftiger Generationen

2.1 Auswirkungen von KI auf das Arbeitsumfeld

"Anything that's invented between when you're fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it."

Douglas Adams, *The Salmon of Doubt*

Künstliche Intelligenz wird die Zukunft der Arbeit maßgeblich beeinflussen, den Einfluss werden vor allem junge Menschen und kommende Generationen spüren. Zum einen werden bestimmte Arbeitsprozesse zunehmend automatisiert ablaufen. Zum anderen können sie durch KI so unterstützt werden, dass die menschliche Arbeitsleistung durch das Zusammenspiel von Mensch und Maschine verbessert wird. Auch wenn KI aktuell noch nicht flächendeckend angewandt wird, kommen KI-Tools, je nach Branche, bereits heute in bis zu jedem fünften Unternehmen in mindestens einem Aufgabenbereich zum Einsatz (McKinsey, zitiert in Shoham et al. 2018: 36). Allein im Jahr 2018 hat sich der Einsatz von KI in Unternehmen verdreifacht (Gartner 2019b). Auf den KI-Einsatz speziell in Deutschland geht eine Studie des Bundeswirtschaftsministeriums ein (BMWi 2020). Es ergibt sich zum einen großes Potenzial für eine Steigerung von Effizienz, Produktivität, Präzision, menschlicher Sicherheit oder Komfort bei der Arbeit. Zum anderen steht jedoch auch eine massive Transformation von Berufsbildern, Sektoren und der Gesellschaft als Ganzes bevor.

Herausforderung: Sind Algorithmen die besseren Chefs?

Ein Beispiel für Künstliche Intelligenz in der Arbeitswelt von heute ist die Anwendung bei der Vermittlung zwischen Bewerberprofilen und offenen Stellen. Die Zuteilung einer Bewerbung erfolgt hier durch Algorithmen, die mit Künstlicher Intelligenz ausgestattet sind und kontinuierlich weiterlernen. Die Grundlage auf deren Basis die Vermittlung erfolgt, bleibt für die Beteiligten in der Regel unbekannt. Bei gewissenhafter Programmierung kann jedoch die Zuteilung der Bewerber/innen und Stellen nach klar definierten Regeln erfolgen. Subjektivität könnte somit minimiert werden, solange die Entscheidungsgrundlage transparent gemacht wird. Allerdings ist dies aktuell nicht die Regel. Somit kann die Diskriminierung oder Bevorzugung bestimmter Personengruppen auch systematisch werden (Dornis 2018).

Sozioökonomische Folgen abfedern

Künstliche Intelligenz wird menschliche Arbeitskraft nicht überflüssig machen, jedoch zu einer starken Transformation von Erwerbsarbeit führen. KI wird gesamtgesellschaftlich betrachtet voraussichtlich zu großen Produktivitätsgewinnen führen, neue Produkte und neue Geschäftsmodelle möglich machen. Jedoch birgt KI auch das Potenzial, bestehende Ungleichheiten zu verfestigen und auszuweiten. Die Rolle der Politik muss es sein, die Nutznießer*innen einer solchen digitalen Transformation in die Verantwortung zu nehmen,

um die gesamte Gesellschaft, inklusive junge Menschen und zukünftige Generationen, an diesen Vorteilen teilhaben zu lassen.

Arbeiter*innen müssen in diesem digitalen Transformationsprozess geschützt werden, um negative sozioökonomische Folgen abzufedern. Junge Menschen sind heute stärker von prekären Arbeitssituationen betroffen (Gründinger et al. 2017); für sie ist dies deshalb besonders kritisch. Zusätzlich zu ökonomischen Nachteilen hinterlassen prekäre Arbeitsverhältnisse, in Kombination mit sich rapide verändernden Technologien, bereits heute psychologische und gesundheitliche Spuren bei jungen Menschen (vgl. Denkschmiede Gesundheit 2019). Wenn die negativen Folgen von KI-getriebener digitaler Transformation unverhältnismäßig auf bestimmte Gruppen entfallen, könnten sich solche Trends verstärken.

Kurzfristig sind von Digitalisierung, inklusive KI, vermehrt ältere Generationen betroffen, beispielsweise durch weniger Flexibilität oder weniger verbleibender Arbeitszeit um Umschulungen zu amortisieren. Die daraus resultierenden Probleme können und müssen von jetzigen Sozialsystemen abgefangen werden. Die Herausforderungen für zukünftige Generationen hingegen sind strukturell. Wenn die Wertschöpfung immer mehr von Erwerbsarbeit entkoppelt wird, geht dies zu Lasten junger Menschen ohne Kapital. Politische Unterstützungsmaßnahmen müssen beides im Blick haben.

In der jüngeren Vergangenheit waren Streiks und Arbeitsverweigerung eine wirksame Methode für Arbeitnehmer*innen, um auf prekäre Arbeitsbedingungen aufmerksam zu machen, denn ihre Arbeitskraft war letztlich unverzichtbar. Durch KI könnten bestimmte Arbeiten redundant werden, wie beispielsweise Fahrer*innen in Folge einer zunehmend automatisierten Logistikbranche. Die Personen, die diese Arbeiten bisher ausgeführt haben, werden – zumindest was ihre Lohnarbeit betrifft – ersetzbar (Harari 2018: 7ff). Gleichzeitig schwindet auch ihre Einflussmacht, um ihre arbeitsrechtlichen Interessen durchzusetzen. Solche Veränderungen werden manche Berufsgruppen stärker betreffen als andere, und auf dieser Basis könnte es schwieriger sein, zwischen verschiedenen Interessensgruppen zu vermitteln. Gewerkschaften haben historisch eine wichtige Rolle bei der Verbesserung von Arbeitsbedingungen gespielt. Durch die Digitalisierung müssen auch sie sich verändern. Eine Ausweitung ihrer Arbeit auf Berufsgruppen, die sich nicht einfach gewerkschaftlich organisieren können (z.B. Arbeiter*innen der Gig-Ökonomie und in prekären Arbeitsverhältnissen) wird in Zukunft immer notwendiger werden. Durch den Trend zur Dienstleistungsgesellschaft kommt hier den Gewerkschaften in diesem Bereich eine besondere Verantwortung zu.

Herausforderung: Sind Roboter die besseren Autofahrer*innen?

Teil-autonome Autos sind bereits heute Realität und werden in bestimmten Anwendungsgebieten schon eingesetzt. Neue Gesetzgebungen werden einige Berufsbilder maßgeblich beeinflussen. Einerseits birgt autonomes Fahren das Potenzial, die Verkehrssicherheit allgemein zu verbessern, denn bei 75-95% aller Unfälle zählt

menschliches Versagen zu den Ursachen (Europäische Kommission 2016). Im Gegensatz zu einzelnen Menschen, die hauptsächlich durch ihre eigenen Erfahrungen lernen, können KI-gesteuerte Fahrzeuge zudem durch jeden geschehenen Unfall insgesamt dazu lernen und können so verhindern, dass sich gleiche Unfalltypen in Zukunft wiederholen (WDR 5 2019a). Andererseits sind Herausforderungen zur sicheren Datenverarbeitung, zur Straßenverkehrssicherheit von Fußgänger*innen, sowie Fragen zur Haftung im Falle von Unfällen noch nicht geklärt (Europäisches Parlament 2019). Hinzu kommen außerdem ethische Fragen, wie Entscheidungen zwischen verschiedenen potenziellen Schäden, und Fragen zur Sozialverträglichkeit.

Arbeitsplätze sind kein Null-Summen-Spiel. Wenn durch KI bis 2020 geschätzte 1,8 Millionen Arbeitsplätze wegfallen und 2,3 Millionen neue entstehen (Gartner 2017), bedeutet dies nicht, dass Arbeitskräfte einfach von einer Branche in die nächste „wandern“ können. Ein Arbeitsplatzverlust durch KI stellt Individuen wie Gesellschaft vor Herausforderungen. Arbeitsmarkt- und bildungspolitische Investitionen werden unabdingbar. Eine angemessene Grundsicherung für nicht-arbeitende Menschen muss ebenso gewährleistet werden wie Anpassungs- und Aufstiegsfortbildungen, Umschulungen und Weiterbildungen, die der neuen Arbeitsrealität entsprechen. Die Bundesagentur für Arbeit sollte über Fortbildungsmaßnahmen beraten sowie psychologische Unterstützung für Menschen in sich drastisch verändernden Berufen vermitteln. Zudem sollte sie ihre Bewerbungstrainings an die bereits heute eingesetzten, KI-gestützten Personalauswahlverfahren anpassen und hierfür sensibilisieren. Der Diskriminierung durch KI- und datenbasierten Lösungen – dem Bias – muss entgegengewirkt werden (vgl. Feast 2019 und siehe Kapitel 2.3 zu Transparenz).

Herausforderung: Diskriminierende Algorithmen – der Bias

Ein massives, bislang noch ungelöstes Problem von Algorithmen und Künstlicher Intelligenz ist der Bias. Dieser Anglizismus beschreibt eine subjektive und teils vorurteilsbehaftete Verzerrung der Realität, entstehend durch unausgewogene oder unvollständige Daten, die für das Trainieren des Algorithmus verwendet werden; oder durch fehlerhaftes Design, welches die voreingenommene oder beschränkte Sicht von homogenen Gruppen von Analysten und Entwicklern widerspiegelt (Bitkom, 2017).

Wenn marginalisierte Gruppen in Datensätzen und Entwicklerteams nicht ausreichend repräsentiert sind, führt dies dazu, dass bestehende strukturelle Ungleichheiten für zukünftige Generationen noch weiter vergrößert werden. Viele Algorithmen werden mit denselben Bilddaten trainiert und diese wiederum stammen aus einigen wenigen Ländern (Zou und Schiebinger 2018) und repräsentieren damit nur bestimmte Gesellschaftsteile. Potentielle Lösungen beinhalten die interdisziplinäre Zusammenarbeit von KI Forschern mit Experten anderer Disziplinen (z.B. Sozialwissenschaften) sowie eine kontinuierliche

Weiterentwicklung der Algorithmen.

Beispiele für Bias in Algorithmen der KI sind KI-gesteuerte Bewerbervermittlungstools, die bestimmte Personengruppen (wie Frauen oder nicht-weiße Menschen) systematisch schlechter bewerten, beispielsweise basierend auf vorherrschenden strukturellen Missständen, die sich in den verwendeten Datensätzen widerspiegeln. Gesichtserkennungstools, die Menschen mit schwarzer Haut mit Gorillas verwechseln (Kaltheuner / Obermüller 2018); oder KI-unterstützte „präventive“ Polizeiarbeit, welche bestimmte Personengruppen anders beurteilt oder schneller verdächtigt. Solche Problematiken können nur mit Hilfe von diversen, interdisziplinären und intersektionalen Teams eingedämmt werden.

Insgesamt gilt es, die maßgeblichen Nutznießer der „KI-Revolution“ in angemessener Weise an den gesellschaftlichen und finanziellen Kosten zu beteiligen, ohne Innovationsanreize zu verringern. Eine stetig zunehmende Entkopplung von Wertschöpfung und Erwerbsarbeit verfestigt bestehende Ungleichheiten und geht langfristig auf Kosten der gesamten Gesellschaft, und im Besonderen zu Lasten zukünftiger steuerzahlender Generationen. Mögliche Finanzierungsmodelle könnten eine vielfach geforderte Algorithmus- oder Roboter-Steuer sein. Unabhängig davon welches Modell gewählt wird: eine zukunftsorientierte Besteuerung muss sicherstellen, dass der Wegfall von Steuern aus menschlicher Erwerbsarbeit ausgeglichen wird, und dass ausreichend finanzielle Ressourcen für gesellschaftlich notwendige Aufgaben zur Verfügung stehen wie Bildung, nachhaltige Entwicklung, Gesundheit, Infrastrukturschaffung etc. In bestimmten Branchen müssen neue Modelle zur Vergütung von gesellschaftlich wertvoller Arbeit gefunden werden. Denn bereits heute lässt sich nicht mehr leugnen, dass aktuelle, rein monetär ausgerichtete Modelle zu kurz greifen, wie beispielsweise in der Gesundheit und Pflege, Erziehungsarbeit oder in der Grundversorgung.

Zugang zu Arbeitsplätzen

Der Arbeitsmarkt wird sich strukturell verändern, und somit auch die potenziellen Zugänge für viele Menschen (Gartner 2019c). Neue und digitale Arbeitsplätze setzen meistens einen hohen Bildungs- und Ausbildungsgrad voraus. Deshalb müssen gleiche Chancen und Zugang zu hochwertigen und bezahlbaren Bildungsangeboten sichergestellt werden.

Das (duale) Ausbildungssystem, das gerade für Deutschland einen wesentlichen Erfolgsfaktor darstellt, sollte folglich schnellstmöglich neue KI-bedingte Qualifikationserfordernisse vorausschauend abbilden. Gegenwärtig werden Ausbildungsberufe auf der Basis von Vorgesprächen der verschiedenen Akteure initiiert und modernisiert; hierbei wird die Ausbildungsordnung festgelegt. Bis zur Anerkennung einer neuen oder modernisierten Ausbildungsordnung vergehen üblicherweise Jahre – von der Bedarfserkennung auf Seiten der Unternehmen bis zum Erlass der Verordnung. Auch das

Bundesinstitut für Berufsbildung kann auf Grundlage von Forschungsergebnissen und Gutachten die Eckwerte für neue Ausbildungsordnungen erarbeiten und zur Prüfung vorlegen. Künstliche Intelligenz durchdringt mittlerweile so viele Arbeitsbereiche, dass die Einrichtung einer speziellen Forschungseinheit am Bundesinstitut für Berufsbildung zu diesem Thema notwendig ist. Diese Einheit sollte ausschließlich und fortwährend die konkreten Auswirkungen von Künstlicher Intelligenz auf die Qualifikationserfordernisse im deutschen Arbeitsmarkt und im beruflichen Bildungssystem analysieren und antizipieren. Hierdurch steigt die Chance, dass jungen Auszubildenden rechtzeitig die Kompetenzen vermittelt werden, die für die Digitalökonomie relevant sind.

Nicht nur das Ausbildungssystem muss ergänzt werden. Auch die berufliche (Weiter-) Bildung muss weiterentwickelt werden, um mit dem von KI getriebenen Strukturwandel mithalten zu können. Laut dem Bericht über das Bildungswesen in Deutschland, der von der Kultusministerkonferenz der Länder und dem Bundesministerium für Bildung und Forschung in Auftrag gegeben wurde, bildeten sich im Jahre 2016 nur ca. 50% der Beschäftigten weiter, davon allerdings nur 26% außerhalb ihres jetzigen thematischen Umfelds (Maaz et al. 2018: 174). Dies verdeutlicht, dass sich die Kultur der Weiterbildung wandeln muss und Faktoren, die die Weiterbildung hemmen, beseitigt werden müssen. Junge Menschen benötigen größere zeitliche, finanzielle und strukturelle Freiräume für Weiterbildungsmaßnahmen, um sich an neue Anforderungen im Arbeitsmarkt anpassen zu können:

- a. Zeitliche Freiräume: Zeit⁷ ist die Grundvoraussetzung für zusätzliche oder weiterführende Qualifizierung. Eine Bildungskarenzperiode beispielsweise schafft einen entsprechenden zeitlichen Freiraum. Hierbei könnten Arbeitnehmer*innen im Einvernehmen mit dem Arbeitgeber für eine Weiterbildung von bis zu einem Jahr ohne Lohnfortzahlung alle sechs Jahre freigestellt werden. So könnten sie berufliche Aufstiegsfortbildungen oder Studiengänge im In- und Ausland absolvieren, die modularer als bisher aufgebaut sein sollten.
- b. Finanzielle Freiräume: Junge Menschen können gerade in der Anfangsphase ihrer beruflichen Laufbahn nur wenig Vermögen bilden, von dem sie in einer Weiterbildungsphase zehren können. Wenn früh Weiterbildungsphasen nötig sind, brauchen sie Unterstützung. Darüber hinaus sollten Aus- und Weiterbildungen von Förderungswerken finanziell unterstützt werden können. Alters- und Einkommensbeschränkungen für verzinsliche ausbildungs- und weiterbildungsbezogene Förderprogramme der Förderbanken sollten beseitigt werden.
- c. Strukturelle Freiräume: Damit berufliche Weiterbildungsangebote am besten in Anspruch genommen werden können, sind Modularität, Flexibilität sowie deren Qualität und Vergleichbarkeit sehr wichtig. Die Politik sollte daher die Berufsschulen zu regionalen und branchenspezifischen Digitalkompetenzzentren

⁷ Zeit, in der keiner Erwerbsarbeit nachgegangen werden muss.

weiterentwickeln, in denen Aufstiegs- und Anpassungsfortbildungen arbeitsstättentnah begonnen und beendet werden können. Auch gilt es, einzelne universitäre Veranstaltungen durch den Nachweis relevanter und einschlägiger Berufserfahrung oder die entsprechende formale Qualifikation zugänglich zu machen und durch eine Prüfung abschließen zu können.

Herausforderung: Anpassung der Bildungseinrichtungen auf veränderte Lebensrealitäten

Bildungseinrichtungen müssen sich an die veränderten Spielregeln und Anforderungen anpassen. Konkrete Vorschläge, wie bestehende (Berufs-) Bildungseinrichtungen angepasst werden können beinhalten:

- Aufbau einer speziellen Forschungseinheit zu KI am Bundesinstitut für Berufsbildung einrichten.
- Berufsschulen zu regionalen und branchenspezifischen Digitalkompetenzzentren weiterentwickeln, in denen Aufstiegs- und Anpassungsfortbildungen arbeitsstättentnah begonnen und beendet werden können.
- Einzelne universitäre Veranstaltungen durch den Nachweis relevanter und einschlägiger Berufserfahrung oder die entsprechende formale Qualifikation zugänglich zu machen und durch eine Prüfung abschließen zu können.
- Vermittlung von Kompetenzen zur Unternehmensgründung, um eine Gründerkultur in Deutschland zu stärken und einen pluralistischen Arbeitsmarkt erhalten und ausbauen zu können. Dies kann nicht nur zu Wirtschaftswachstum, sondern auch zur Lösung von gesellschaftlichen Herausforderungen beitragen.

Neben der Reform des beruflichen Bildungssystems ist die Stärkung einer Gründerkultur ein essenzieller Aspekt, um einen pluralistischen Arbeitsmarkt erhalten und ausbauen zu können. Wichtig sind zudem nicht nur Arbeitsplätze bei bereits bestehenden Unternehmen, sondern auch die Gründung von neuen Startups im Bereich KI, welche in den letzten Jahren ein großes Wachstum verzeichneten (Shoham et al. 2018: 31). Unternehmensgründungen in diesem Bereich sollten umfassend gefördert werden, um deren Potenzial voll auszuschöpfen – nicht nur für Wirtschaftswachstum, sondern auch zur Lösung gesellschaftliche Herausforderungen (z.B. Startups in den Bereichen Klimaschutz, soziale Startups, etc.). Hierzu gehört nicht nur eine ideelle und finanzielle Förderung. Insbesondere Startups, die Anwendungen im Bereich des maschinellen Lernens entwickeln, sollten einen besseren Zugang zu entsprechenden Trainingsdaten (Daten um Algorithmen zu trainieren welche maßgeblich die Qualität beeinflussen) erhalten. Um dies zeitnah zu realisieren, könnte die öffentliche Hand die ihr zur Verfügung stehenden nicht-personenbezogenen und sicherheitspolitisch nicht-sensiblen Daten zuvor akkreditierten Forschungseinrichtungen und Unternehmen zur freien Nutzung zugänglich machen. Hierbei dürfen nur solche Daten bereitgestellt werden, die keine Rückschlüsse auf den Einzelnen erlauben. Dies erfordert eine europaweite Definition von Standards zu Methoden der Anonymisierung und

Verschlüsselung sowie zur Zugänglichkeit, Speicherung und Verarbeitung sämtlicher Daten. Die Voreingenommenheit und dadurch bedingte Diskriminierung, der sogenannte Bias, muss hierbei berücksichtigt, beziehungsweise ihm effektiv gegengewirkt werden.

Details zu den Auswirkungen von KI auf die schulische Ausbildung die Anwendung von KI in der Bildung selbst sind nicht Teil dieses Papiers.

Art und Bedingungen von Arbeit

Es wird erwartet, dass die Digitalisierung, inklusive Künstlicher Intelligenz, die Produktivität menschlicher Arbeitskraft massiv erhöhen wird (Accenture 2016 und Siemens 2019: 24). Dies wirft die Frage auf, ob und wie die Teilhabe von Arbeitnehmer*innen an dieser gesteigerten Produktivität sichergestellt wird, und wie die Gesellschaft als Ganzes zusätzlich zu den nutznießenden Industrien mit-profitieren kann.

Herausforderung: Wem soll eine gesteigerte Produktivität nutzen?

Eine potenzielle Steigerung der Produktivität in Verbindung mit einer Entkopplung von der Lohnarbeit, stellt in unserem jetzigen System eine Herausforderung dar. Eine Reihe von Fragen diesbezüglich muss im öffentlichen Diskurs (vgl. 2.4 adressiert) werden:

- Ist im Hinblick auf eine erhöhte Produktivität pro menschliche Arbeitskraft eine reduzierte Regelarbeitszeit denkbar?
- Oder wird dies Kapazitäten und Ressourcen für solche Aufgaben freimachen, die ökonomisch schwer messbar sind, die jedoch für menschliche Interaktion oder die Gesellschaft als Ganzes wertvoll sind – wie beispielsweise Aufgaben der Grundversorgung, emotionale, soziale oder Care-Arbeit?
- Könnte eine gesteigerte Netto-Produktivität genutzt werden, um klima- und umweltfreundliche Lösungen für heutige Probleme zu finden?
- Welche Aufgaben sollen weiterhin oder verstärkt von Menschen ausgeführt werden und welche können an KI ausgelagert werden?

Das Ziel des Verwendens von Künstlicher Intelligenz sollte sein, Menschen zu unterstützen und Wohlstand nachhaltig zu sichern. Die Künstliche Intelligenz sollte stets eine dienende niemals dominierende Rolle einnehmen. Bei alledem müssen der Mensch und menschliche Werte im Fokus bleiben, wie auch aus der KI-Strategie und den KI-Ethik-Leitlinien der Europäischen Kommission hervorgeht (Europäische Kommission 2018 und 2020). Um dies zu gewährleisten, könnte ein ständiger Ethikrat für Künstliche Intelligenz eingeführt werden. Dieser sollte sich aus verschiedenen gesellschaftlichen Interessenvertreter*innen auf europäischer Ebene zusammensetzen, welche diese KI-Ethik-Leitlinien weiterentwickeln. Deren Einhaltung könnte durch eine europäische Aufsicht sichergestellt werden (siehe Kapitel 2.4).

Obwohl junge Menschen sowie zukünftige Generationen von solchen Veränderungen massiv betroffen sein werden, haben sie in der Regel schlechten Zugang zu kollektiven Entscheidungsprozessen, da sie politisch stark unterrepräsentiert sind (Inter-Parliamentary Union 2018). Somit ist gesondert darauf zu achten, junge Menschen in einen solchen öffentlichen Diskurs und partizipativen Prozess einzubeziehen – beispielsweise durch Quoten oder spezielle Jugendgremien (siehe Kapitel 2.4 zu Diskurs). Zudem ist breite Bildung zu ethischer, sozialer und gesellschaftlicher Verantwortung im Hinblick auf KI notwendig – sowohl von jungen Menschen im Lauf ihres Bildungsprozesses als auch von Entscheidungsträger*innen in Wirtschaft und Politik.

Forderungen zum Einsatz von KI in der Arbeitswelt:

- Wertschöpfung und Erwerbsarbeit können voneinander durch KI entkoppelt werden, die **sozioökonomischen Folgen** hiervon müssen abgefedert werden.
- Die gesellschaftlichen und finanziellen Kosten und **Nutzen von KI** müssen verantwortungsvoll und **nachhaltig verteilt werden**, damit menschliche Erwerbsarbeit nicht abgewertet wird, vor allem in gesellschaftsrelevanten Berufen, deren Vergütung bereits heute unsozial gestaltet ist (z.B. Pflege, Erziehungsarbeit oder Grundversorgung).
- Eine angemessene Grundsicherung für nicht-arbeitende Menschen muss ebenso gewährleistet werden wie **Umschulungen und Weiterbildungen**, die der neuen Arbeitsrealität entsprechen. Gleiche Chancen und Zugang zu hochwertigen und bezahlbaren Bildungsangeboten müssen sichergestellt werden. Bestehende Bildungseinrichtungen müssen angepasst werden.
- **Breite Bildung** über ethische, soziale und gesellschaftliche Verantwortung im Hinblick auf KI muss verpflichtend sein, insbesondere für Entscheidungsträger*innen in Wirtschaft und Politik.
- Alle Menschen müssen sowohl Zugang als auch genügend zeitliche, finanzielle und strukturelle Freiräume haben, damit eine den Umständen angemessene Auseinandersetzung und ein **inklusive Diskurs** mit KI für jede*n möglich ist. Besonders diskutiert werden müssen die Art und Bedingungen von Arbeit, um bspw. zu entscheiden, wie eine erhöhte Produktivität pro menschliche Arbeitskraft der gesamten Gesellschaft dienen kann, sowie welche Arbeiten weiter von Menschen verrichtet werden sollen, und welche an KI „ausgelagert“ werden können.
- **Gewerkschaften** müssen sich stärker mit den Auswirkungen von KI beschäftigen und ihre Vertretungsrolle weiterdenken. Insbesondere müssen Gewerkschaften auch Berufsgruppen vertreten, für die eine gewerkschaftliche Organisation schwierig ist.

2.2 Digitales Eigentum

„Denn es ist zuletzt doch nur der Geist, der jede Technik lebendig macht!“

Johann Wolfgang von Goethe

Das folgende Kapitel beschäftigt sich mit der Frage nach digitalem Eigentum im Bereich von KI. Da Eigentum und die Weitergabe von Eigentum ein wichtiger Aspekt der Generationengerechtigkeit ist, stellt dies einen zentralen Punkt des Positionspapiers dar.

Jährlich steigt die Anmeldung neuer KI-Patente (Shoham et al. 2018: 31); gleichzeitig schreibt sich die Tech-Industrie eine Open-Source-Kultur auf die Fahnen. Das Bestreben, Forschung und Innovation voranzutreiben, steht im Dilemma mit dem Anspruch auf Eigentum bzw. dem Anspruch auf die alleinige kommerzielle Nutzung von Technologien. In der Praxis kann dieses Dilemma verschiedene Formen annehmen.

Herausforderung: Urheberrechte von Weiterentwicklung von KI Algorithmen

Der KI-Künstler Robbie Barrat kreierte einen Algorithmus, der Gemälde selbstständig „malt“. Das Künstler-Trio „Obvious“ modifizierte diesen Algorithmus leicht, und versteigerte ein damit entstandenes Gemälde für \$432,000. Obwohl die Obvious-Künstler nicht die Urheber des Algorithmus waren, behielten sie das Anrecht auf den Kaufpreis. Der 19-jährige Urheber Robbie Barrat hatte das Nachsehen und wurde am Gewinn nicht beteiligt (Cohn 2018).

Wie können Urheberrechte an Algorithmen gerecht fair und nachhaltig gestaltet werden? Wie kann erreicht werden, dass zukünftige Entwickler genug Gestaltungsspielraum bei ihren (Weiter-)Entwicklungen haben?

Fest steht: die Entwicklung und Modifizierung von KI Software ist komplex, und die letztliche Verwendung und Kommerzialisierung oftmals unvorhersehbar. Es ist bislang noch unklar, welche Akteure ihre Ansprüche auf welche Weise geltend machen können; dies gilt für Fragen des Eigentums, der Anerkennung sowie der Beteiligung an der entstandenen ökonomischen Wertschöpfung. Momentan ist Software und somit auch jegliche Form von Algorithmen nicht direkt patentierbar. Zu restriktive Regelungen im Bereich des digitalen Eigentums verringern Anreize für Innovation und Weiterentwicklungen; der Mangel einer sicheren Gesetzesgrundlage birgt jedoch ebenfalls Herausforderungen. Es ist nicht nur wichtig, angemessene Regulierungen in diesem Zusammenhang festzulegen, sondern es sollte auch die Möglichkeit gegeben sein, sie an künftige Bedingungen anzupassen – zumal Veränderungen in diesem Sektor rasant schnell geschehen können. Denn Regelungen, die heute festgelegt werden, werden i.d.R. nicht nur heutige, sondern auch zukünftige Generationen von Entwickler*innen betreffen.

Einen kollaborativen Weg gehen Entwickler*innen und Unternehmen, die ihre Programme als Open Source und Datengrundlagen als Open Data für alle zugänglich machen, und an denen eine Community von freiwilligen Open Source Entwickler*innen mitarbeiten können

oder die Inhalte weiterverwenden kann. Eine Vielzahl von Firmen sind mit ihrem Open Source-Modell profitabel, obwohl das Kernprodukt kostenfrei angeboten wird. Beispielsweise können unterschiedliche kostenpflichtige Premium-Leistungen oder das Sammeln von begehrten Nutzer*innendaten die Profitabilität des Unternehmens sicherstellen. Da im Gegensatz zu konventionellen Firmen der Großteil des Codes und der Algorithmen offengelegt sind, wird diesem Modell ein demokratisierender Effekt und eine höhere Transparenz zugeschrieben (Bradford 2018)⁸. Vor allem Social Entrepreneurship und die Entwicklung neuartiger Ideen werden durch die niedrigen Eintrittsschwellen durch Open Source und Open Data profitieren.

Die Digitalisierung, und vor allem der Einsatz von Algorithmen zur ökonomischen Wertschöpfung birgt große Skaleneffekte: je mehr Daten jemand zur Verfügung hat oder besitzt, desto mehr Wertschöpfung kann daraus kreiert werden. Dies begünstigt wiederum die Bildung von Monopolen und Oligopolen (Bartlett, 2018). Dadurch findet in diesen Unternehmen oft auch die Entwicklung von Grundlagen statt, die nicht direkt wirtschaftlich verwertbar sind. Dies birgt die Gefahr, dass Markteintrittshürden deutlich erhöht werden können. Eine Überwachung von diesen nicht direkt wirtschaftlichen Vorteilen durch ein erweitertes Kartellrecht ist zu prüfen.

Wenn eine von KI erzielte Wertschöpfung wiederkehrend ökonomische Wertschöpfung erzielt (und somit Produktivität steigert), kann sie gleichzeitig bestimmte menschliche Aufgaben überflüssig machen. So müssen neue Lösungen dafür gefunden werden, wie die gesamte Gesellschaft, also auch die „Verlierer“ der „KI-Revolution“ an den Vorteilen beteiligt werden. Es ist also sicherzustellen, dass das digitale Eigentum welches einigen wenigen gehört, bei wiederkehrender Wertschöpfung in angemessenem Maße der Gesamtheit dienen kann. Beispielsweise wird die Erwerbsarbeit von LKW-Fahrer*innen zum Allgemeinwohl der Gesellschaft versteuert. Wenn solche Logistikarbeit in Zukunft von KI-gesteuerten autonomen Fahrzeugen verrichtet wird, muss auch diese Wertschöpfung entsprechend versteuert werden – sonst ist die Finanzierbarkeit der sozialen Sicherungssysteme und anderer staatlicher Leistungen in Gefahr. Politische Akteure müssen deshalb die fiskal- und sozialpolitischen Implikationen von KI ganzheitlich und vorausschauend evaluieren. Zusätzliche Herausforderungen entstehen durch stark globalisierte Wertschöpfungsketten. Dies wirft weitere Herausforderungen dazu auf, wie und wo die verschiedenen Etappen der KI-Wertschöpfung versteuert werden.

Auch der Umgang mit öffentlichen und personenbezogenen Daten führt zu Herausforderungen. Daten sind ein Rohstoff, der große finanzielle Gewinne generieren kann, beispielsweise durch Werbeeinnahmen, persönliches Profiling oder auch zur Nutzung in der politischen Kommunikation. Es ist noch offen, wie Privatpersonen einerseits vor

⁸ Ein Beispiel für ein gelungenes Open Source Projekt, wenn auch nicht mit explizitem Einsatz von KI Technologie, ist die deutsche Corona Warn-App von SAP im Auftrag der deutschen Bundesregierung. Durch die lückenlose Veröffentlichung des Codes konnte vor allem ein hohes Level an Sicherheit und Datenschutz erreicht werden (SAP 2020).

einem Missbrauch ihrer Daten geschützt (auch wenn die EU Datenschutz-Grundverordnung hier erste Schritte geht, siehe auch Kapitel 2.3), und andererseits an den Gewinnen, die durch ihre eigenen Daten entstehen, beteiligt werden können. Auch auf den politischen bzw. gesellschaftlichen Diskurs kann KI massiv einwirken; dies wird deutlich anhand der aktuellen Fake News Debatte oder zahlreicher Wahlen in den vergangenen Jahren, bei denen die Einflussnahme von außen angenommen wird – beispielsweise die US Präsidentschaftswahlen 2016 (Bartlett 2018), die Bundestagswahl 2017 (Zeit Online 2017) sowie die Europawahlen 2019.

Insgesamt ist festzuhalten, dass KI eine disruptive Technologie ist, die weitreichende Auswirkungen auf unser Wirtschaftssystem haben wird. Historisch gesehen wurde dann der meiste und am besten verteilte Wohlstand generiert, wenn sowohl wirtschaftliche als auch politische Institutionen inklusiv ausgelegt waren (Acemoglu und Robinson 2012: 126ff). Da dies jedoch nicht automatisch sichergestellt ist, muss es aktiv gestaltet werden. Bahnbrechende Technologien wie KI bergen neben großen Chancen aber auch das Risiko, das Wirtschaftssystem von einem relativ inklusiven zu einem verstärkt exklusiven zu transformieren – wenn nicht rechtzeitig nötige Änderungen in den Gesetzen und Regularien vorgenommen werden. Im Rahmen von KI handelt es sich hierbei sowohl um private Daten als auch die von Algorithmen generierte Wertschöpfung.

Forderung bezüglich des digitalen Eigentums an künstlicher Intelligenz:

- Eigentum und die Weitergabe von Eigentum sind ein wichtiger Aspekt der Generationengerechtigkeit, gleichzeitig ist Produktivitätssteigerung immer mehr an Ressourcenbesitz, und weniger an menschliche Arbeit gebunden. Digitales Eigentum, welches einigen wenigen „gehört“, und wiederkehrende **Wertschöpfung** erzeugt, muss – ähnlich wie auch steuerpflichtige menschliche Arbeitskraft – in angemessenem Maße der **Gesellschaft** dienen. Hierbei muss vor allem geklärt werden, was künftig die Bemessungsgrundlage für Steuern sein soll.
- Um den meisten Wohlstand zu generieren, müssen wirtschaftliche und politische **Institutionen inklusiv** ausgelegt sein.
- Eine **angemessene Beteiligung** an der (Weiter-) Entwicklung von KI muss möglichst vielen Beteiligten ermöglicht werden. Open Data Ansätze sollten deshalb gefördert werden – auch, um die Entwicklung von Lösungen zu unterstützen, die gesellschaftliche Herausforderungen lösen können.
- Eine Ausweitung der **kartellrechtlichen Überwachung** auf den Bereich der nicht direkt wirtschaftlichen Algorithmen sollte geprüft werden um Monopole zu verhindern.

2.3 Transparenz von Algorithmen

Wenn wir etwas in die Google-Suchmaschine eingeben, erscheinen einige Ergebnisse vor anderen. Wenn wir Facebook öffnen, erscheint in unserer Timeline eine bestimmte Reihenfolge. Ein Bewerber*innen-Matching-Tool empfiehlt bestimmte Bewerber*innen anstelle von anderen Kandidat*innen. Oder vorausschauende Polizeimaßnahmen (engl.: predictive policing) schreiben bestimmten Menschen eine höhere Wahrscheinlichkeit für kriminelle Taten zu. Die Anordnung und Auswahl erfolgt über Algorithmen, die sich Künstlicher Intelligenz bedienen (Lenzen 2019: 163). Doch wie sollen Nutzer*innen oder Entscheider*innen verstehen, wie solche Anordnungen von Ergebnissen entstehen, sofern man diese nicht selbst programmiert hat? Und wie können sie gegen Benachteiligung vorgehen, wenn sie gar nicht davon wissen? Es muss also Transparenz geschaffen werden, um Algorithmen verständlich und erklärbar zu machen, denn schließlich nehmen diese einen großen Teil unseres Lebens ein (Zierer 2018). Dabei reicht es nicht aus, die Algorithmen einfach offenzulegen, denn dieser komplizierte Quellcode ist für Allgemeinverbraucher*innen unverständlich und kann nicht gelesen oder gar verstanden und interpretiert werden - auch weil in vielen KI-Anwendungen mehrere Algorithmen miteinander interagieren. Zudem sind auch Algorithmen digitales Eigentum und können nicht immer einfach veröffentlicht werden (s. Kapitel digitales Eigentum). Auch die Daten, die zum Trainieren dieser Algorithmen verwendet wurden, sind üblicherweise nicht einsehbar.

Um KI-Anwendungen, die sich verschiedener Algorithmen bedienen, verständlich und transparent zu machen, benötigt man also bestimmte Kriterien und rechtliche Rahmenbedingungen, nach denen sie gegliedert und bewertet werden. Doch eine übergeordnete und unabhängige Prüfstelle, die solche Kriterien festlegen und KI-Anwendungen überprüfen könnte, gibt es nicht. Informatiker wie Viktor Mayer-Schönberger forderten bereits 2013 einen sogenannten „Algorithmus-TÜV“ (Fischermann und Hamann 2013), um eine Instanz zu schaffen, die KI-Anwendungen transparent und erklärbar macht. Vor allem bei neueren Verfahren wie Deep Learning ist allerdings eine komplette Prüfung des Algorithmus und Nachvollziehbarkeit selbst für eine solche Instanz nicht ohne weiteres möglich. In den letzten Jahren hat sich das Forschungsfeld der Erklärbaren KI (englisch: Explainable AI) gebildet, bei dem versucht wird den Zusammenhang zwischen Eingangs- und Ausgangsgrößen herzustellen. Dieses Gebiet der KI Forschung sollte in Kombination mit Maßnahmen zur Steigerung des allgemeinen KI Verständnisses gefördert werden.

Herausforderung: Wie kann jede*r Einzelne Algorithmen bewerten?

Es könnte die Aufgabe einer Art Prüfstelle sein, KI-Anwendungen anhand von festgelegten Kriterien zu beurteilen. Jedoch sind die persönlichen Gewichtungen der Bewertungskriterien hierbei nicht für jede Person gleich. Somit muss jeder für sich selbst Wissen und Kenntnisse über Algorithmen haben, um die sachlichen Urteile auch für sich selbst persönlich analysieren und bewerten zu können (Trottenberg 2016). Durch die

Sortierung und Anordnung der Inhalte durch Algorithmen befindet man sich schnell in der sogenannten Filterblase (Ovens 2017). So werden Inhalte, die einem potentiell gefallen könnten angezeigt und da man auf diese dann immer wieder reagiert, entwickelt sich eine Schleife, sodass häufig nur noch Inhalte dieser Kategorie angezeigt werden. So werden einer 15-Jährigen, die sich für Mode oder Computerspiele interessiert keinerlei politischen Inhalte angezeigt, da diese nicht Bestandteil ihrer Filterblase sind. Daher müssen Algorithmen transparent behandelt werden, um herausfinden zu können, wieso welche Inhalte angezeigt werden und welche möglichen Inhalte es außerhalb der eigenen Filterblase gibt (Lenzen 2018: 163).

Bei einer Analyse von Algorithmen muss zudem transparent gestaltet sein, welche und auf welche Art und Weise persönliche Daten verwendet werden – sodass ein Bewusstsein dafür geschaffen wird, wofür die eigenen Daten verwendet werden können. Ein Grundstein hierfür ist bereits in der Datenschutz-Grundverordnung der EU Art. 15 gelegt (Europäische Kommission 2016b). Diese gibt betroffenen Personen das Recht auf:

- Auskunft darüber wann eine automatisierte Entscheidungsfindung stattgefunden hat,
- Aussagekräftige Informationen über die involvierte Logik, sowie
- Auskunft über die Tragweite der Entscheidungen.

Wie dies konkret gewährleistet werden kann, muss erarbeitet werden. Die Offenlegung der Algorithmen soll digitale Eigentumsrechte nicht verletzen, jedoch wird eine Transparenz benötigt, um Erkenntnisse über Entscheidungsprozesse zu gewinnen. Durch diese Transparenz sollen vor allem auch Ungleichheiten, die durch Algorithmen entstehen können, verhindert werden (Dornis 2018). Eine effektive Anonymisierung von persönlichen Daten wird durch die steigende Individualisierung und wachsende Menge an historischen Daten zunehmend schwieriger (Fry 2018) – sie ist und bleibt jedoch essenziell.

Eine grundlegende Arbeitsweise von KI-Algorithmen ist das maschinelle Lernen, also das Nutzen von Daten, um durch statistische Korrelationen Aussagen über neue Anfragen zu treffen. Sie basieren somit auf Daten aus der Vergangenheit. Diese Daten können unvollständig sein (vgl. Criado-Perez 2019) oder Einstellungen und Lebensweisen repräsentieren, welche dann bei zukünftigen Fragestellungen oder Anwendungen reproduziert werden. Allerdings sind unsere Werte und Einstellungen – beziehungsweise die der Entwickler*innen – zum einen subjektiv und voreingenommen. Zum anderen entwickeln sie sich kontinuierlich weiter. Somit kann dies theoretisch in einer Sackgasse von sich immer wiederholenden, gleichen Schleifen enden. Bestehende strukturelle Ungleichheiten und Diskriminierungen verfestigen und reproduzieren sich so. Je festgefahrener solche Ungleichheiten werden, umso geringer werden die Einflussmöglichkeiten künftiger Generationen.

Doch Transparenz und Erklärbarkeit von Algorithmen alleine reichen noch nicht aus. Selbst ausgeglichene und alle demographischen Gruppen repräsentierende Datensätze garantieren noch nicht, dass die daraus lernende KI-Anwendung nicht diskriminiert. Zusätzlich zur Transparenz und Erklärbarkeit müssen die Ergebnisse – sprich der Output – von KI-Anwendungen gründlich getestet werden, um zu verstehen, ob sie diskriminierend wirken, und wenn ja, wie dem entgegengewirkt werden kann (Barton et al. 2019).

Für die Transparenz von Algorithmen ist es weiterhin essenziell, dass die Entscheidungsgrundlage einsehbar ist. Neben der Logik des Algorithmus sind hierfür die verwendeten Daten wichtig. Und diese liegen vermehrt nur noch Großkonzernen vor, die sich diese monopolähnlich aufteilen (Moorstedt 2019). Die Konzentration von persönlichen und nicht-persönlichen Daten bei einzelnen Akteuren ist unvorteilhaft und kreiert Abhängigkeiten. Diese Abhängigkeiten wiederum verstärken den Prozess der Konzentration und haben unter Umständen exponentiell steigende Auswirkungen und sind somit für zukünftige Generationen besonders belastend.

„Eigentum verpflichtet“ (Grundgesetz, Artikel 14, Absatz 2). Bei physischem Eigentum ist dies bekannt und wird zumindest sporadisch eingefordert. Der Artikel des Grundgesetzes ist aber auch auf digitales Eigentum anwendbar. Dateneigentum verpflichtet somit auch zu einem Einsatz zum Wohle der Gesellschaft. Um dies evaluieren zu können ist es essenziell, bei gesellschaftlichen Entscheidungen die Datengrundlage transparent und einsehbar zu halten. Zivilgesellschaftliche Akteure wie die Open Source Business Alliance setzen sich für den verstärkten Einsatz von Open Source Lösungen ein (OSB Alliance 2019). Ähnliche Lösungen für Daten sind wünschenswert und sollten für alle öffentlich-geförderten und staatlich-politischen Entscheidungen verpflichtend sein. Um Transparenz zu gewährleisten ist zudem wesentlich, dass dem Einzelnen explizit offengelegt wird, wann er mit einem künstlich intelligenten Akteur (beispielsweise ein Chatbot) kommuniziert und interagiert oder wenn Inhalte von intelligenten Systemen selektiert werden. Nur so ist eine kritische Einordnung der Entscheidungsgrundlage möglich.

Forderungen im Bereich Transparenz von Algorithmen:

- Juristische Prüfung der Möglichkeit und anschließend Entwicklung einer **unabhängigen, übergeordneten Prüfstelle**, die Algorithmen anhand Kriterien und rechtlichen Rahmenbedingungen bewertet und Zugriff auf staatliche und nicht-staatliche Akteure hat.
- Art. 15 der **europäischen Datenschutz-Grundverordnung** muss effektiv umgesetzt werden. Diese gibt betroffenen Personen sowohl das Recht auf Auskunft darüber wann eine automatisierte Entscheidungsfindung stattgefunden hat als auch auf aussagekräftige Informationen über die involvierte Logik sowie die Tragweite der Entscheidungen. Es muss Menschen explizit offengelegt werden, wann sie mit künstlich intelligenten Akteuren kommunizieren.

- **Transparenz** bezüglich der **Eingangsdaten** muss geschaffen werden um die Tendenz zur Fokussierung auf vergangene Ereignisse adressieren zu können. Um der Reproduzierung struktureller Diskrimination vorzubeugen, sind diverse, interdisziplinäre und intersektionale Teams notwendig.
- Die **Ergebnisse** von KI-Anwendungen müssen im Hinblick auf potenzielle **Diskriminierung getestet** werden, um ihren tatsächlichen Effekt bewerten zu können.
- Der voranschreitenden Monopolisierung von Daten muss jetzt entgegengewirkt werden um unumgängliche Spätfolgen zu vermeiden. **Open Data** soll aktiv gefördert werden, damit auch Personengruppen ohne primär wirtschaftliche Ausrichtung einen Mehrwert beitragen können.
- Für Algorithmen die in staatlichen und staatlich geförderten Einrichtungen entwickelt oder verwendet werden müssen **Quellcodes** sowie aussagekräftige Erklärungen **offengelegt** werden.

2.4 Diskurs zu KI schaffen

Die Geschwindigkeit mit der KI in alle Bereiche unserer Gesellschaft einzieht und diese verändert, kreiert große Herausforderungen. Auf diese müssen Gesetzgeber, Administration, Wissenschaft und Bildung, Wirtschaft und letztlich unsere gesamte Gesellschaft vorbereitet werden (Bitkom 2017: 36). Stephen Hawking geht sogar so weit, dass unser Umgang als Gesellschaft mit dem Thema darüber entscheidet ob die Entwicklung eine positive oder negative wird (vgl. Zitat am Anfang des Positionspapiers, Hawking 2017).

Als Gesellschaft haben wir die Pflicht, durch verschiedene Organe dafür zu sorgen, dass KI eine positive Entwicklung antreibt. Hierbei muss ein besonderer Fokus auf den Generationen liegen, die am längsten mit den Folgen leben müssen. Gleichzeitig sind junge Menschen aber nicht nur demografisch zahlenmäßig unterlegen, sondern momentan auch in sämtlichen politischen Entscheidungsgremien unterrepräsentiert – Menschen unter 30 stellen beispielsweise nur 2% aller Parlamentsabgeordneten weltweit dar (Inter-Parliamentary Union, 2018). Somit müssen besonders junge Menschen deutlich aktiver in den Diskurs eingebunden werden, und ihn an vielen Stellen anführen und mitgestalten. Denn fundamentaler technologischer Wandel geht immer auch mit einem grundlegenden gesellschaftlichen Wandel einher (Bitkom 2017). Die Einführung von speziellen Institutionen und Quoten für junge Menschen in der Mitbestimmung wäre zu begrüßen. Ein entsprechender KI-Ethikrat sollte diesen Wandel moderieren; die Ausgestaltung und Platzierung in der politischen und gesellschaftlichen Landschaft ist noch zu prüfen.

Best Practice: Urheberrechte von Weiterentwicklung von KI Algorithmen

In der kanadischen Regierung gibt es ein fest verankertes Komitee aus 16-20-jährigen, die bei digitalen Fragestellungen konsultiert werden (ab der nächsten Komiteeperiode bis 26),

das Digital Youth Advisory Committee. Die Mitglieder des Komitees treffen Regierungsmitglieder in regelmäßigen Abständen zum gegenseitigen Austausch. Die Regierung bekommt Einblicke in die Belange junger Generationen und die jugendlichen Mitglieder bekommen früh wertvolle Einblicke in das politische System (Government of Canada 2019). Der Austausch ist verbindlich, Handlungen müssen daraus allerdings nicht entstehen. Aber eine institutionalisierte, gemeinsame Diskussion ist bereits ein großer Schritt der in vielen Nationen noch aussteht.

Ein inhärentes Merkmal von Künstlicher Intelligenz ist, dass sie spezialisiertes Fachwissen benötigt, was in Gänze von einem einzelnen Menschen nur grundlegend verstanden werden kann. Damit wird die in unserer Gesellschaft vorherrschende Spezialisierung noch weiter vorangetrieben. Für KI wird dies vor allem eine Veränderung der Arbeitswelt bedeuten, aber auch den Einzug von Algorithmen in sämtliche Bereiche des privaten Lebens. Diese Veränderungen sind mit großen Unsicherheiten behaftet; somit braucht es großes Vertrauen in die Gestalter*innen von KI. Ein tiefgehender Diskurs ist deshalb wichtig, um das nötige Vertrauen in die Mitmenschen oder auch staatlichen Institutionen die sich mit KI befassen, zu schaffen.

Best Practise: Frühe und breite Teilhabe an Fachwissen zu Künstlicher Intelligenz

Die Baden-Württemberg Stiftung hat Ende 2018 angekündigt ihre Fördertöpfe für die Ausbildung junger Menschen im Bereich KI zu vergrößern. Dies beinhaltet die Förderung von Fortbildungen, aber auch die Ausbildung von KI Schulbegleiter*innen (Baden-Württemberg Stiftung 2018). Ebenfalls Ende 2018 hat das Max Planck Institut für Intelligente Systeme einen Bundeswettbewerb Künstliche Intelligenz ausgerufen. Durch den Kontakt mit dem Thema können Schüler*innen selbst zu Multiplikator*innen werden.

Die finnische Ratspräsidentschaft der EU (Juli-Dezember 2019) hat sich zum Ziel gesetzt bis 2021 1% der europäischen Bevölkerung zu den Grundlagen künstlicher Intelligenz auszubilden (Elements of AI, 2018). Solche Initiativen sind lobenswert und haben das Potenzial, eine notwendige Grundkompetenz für einen breiten gesellschaftlichen Diskurs zu schaffen.

Seit November 2018 hat Deutschland eine offizielle und transparent kommunizierte KI-Strategie (Bundesregierung Deutschland 2018). Weiterhin gibt es mit der Datenethikkommission (Bundesministerium des Inneren, für Bau und Heimat 2019) ein Organ, das für die Erstellung eines Leitfadens und ethischer Richtlinien verantwortlich ist. Dies ist ein guter Anfang, um die verfolgten Ziele transparent zu kommunizieren. Allerdings wird es beim Thema KI wichtig sein, einen gesamtgesellschaftlichen Diskurs zu schaffen und bei den Diskussionen möglichst niemanden zurückzulassen. KI könnte sonst zur Elitentechnologie werden, da ein hoher Grad von Spezialisierung und Bildung vorausgesetzt wird. Eine wichtige Voraussetzung dafür ist eine grundlegende Bildung im Bereich digitaler

Technologien. Darüber hinaus ist es dennoch wichtig den Austausch auf verschiedenen Ebenen anzuregen (z.B. Schule, Ortsgruppen, Massenmedien oder Diskussionsveranstaltungen).

Die KI Strategie Deutschland sieht KI-Trainer*innen in den Mittelstandszentren für Unternehmen vor (Bundesregierung Deutschland 2018, p.23). Dies ist die einzige verankerte Transferstrategie für KI Wissen. Eine gesetzliche Verankerung für den Diskurs und das Training bezüglich KI in der Gesellschaft wären wünschenswert. Die lebhafte Diskussion über KI ist eine einmalige Chance, unsere ethischen und gesellschaftlichen Ideale in die Diskussion um Künstliche Intelligenz einzubringen.

Forderungen zur Stärkung des Diskurses zu Künstlicher Intelligenz:

- Künstliche Intelligenz ist ein Zukunftsthema, dessen volle Ausmaße nicht abzuschätzen sind. **Junge Menschen** müssen aktiv in ihrer **politischen Teilhabe** unterstützt werden, z.B. durch spezielle Institutionen und Quoten für junge Menschen bei der Mitbestimmung.
- Ein **Austausch** über KI muss **auf verschiedenen Ebenen** angeregt werden - z.B. Schule, Ortsgruppen, Massenmedien.
- Konkrete **Förderprogramme** können dabei helfen, (jungen) Menschen Zugang zu und Wissen über die Technologie zu ermöglichen, und sie so in einem gesellschaftlichen Diskurs einzuschließen. Ein essenzieller Beitrag zu einem mündigen Diskurs ist die Bereitstellung entsprechender Bildungsangebote.
- Zur Institutionalisierung des Diskurses soll ein **KI-Ethikrat** geschaffen werden oder ein bestehender Rat um diese Kompetenz erweitert werden.

2.5 Zukunftsinvestition KI

Ein massives Investment in KI wird in der nahen Zukunft nötig sein, damit Deutschland im vorherrschenden Wettbewerb bestehen kann (Döbel et al. 2018; Mercator Institute for China Studies 2018). Im oben angesprochenen Diskurs wird zu erörtern sein, wo wir als Gesellschaft unsere Rolle in der Erforschung, Anwendung und Bewertung von KI sehen.

Deutschland hat in den frühen Phasen der KI Forschung durch seine Arbeiten, vor allem im Feld der Support Vector Machines, wertvolle Arbeit geleistet. Aktuell sind die Vereinigten Staaten von Amerika und China, auch durch andere Förderbedingungen, hierbei deutlich weiter und haben Deutschland abgehängt (Döbel et al. 2018). Bei der Bereitstellung von Nutzer*innendaten für die Anwendung von KI kristallisiert sich immer mehr China als absoluter Primus (aus Sicht von KI Anwender*innen) heraus⁹. Durch wenig restriktive Datenschutzgesetze und die umstrittene Einführung des „Social Credit Systems“ (einem

⁹ Die Einordnung als Primus geschieht hier unter den rein technischen Aspekten der Datenverfügbarkeit zum Training der Algorithmen. Eine ethisch-moralische Einordnung findet bewusst nicht statt.

Datenpool über das Verhalten chinesischer Bürger*innen) und die dadurch vorhandene Datenfülle hat sich vor allem die Region um Shenzhen zur Vorreiterin in der KI Anwendung gemauert (Mercator Institute for China Studies 2018). Hier könnte allerdings Deutschland als starke Industrienation mit starken gesellschaftlichen und Traditionen im Hinblick auf den Schutz des Individuums als Vorreiter fungieren. Mit seinen vielen großen und aber vor allem auch kleinen und mittelständischen Unternehmen hat Deutschland hervorragende Voraussetzungen in der industriellen Anwendung von KI. Hier plädieren wir für eine bewusste und von ethischen Überlegungen geleitete Entscheidung für einen klaren Fokus, um nicht zukünftig im Wettbewerb unterzugehen. Weitergedacht könnte auch um Deutschland herum eine starke europäische Allianz zur Anwendung von KI im Industrieumfeld entstehen. Die deutsche EU Ratspräsidentschaft (Juli - Dezember 2020) bietet hier Chancen.

Konkret sind jetzt Investitionen wichtig, um die Reproduktion struktureller Diskriminierung einzuschränken. Dies kann nur mit Hilfe von diversen, interdisziplinären und intersektionalen Ansätzen gewährleistet werden (siehe: Kapitel 2.3 unter Bias). Im Besonderen sollte die Steuerung des Diskurses bezüglich moralischer und ethischer Fragestellungen im Bereich KI nicht privatwirtschaftlichen und unzureichend divers aufgestellten Förder*innen überlassen werden. Ein Negativbeispiel wäre hier die massive Förderung des Institute for Ethics in Artificial Intelligence der TU München durch Facebook (Dornis 2019).

Das von der Bundesregierung angestrebte Siegel „AI made in Germany“ ist ein wichtiger Baustein, welches allerdings noch mit praktischen Inhalten gefüllt werden muss. Vor allem die deutlichen Investitionen in Schul- und Hochschulbildung (Bundesregierung Deutschland 2018, 15 und 31) können hierzu einen wichtigen Beitrag leisten.

Ein weiterer wichtiger Punkt ist jetzt sicherzustellen, dass die digitale Infrastruktur zukunftsfähig ist. Zum einen bedeutet dies Investitionen in Bildung, Technik und Cyber-Sicherheit. Zum anderen bedeutet es aber auch, dass KI ressourcenschonend und klimaverträglich in der Anwendung und Ableitung von Maßnahmen/Ergebnissen ist (vgl. WBGU 2019: 37). Gleiches gilt für die Produktion der technischen Maschinen und Computer. Es ist nachgewiesen, dass digitale Technologien deutliche Umweltauswirkungen hervorrufen (Morley et al. 2018), trotzdem gibt es zu den Auswirkungen von KI auf Umwelt und Ressourcenverbrauch bisher kaum quantifizierte Informationen. In Zukunft könnten die nur sehr begrenzt verfügbaren seltenen Erden und andere Rohstoffe zum Flaschenhals bei der Erweiterung der technischen Infrastruktur werden. Eine Möglichkeit, die begrenzten Ressourcen unseres Planeten zu berücksichtigen und ebenso die Bedürfnisse zukünftiger Generationen, ist die Bewertung nach den planetaren Grenzen (engl.: Planetary Boundaries) (Stockholm Resilience Center 2019) auf welche der Abbau und die Verarbeitung der benötigten Materialien direkten oder indirekten Einfluss haben. Denn wenn heute alle zur Verfügung stehenden Ressourcen aufgebraucht werden, können diese für zukünftige technische Lösungen nicht mehr eingesetzt werden. Eine andere Ebene der Diskussion um die Nachhaltigkeit des Einsatzes von KI bildet die Frage nach den

Anwendungen. Im Sinne des ethisch und generationengerechten Einsatzes von KI ist es wünschenswert, dass KI dafür eingesetzt wird, nachhaltigere Lösungen zu schaffen. Die Nachhaltigkeit bezieht sich hierbei vor allem auf die ökologische und soziale Dimension. Beispiele hierfür wären der Einsatz von KI zur Optimierung von Transportrouten um damit Emissionen zu vermeiden oder die Übernahme von für Menschen gefährlichen Arbeiten. Etwaige Reboundeffekte müssen in die Bewertung der Maßnahmen einfließen.

Der nachhaltige Einsatz von KI bedeutet auch, dass die Resilienz – also die Widerstandsfähigkeit gegenüber äußerer Einflüsse – aller beeinträchtigten Gebiete, Themenfelder und Personen gewährleistet wird. Besonderen Einfluss auf die Resilienz könnte der Einsatz von KI in der Versorgung (z.B. Stromversorgung, Kommunikationsnetze) haben (vgl. Germanwatch 2019). Auch die psychische Auswirkung auf den Menschen durch den Einsatz von KI bzw. die Vorstellung davon sollte betrachtet werden (siehe Kapitel 3 und vgl. Auerbach 2014).

Die menschliche Resilienz wird vor allem durch den gezielten Einsatz von Bildung gefördert. Die wirtschaftliche und gesellschaftliche Resilienz kann durch gute Spielregeln bezüglich des digitalen Eigentums gefördert werden sowie einer verantwortungsvollen und nachhaltigen Verteilung der Kosten und Nutzen von KI. Am schwersten einzuschätzen ist die Versorgungsresilienz unter Anwendung von KI: Der Einsatz von Algorithmen, zum Beispiel bei der Regelung des Strommarktes, kann natürlich menschliche Fehler eliminieren. Allerdings bedingt ihr Einsatz auch die Abhängigkeit von ihnen und potentiellen systematischen Fehlern. Diese Auswirkungen und sogar ihre Existenz sind bis zum Eintreten des Fehlers selbst oft nicht vorhersehbar. Detaillierte Ausführungen hierzu finden sich bei Zimmermann und Frank (2019). Um die Abhängigkeit von Algorithmen und Künstlicher Intelligenz zu minimieren, sollen vor allem Forschungsvorhaben die Mitbetrachtung potenzieller systematischer Fehler fördern und fordern. Eine spezielle Ausschreibung des Ministeriums für Bildung und Forschung zum Thema Künstliche Intelligenz in Kommunikationsnetzen berücksichtigt bereits genau dies, indem sie die „Steigerung ihrer Leistungsfähigkeit, Resilienz sowie ihrer Ressourcen- und Energieeffizienz“ (BMBF 2018) fordert.

Forderungen für die Zukunftstauglichkeit von KI:

- Um in Zukunft eine **signifikante Stimme** bei der ethisch-korrekten Anwendung und Entwicklung von KI zu haben, muss sich Deutschland jetzt gut bei der Anwendung und Forschung von KI platzieren und europäische sowie internationale Partnerschaften ausbauen.
- Um in Zukunft im Feld der Künstlichen Intelligenz mitgestalten zu können, muss Deutschland sich mit **gleichgesinnten Partner*innen** verbünden. KI ist keine lokale Herausforderung, sondern eine globale.
- KI muss dazu eingesetzt werden, eine **nachhaltige Entwicklung** zu unterstützen. Dies beinhaltet die Entwicklung einer nachhaltigen digitalen Infrastruktur. Dabei ist

zu beachten, dass KI Nachhaltigkeit fördern soll und dabei selbst nachhaltig zu gestaltet ist, unter anderem durch eine Bewertung innerhalb der planetaren Grenzen.

- Die Förderung von Projekten und Initiativen, die **ethische Aspekte** von KI behandeln muss unbedingt **unabhängig** gestaltet sein.

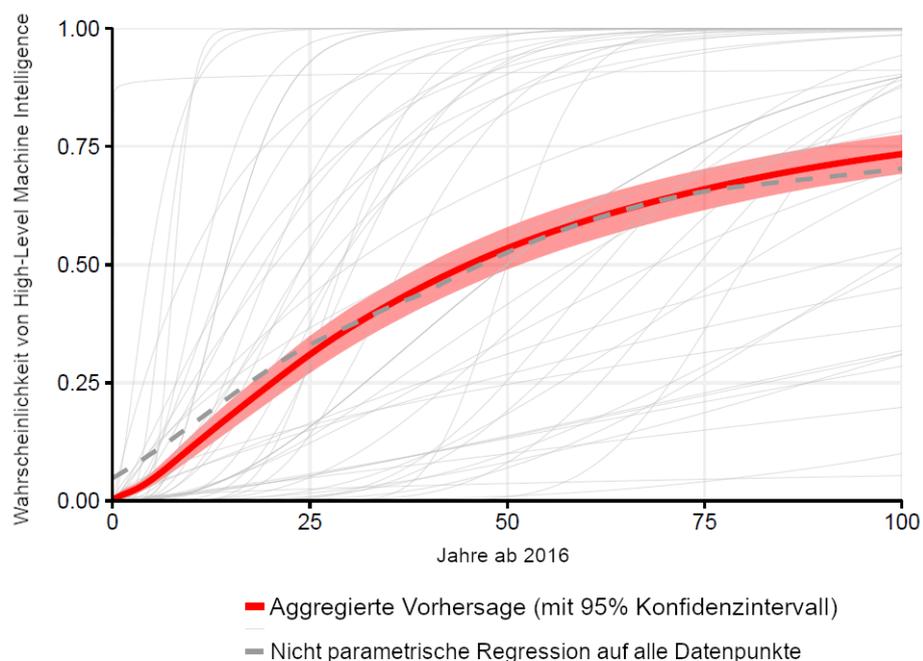
3. Starke KI als Gefahr für zukünftige Generationen

Heute basieren KI-Anwendungen primär auf den Prinzipien der schwachen KI (Germanwatch 2019: 12). Das bedeutet: KI unterstützt den Menschen beim Lösen von bestimmten Aufgaben und erzielt dabei hauptsächlich Geschwindigkeitsvorteile. Nach dem Moravec Paradoxon (Moravec, 1988) lässt sich KI auf Grund seiner Struktur zum Erkenntnisgewinn dabei besonders gut für Aufgaben einsetzen, die dem Menschen schwerfallen – aber für den Menschen einfachere Aufgaben, können für KI unlösbar werden. In Zukunft könnte diese Barriere durch steigende Rechenleistung und den vermehrten Einsatz neuronaler Netze beim Deep Learning überwunden werden. So könnten ganz neue Formen der KI entstehen (vgl. WBGU 2019: 75-76). Diese Formen der starken KI finden sich im heutigen Diskurs in verschiedenen Ausprägungen wieder.

Nick Bostrom (2014) zum Beispiel rechnet im Jahr 2105 mit der ersten Superintelligenz basierend auf KI oder der Whole Brain Simulation (vgl. Stiefel and Brooks 2019, die parallele Forschungsrichtung, bei der die Funktionsweise des menschlichen Gehirns nachgebildet werden soll, wird hier nicht weiter ausgeführt). Bostrom vergleicht in seinem Buch die Entstehung der Superintelligenz mit den global disruptiven Ereignissen der Weltgeschichte (Erkenntnis Feuer zu entzünden, Einführung der Landwirtschaft, Industrialisierung, etc.) (Bostrom 2014). Als besonders wichtig erachtet Bostrom die Frage, wer es schafft, die Superintelligenz zu entwickeln und wer demnach die Entscheidungsgewalt darüber hat. Domingos (2015) sieht die Entstehung einer starken KI in der Zusammenführung bisheriger spezialisierter Algorithmen zu einem Masteralgorithmus. In seiner Vision ist es nur eine Frage der Zeit, bis eine Drittinstitution bisherige Algorithmen zusammenführt. In der Version von Kurzweil (2006) kommt es zu einem bestimmten Zeitpunkt zu der unumkehrbaren und unvorhersehbaren Veränderung durch die Singularität – der Punkt an dem Maschinen und Algorithmen sich so schnell selbst optimieren, dass sie dem Menschen überlegen sind, und dieser sie nicht mehr abschalten kann. Die Singularität steht bei Kurzweil am Ende einer Evolution – sie beginnt bei der biologischen Evolution (durch die Natur), führt über die technische Evolution (durch den Menschen) und endet in einer Fortführung der technischen Evolution durch intelligente Maschinen (Kurzweil 2006). Diesen Zeitpunkt sieht er etwa im Jahr 2045. Im Dataismus nach Brooks (2013) wird der Informationsfluss zum höchsten Gut des fortschreitenden Anthropozän und bei Harari (2015) verändert sich dadurch in Kombination mit der absoluten Selbstoptimierung unser Menschenverständnis so weit, dass die neue Gattung des Homo Deus entsteht.

Die Abschätzung der Auswirkungen von KI wird schwieriger je weiter man in die Zukunft blickt. Somit besteht hier besondere Relevanz für und Verantwortung gegenüber zukünftiger Generationen. Besonders der Zeitpunkt der Entwicklung einer starken KI (siehe 1.2) könnte hier ein Kipppunkt sein. Grace et al. (2018) haben Expert*innen zu der Wahrscheinlichkeit des Eintretens starker KI¹⁰ befragt (siehe Abbildung 1). Die Schätzungen zum Zeitpunkt des Eintretens einer starken KI streuen stark. Aber in 100 Jahren (ab 2016) gehen die Expert*innen zu 75% von der Existenz einer starken KI aus. Eine Gruppe von Expert*innen mit ähnlicher Expertise wurde in Müller und Bostrom (2014) zu den möglichen Auswirkungen befragt. 53% der Befragten gehen von extrem guten oder guten Auswirkungen aus und 31% von extrem schlechten oder schlechten. In Kombination mit dem Ergebnis der selben Umfrage, dass starke KI sehr wahrscheinlich aus den schon existierenden und beforschten Technologien der KI entstehen wird, leitet sich die Anforderung ab, schon jetzt die Entwicklungen in der KI so zu regulieren, dass es in der Zukunft nicht zu schlechten oder extrem schlechten Auswirkungen kommen kann, egal ob für zeitlich nah oder fern lebende Generationen.

Abbildung 1 Wahrscheinlichkeit der Erschaffung einer starken KI (High-Level Machine Intelligence) in Abhängigkeit von den vergangenen Jahren nach Expert*innenbefragung (Grace et al. 2018)



¹⁰ Anm.: In Grace et al. (2018) wird diese neutraler als High-Level Machine Intelligence bezeichnet.

Herausforderungen: Unsachgemäßen militärischen Einsatz von KI gemeinschaftlich verhindern

Im militärischen Kontext wird KI vor allem im Bereich von autonomen Waffensystemen und in der Lagebeurteilung eingesetzt. In der Lagebeurteilung wird Bilderkennung verwendet, um Gefahrenlagen zu evaluieren und im besten Falle schneller zu handeln als es Menschen möglich wäre. Im schlechten Falle eskaliert eine ruhige Lage durch eine Fehleinschätzung (Welchering 2019).

Bei der Definition von autonomen Systemen gibt es noch keinen Konsens darüber, wie viel menschliche Interaktion nötig ist, damit ein System ‚nur‘ als teilautonom gilt. Im extremsten Fall reicht das Einschalten durch einen Menschen dazu aus; Angriffsentscheidungen werden dann nur noch von Algorithmen gesteuert (Welchering 2019). Zur Regulierung dessen gibt es Bestrebungen einiger Staaten im Rahmen der UN-Waffenkonvention. Allerdings ist die Staatengemeinschaft noch weit von einem Konsens entfernt. Seit der letzten Konvention im November 2019 gibt es aber zumindest Leitprinzipien für letale autonome Waffensysteme, denen aber bisher nur 125 Staaten zugestimmt haben (Auswärtiges Amt 2019). Das Problem der unscharfen Definition von autonomen Waffensystemen besteht weiterhin.

Ein möglicher Einsatz von KI im Militär beispielsweise wirft viele ethische, strategische sowie menschenrechtliche Fragen auf, die bisher weitestgehend ungeklärt bleiben. Da Entscheidungen durch KI bei Waffensystemen oft endgültig sind, ist effektive Regulierung hier besonders wichtig. Und da die Entwicklung rasant fortschreitet, kann die Gesellschaft sich nicht auf einzelne Institutionen bzw. Kontrollorgane verlassen. Ganz im Sinne von Hawkings Eingangszitat muss sie vielmehr gemeinschaftlich die Handlungsspielräume und ethischen Richtlinien vorgeben, um eine lebenswerte Zukunft für alle zu gewährleisten. Ein Beispiel für den militärischen Einsatz von KI ist (bzw. war) das Google-Projekt „Maven“ zur Lagebeurteilung. Da die Anwendungsfelder der Projektergebnisse nicht absehbar waren haben sich über 3000 Google-Mitarbeiter in einem öffentlichen Brief gegen das Projekt ausgesprochen (New York Times 2018). Das Unternehmen reagierte: Das Projekt wurde infolgedessen nicht verlängert (Brühl 2018) und seitdem untersagen es Googles „KI Prinzipien“ (Pichai 2020) explizit, „Waffentechnik oder Technologien, die dazu angetan sind, Menschenrechte oder internationale Rechte zu brechen“ zu unterstützen. Doch auch andere Bereiche wie zum Beispiel militärische Logistik, Cyber-Operationen oder das Treffen von Vorhersagen müssen von demokratischen Entscheider*innen sowie der Gesellschaft als Ganzes diskutiert und reguliert werden (Franke 2019).

In den oben beschriebenen Ausführungen nimmt die Zukunft meist dystopische Züge an: Bei Bostrom (2014) ist es wichtig, dass sich Maschinen fühlen, als wären sie aus dem Urlaub gekommen; bei Kurzweil (2006) erschaffen sich sogenannte Nanobots – eine von starker KI selbstgeschaffene Roboterart – selbst, und fressen in letzter Konsequenz sogar Menschen auf, wenn dies für sie opportun ist. An den dystopischen Versionen starker KI gibt es auch

deutliche Kritik. So sind laut Hannah Fry (2018) Algorithmen immer so fehleranfällig (und demnach genauso (un)fähig) wie ihre Erschaffer, und es wird nach ihrer Einschätzung in absehbarer Zeit keinen vollkommenen Algorithmus geben. Weiterhin ist mittlerweile bewiesen, dass selbst das moderne Deep Learning mit neuronalen Netzen immer noch nach den Prinzipien einer Turing Maschine funktioniert und deshalb keine Abkehr vom Prinzip der Algorithmik ist (WDR 5 2019b). Nach Ullrich (2019) liegt hier der besondere Unterschied zum Menschen, der Entscheidungen auch heuristisch (also mit begrenztem Wissen) treffen kann. Bei der Algorithmik ist dies nicht der Fall. Nach Ullrich (2019) sind Maschinen somit stets neutral und bilden immer nur die Vorgaben der Entwickler ab.

Unter KI Expert*innen herrscht also Uneinigkeit bezüglich der Form einer starken KI, des möglichen Eintrittsdatums und der Frage ob es sie jemals geben wird – was auch von der Form bzw. der Definition abhängt (Müller 2014, Grace 2018). Egal welcher Philosophie und potenziellen Zukunftsszenarien man bei starker KI Glauben schenken mag: die theoretische Existenz einer solchen in der Zukunft bestärkt alle für schwache KI gültigen Kriterien, mit stärkerem Fokus auf zukünftige Generationen. Besonders da Expert*innen sich nicht über die Auswirkungen auf die Menschheit einig sind (Vgl. Müller 2014; 31% erwarten extrem schlechte oder schlechte Folgen, 52% extrem gute oder gute Folgen), müssen die potenziellen Folgen heute schon mitbeachtet werden. Die signifikanten Auswirkungen von starker KI mögen weit in der Zukunft liegen und im Moment von anderen dringenden Problemen überschattet werden (Singer 2016), aber aus Fairness gegenüber zukünftiger Generationen müssen wir schon jetzt dafür sorgen, dass die Weichen richtig gestellt werden. Dabei sind folgende Felder besonders relevant (vgl. WBGU 2019: 77):

- Keine Abwertung der menschlichen Fähigkeiten
- Bewahren der menschlichen Verantwortung
- Bewahren der menschlichen Kontrolle
- Bewahren der Selbstbestimmung jedes Einzelnen

Um dies zu gewährleisten, sollten signifikante Entwicklungstätigkeiten in der KI-Forschung nie unkontrolliert und geheim stattfinden (vgl. Bostrom 2014, WBGU 2019: 77). Weiterhin sollte, so lange nicht klar ist was KI im Stande ist zu leisten, niemals von einer oberen Leistungsgrenze ausgegangen werden (vgl. Future of Life Institute 2020). Damit kann sichergestellt werden, dass der Tragweite und potenziellen Risiken starker KI bestmöglich begegnet werden kann – auch noch durch zukünftige Generationen.

Forderungen zur starken KI:

- Es ist wichtig, beim Thema KI zu jeder Zeit handlungsfähig zu bleiben und die **Handlungsspielräume** für zukünftige Generationen **nicht zu verringern**.
- Signifikante Entwicklungen im Feld der KI sollten unter **Beobachtung der Gesamtgesellschaft** stattfinden.
- Heute drängende Fragen zu **schwacher KI** dürfen relevante Fragestellungen über eine potenzielle **starke KI** in der Zukunft nicht in den Hintergrund rücken.

- Um dies zu gewährleisten, sollten zu jeder Zeit und auf **globaler Ebene** die bisher aufgeführten Forderungen umgesetzt werden.
- Nach dem Vorsorgeprinzip sollte bei allen Überlegungen und Maßnahmen keine Obergrenze bzgl. der Leistungsfähigkeit von KI angenommen werden, sondern einem **Worst Case Szenario** schon jetzt entgegengewirkt werden.

4. Fazit

Künstliche Intelligenz ist ein dynamisches Feld, das unsere Zukunft nachhaltig verändern wird. Da Lebensumstände und -präferenzen in der Zukunft unter Umständen nicht mehr so sein werden wie heute, muss KI so umgesetzt werden, dass die Gestaltungsmöglichkeiten in Zukunft nicht sinken und die Bedürfnisse zukünftiger Generationen nicht gefährden werden. Bezüglich der prozeduralen Ebene von KI gibt es viele sinnvolle Leitsätze. Hier seien vor allem die KI-Leitsätze von Asilomar (Future of Life Institute 2020) genannt, die wir komplett unterstützen, vor allem da sie explizit auch langfristige Herausforderungen adressieren. Damit die Entwicklung von KI nicht zum Nachteil zukünftiger Generationen geschieht, stellt dieses Papier die ersten Forderungen bezüglich der Rahmenbedingungen speziell im Kontext von Deutschland. Dabei ist klar, dass viele der Herausforderungen auf übernationaler Ebene angegangen werden müssen. Dennoch kann und sollte Deutschland zusammen mit der EU hier Vorreiter werden.

In den Themenfeldern Bildung und Arbeit bedeutet dies vor allem, junge Menschen dazu zu ermächtigen, ihren Gestaltungsspielraum zu erkennen und wahrnehmen zu können, indem sie ihren politischen Einfluss geltend machen. Aus- und Weiterbildungsformen müssen schnell an die Herausforderungen, die KI bereithält, angepasst werden. Viele Fragen bezüglich Eigentums sind im Bereich des Digitalen, inklusive KI, noch nicht geklärt. Hier müssen Regeln gefunden werden, wie geistiges Eigentum an Algorithmen und deren Weiterentwicklung fair gehandhabt werden. Weiterhin wird die zunehmende Entkopplung von physischer und digitaler Wertschöpfung Herausforderungen aufwerfen, die großen sozialen Folgen hiervon müssen für den/die Einzelne/n abgefedert werden. Algorithmen müssen transparent werden bzw. bleiben, um struktureller Diskriminierung anhand von demographischen Merkmalen (sowohl ethnisch als auch altersbezogen) vorzubeugen. Zudem darf die Monopolisierung digitaler Daten nicht zu Intransparenzen führen. Ein intergenerationaler Diskurs ist notwendig, um die anstehenden Entscheidungen zu moderieren und zu begleiten. Institutionen hierfür müssen geschaffen oder ausgebaut werden, um letztendlich die Ausarbeitung von KI zukunftstauglich zu gestalten. Dies beinhaltet vor allem die ökologische Nachhaltigkeit und den ethisch korrekten Einsatz. Hierbei soll Deutschland zusammen mit Europa als Vorreiter und Schrittgeber agieren.

Die Möglichkeit einer starken KI mit disruptiven Veränderungen für die Menschheit verstärkt den Appell an einen bewussten Umgang mit KI im Sinne der Interessen

zukünftiger Generationen. Es ist essenziell, Handlungsspielräume bezüglich KI nicht zu verringern und stattdessen sogar zu vergrößern.

Deshalb fordern wir, dass alle Menschen eine faire Chance bekommen KI zu verstehen und zu bewerten, um dann an der Entwicklung und Gestaltung teilhaben zu können (Kapitel Arbeit, 2.1). Maßgebliche ethische Entscheidungen müssen von der gesamten Gesellschaft getroffen werden (Kapitel Diskurs, 2.4), und auf internationaler Ebene soll Europa eine Vorreiter- und Vorbildrolle einnehmen, wobei die Gestaltung letztendlich auf globaler Ebene stattfindet (Kapitel Zukunft, 2.5). Nur eine faire Verteilung und Regelung von digitalem Eigentum (Kapitel Eigentum, 2.2) und die Sicherstellung von Transparenz (Kapitel Transparenz, 2.3) können weiterhin dafür sorgen, dass die Entscheidungsgewalt bei der Gesellschaft bleibt, auch und vor allem im Falle der starken KI (Kapitel Starke KI, 3).

Literatur

Accenture (2016): Artificial Intelligence is the future of growth. <https://www.accenture.com/sk-en/insight-artificial-intelligence-future-growth>, letzter Zugriff: 25.07.2019.

Acemoglu, Daron / Robinson, James (2012): Why Nations Fail: The Origins of Power, Prosperity, and Poverty. Crown Publishers, New York, USA.

Auerbach, David (2014): The Most Terrifying Thought Experiment of All Time. <https://slate.com/technology/2014/07/rokos-basilisk-the-most-terrifying-thought-experiment-of-all-time.html>, letzter Zugriff: 22.12.2019.

Auswärtiges Amt (2019): Tödliche Roboter-Waffen regulieren. Artikel vom 15.11.2019. <https://www.auswaertiges-amt.de/de/aussenpolitik/themen/abruestung-ruestungskontrolle/autonome-waffen/2276700>, letzter Zugriff: 08.01.2020.

Baden-Württemberg Stiftung (2018): Aufsichtsrat beschließt 35 Mio. Euro für Programme in 2019. <https://www.bwstiftung.de/news-presse/pressemitteilungen/pressemitteilungen/aufsichtsrat-beschliesst-35-mio-euro-fuer-programme-in-2019/>, letzter Zugriff: 01.03.2019.

Bartlett, Jamie (2018): The People Vs Tech: How the internet is killing democracy (and how we save it), Ebury Press.

Barton, Genie et al (2019): Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>, letzter Zugriff: 02.07.2020.

Bitkom (2017): Entscheidungsunterstützung mit Künstlicher Intelligenz - Wirtschaftliche Bedeutung, gesellschaftliche Herausforderungen, menschliche Verantwortung, Bitkom e.V., DFKI (Hrsg.).

Bostrom, Nick (2014): Superintelligence - Paths, Dangers, Strategies. Oxford University Press, ISBN: 978-0198739838.

Bradford, Laurence (2018): How Open-Source Development Is Democratizing The Tech Industry, Forbes.com. <https://www.forbes.com/sites/laurencebradford/2018/03/26/how-open-source-development-is-democratizing-the-tech-industry/>, letzter Zugriff: 05.08.2019.

Brooks, D. (2013): The Philosophy of Data. <https://www.nytimes.com/2013/02/05/opinion/brooks-the-philosophy-of-data.html>. New York: The New York Times.

Brühl, Jannis (2018): Krieg steckt in der DNA des Silicon Valley. In: Süddeutsche Zeitung. <https://www.sueddeutsche.de/digital/us-militaer-darf-google-beim-toeten-helfen-1.3998295>, letzter Zugriff: 14.01.2020.

Bundesministerium des Inneren, für Bau und Heimat (2019): Datenethikkommission. <https://www.bmi.bund.de/DE/themen/it-und-digitalpolitik/datenethikkommission/datenethikkommission-node.html>, letzter Zugriff: 14.01.2019.

Bundesministerium für Bildung und Forschung (BMBF) (2018): Bekanntmachung - Richtlinie zur Förderung von Forschungsvorhaben zum Thema „Künstliche Intelligenz in

Kommunikationsnetzen" im Rahmen der Hightech-Strategie der Bundesregierung. Bundesanzeiger vom 02.10.2018.

Bundesministerium für Bildung und Forschung (BMBF) (2019): Wissenschaftsjahr 2019 - Künstliche Intelligenz. <https://www.wissenschaftsjahr.de/2019/>, letzter Zugriff: 11.08.2019.

Bundesministerium für Wirtschaft (BMWi) (2020): Einsatz von Künstlicher Intelligenz in der Deutschen Wirtschaft. <https://www.bmwi.de/Redaktion/DE/Publikationen/Wirtschaft/einsatz-von-ki-deutsche-wirtschaft.html>, letzter Zugriff: 20.06.2020.

Bundesregierung Deutschland (2018): Strategie Künstliche Intelligenz der Bundesregierung. https://www.bmwi.de/Redaktion/DE/Publikationen/Technologie/strategie-kuenstliche-intelligenz-der-bundesregierung.pdf?__blob=publicationFile&v=8, letzter Zugriff: 09.01.2019.

Bundesverband Digitale Wirtschaft (BVDW) e.V. (2019): Mensch, Moral, Maschine. Digitale Ethik, Algorithmen und künstliche Intelligenz. https://www.bvdw.org/fileadmin/bvdw/upload/dokumente/BVDW_Digitale_Ethik.pdf, letzter Zugriff: 23.12.2019.

Burgard, Wolfram (2018): Künstliche Intelligenz. Expertise für das WBGU-Hauptgutachten „Unsere gemeinsame digitale Zukunft“. https://www.wbgu.de/fileadmin/user_upload/wbgu/publikationen/hauptgutachten/hg2019/pdf/Expertise_Burgard.pdf, letzter Zugriff: 21.07.2020.

Caney, Simon (2018): Justice and Future Generations. In: Annual Review of Political Science, 2018, 21:475-93.

Cohn, Gabe (2018): AI Art at Christie's Sells for \$432,500. <https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html>, letzter Zugriff: 23.12.2019.

Criado-Perez, Caroline (2019): Unsichtbare Frauen: Wie eine von Daten beherrschte Welt die Hälfte der Bevölkerung ignoriert. Random House, 2019.

Datenethikkommission (2018): Empfehlungen der Datenethikkommission für die Strategie Künstliche Intelligenz der Bundesregierung. Berlin: Datenethikkommission.

Denkschmiede Gesundheit (2019): Eine Gesundheitspolitik - ganzheitlich generationengerecht zukunftsfähig. Positionspapier der Denkschmiede Gesundheit und Stiftung für die Rechte zukünftiger Generationen. https://generationengerechtigkeit.info/wp-content/uploads/2019/08/Positionspapier_Gesundheit_DeGe_SRzG.pdf, letzter Zugriff: 23.12.2019.

Djeffal, Christian (2018): Künstliche Intelligenz. Beitrag zum Handbuch Digitalisierung in Staat und Verwaltung. HIIG Discussion Paper Series No. 3. Berlin: Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG).

Döbel, Inga et al. (2018): Maschinelles Lernen – Eine Analyse zu Kompetenzen, Forschung und Anwendung. Fraunhofer-Gesellschaft e.V., München.

Domingos, Pedros (2016): The Quest for the Master Algorithm, Talk at TEDxUofW in May 2016. <https://www.youtube.com/watch?v=qlZ5PXLVZfo>, letzter Zugriff, 01.12.2019.

Dornis, Valentin (2019): Facebook finanziert Ethik-Institut an der TU München. <https://www.sueddeutsche.de/muenchen/facebook-tu-muenchen-kuenstliche-intelligenz-1.4295434>, letzter Zugriff, 23.12.2019.

Dornis, Valentin (2018): Arbeit aus dem Automaten. <https://www.sueddeutsche.de/digital/digitalisierung-arbeitslosigkeit-jobcenter-1.4178635>, letzter Zugriff: 05.03.2019.

Europäische Kommission (2020): Weißbuch Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf, letzter Zugriff 20.06.2020.

Europäische Kommission (2018): Künstliche Intelligenz: Kommission treibt Arbeit an Ethikleitlinien weiter voran, Pressemitteilung vom 8.4.2018. http://europa.eu/rapid/press-release_IP-19-1893_de.htm, letzter Zugriff: 25.07.2019.

Europäische Kommission (2016a): Rettung von Menschenleben: Mehr Fahrzeugsicherheit in der EU, Bericht der Kommission an das Europäische Parlament und den Rat vom 12.12.2016. <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52016DC0787&from=EN>, letzter Zugriff: 14.01.2020.

Europäische Kommission (2016b): Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG. Datenschutz-Grundverordnung.

Europäisches Parlament (2019): Selbstfahrende Autos in der EU: Science Fiction wird Realität, Pressemitteilung vom 14.01.2019. <https://www.europarl.europa.eu/news/de/headlines/economy/20190110STO23102/selbstfahrende-autos-in-der-eu-science-fiction-wird-realitaet>, letzter Zugriff: 14.01.2020.

Elements of AI (2018): Online-Kurs entwickelt durch die Universität Helsinki und Reaktor. <https://www.elementsofai.com/>, letzter Zugriff 14.01.2020.

Feast, Josh (2019): 4 Ways to Address Gender Bias in AI. Harvard Business Review. <https://hbr.org/2019/11/4-ways-to-address-gender-bias-in-ai>, letzter Zugriff 20.01.2020.

Fischermann, Thomas / Hamann, Götz (2013): Den TÜV fürs Netz, bitte. In: DIE ZEIT Nr. 9/2013. <https://www.zeit.de/2013/09/Internet-Tuev-Viktor-Mayer-Schoenberger-Big-Data/komplettansicht>, letzter Zugriff: 05.08.2019.

Franke, Ulrike Esther (2019): Not smart enough: The poverty of European military thinking on artificial intelligence. https://www.ecfr.eu/publications/summary/not_smart_enough_poverty_european_military_thinking_artificial_intelligence, letzter Zugriff: 14.01.2020.

Fry, Hanna (2018): Hello World - Being Human in the Age of Algorithms. W. W. Norton & Company, ISBN 978-0393357363.

Future of Life Institute (2020): Die KI-Leitsätze von Asilomar. <https://futureoflife.org/ai-principles-german/>, letzter Zugriff: 06.06.2020.

Gabriel, Markus (2018): Wie denken Sie über das Denken? Interviewt von Jürgen Wiebeke für WDR 5 Das philosophische Radio am 28.12.2018. <https://www1.wdr.de/mediathek/audio/wdr5/wdr5-das-philosophische-radio/audio-wie-denken-sie-ueber-das-denken-100.html>, letzter Zugriff: 09.01.2019.

Gartner (2019a): Artificial Intelligence (AI), Gartner IT Glossary, <https://www.gartner.com/it-glossary/artificial-intelligence/>, letzter Zugriff: 23.10.2019.

- Gartner (2019b): A Roadmap to Discover Digital Talent. <https://www.gartner.com/smarterwithgartner/a-roadmap-to-discover-digital-talent-infographic/>, letzter Zugriff: 25.07.2019.
- Gartner (2019c): Gartner Survey Shows 37 Percent of Organizations Have Implemented AI in Some Form. Pressemitteilung am 21.1.2019. <https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have>, letzter Zugriff: 25.07.2019.
- Gartner (2017): Gartner Says By 2020, Artificial Intelligence Will Create More Jobs Than It Eliminates, Pressemitteilung am 13.12.2017. <https://www.gartner.com/en/newsroom/press-releases/2017-12-13-gartner-says-by-2020-artificial-intelligence-will-create-more-jobs-than-it-eliminates>, letzter Zugriff: 25.07.2019.
- Pichai, Sundar (2020): Berechtigte Bedenken. <https://about.google/intl/de/stories/sieben-prinzipien/>, letzter Zugriff: 02.07.2020.
- Government of Canada (2019): Digital Youth Advisory Committee. <https://www.canada.ca/en/treasury-board-secretariat/topics/information-technology-project-management/digital-youth-advisory-committee.html>, letzter Zugriff: 15.01.2019.
- Grace, Katja et al. (2016): When Will AI Exceed Human Performance? Evidence from AI Experts. <https://arxiv.org/pdf/1705.08807>, letzter Zugriff: 06.06.2020.
- Gründinger, Wolfgang et al. (2017): Die Lage der jungen Generation auf dem Arbeitsmarkt. Positionspapier der Stiftung für die Rechte zukünftiger Generationen. https://generationengerechtigkeit.info/wp-content/uploads/2018/06/PP_Junge-auf-dem-Arbeitsmarkt.pdf, letzter Zugriff: 25.07.2019.
- Harari, Yuval Noah (2015): Homo Deus. Harper Perennial, ISBN 978-0062464347.
- Harari, Yuval Noah (2018): 21 Lessons for the 21st Century. Penguin Random House, New York, USA.
- Hawking, Stephen (2017): Key note. Web Summit. 6.11.2017, Lissabon, Portugal.
- Inter-Parliamentary Union (2018): Youth participation in national parliaments 2018. <https://www.ipu.org/our-impact/youth-empowerment/data-youth-participation>, letzter Zugriff: 14.01.2020.
- Kaltheuner, Frederike / Obermüller, Nele (2018): Diskriminierende Gesichtserkennung: Ich sehe was, was du nicht bist. <https://netzpolitik.org/2018/diskriminierende-gesichtserkennung-ich-sehe-was-was-du-nicht-bist/>, letzter Zugriff: 14.01.2020.
- Kurzweil, Ray (2006): The Singularity Is Near: When Humans Transcend Biology. Penguin Books, ISBN 978-0739466261.
- Lenzen, Manuela (2019): Künstliche Intelligenz. Was sie kann und was uns erwartet. C.H.Beck, ISBN 978-3406718694.
- Loske, Reinhard (2007): Allen wohl und keinem wehe. In: Ökologisches Wirtschaften, 2007, Nr. 1.
- Maaz, Kai et al. (2018): Bildung in Deutschland. Ein indikatorgestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung.

<https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2018/pdf-bildungsbericht-2018/bildungsbericht-2018.pdf>, letzter Zugriff: 30.08.2019.

Mainzer, Klaus (2016): Künstliche Intelligenz – Wann übernehmen die Maschinen? Heidelberg, Berlin: Springer.

Max Planck Institut (2018): Bundeswettbewerb Künstliche Intelligenz gestartet. <https://www.is.mpg.de/news/bundeswettbewerb-kunstliche-intelligenz-gestartet>, letzter Zugriff: 23.10.2018.

Mercator Institute for China Studies (2018): Stephan Scheuer und Andre Loesekrug-Pietri über „Silicon China“. <https://www.merics.org/de/podcast>, letzter Zugriff: 01.12.2018.

Moorstedt, Michael (2019): Ein Leben ohne die furchtbaren Fünf ist möglich, aber sinnlos. <https://www.sueddeutsche.de/digital/amazon-facebook-apple-google-microsoft-1.4304826>, letzter Zugriff: 11.08.2019.

Moravec, H. (1988): Mind Children. The Future of Robot and Human Intelligence. Cambridge, MA: Harvard University Press.

Morley, Janine et al. (2018): Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption. In: Energy Research & Social Science, 2018, 38:128–137.

Müller, Vincent / Bostrom, Nick (2014): Future progress in artificial intelligence: A Survey of Expert Opinion, in Vincent C. Müller (ed.), Fundamental Issues of Artificial Intelligence (Synthese Library; Berlin: Springer).

New York Times (2018): Letter by Google staff to Google CEO Sundar Pichai. <https://static01.nyt.com/files/2018/technology/googleletter.pdf>, letzter Zugriff: 14.01.2020.

Open Source Business Alliance (OSB Alliance) (2019): Leitlinien der OSB Alliance. <https://osb-alliance.de/ueber-uns/leitlinien-der-osb-alliance>, letzter Zugriff: 11.08.2019.

Ovens, Carsten (2017): Filterblasen - Ausgangspunkte einer neuen, fremdverschuldeten Unmündigkeit? In: kommunikation @ gesellschaft, 18, 1-25.

Panetta, Kasey (2018): 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies, 2018, Gartner. <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>, letzter Zugriff: 23.10.2018.

Pichai, Sundar (2020): Berechtigte Bedenken. <https://about.google/intl/de/stories/sieben-prinzipien/>, letzter Zugriff: 02.07.2020.

Renn, Ortwin / Knaus, Anya (1998): Den Gipfel vor Augen: Unterwegs in eine nachhaltige Zukunft. Berlin: Metropolis.

Rotenberg, Marc (2020): The AI Policy Sourcebook 2020. Washington DC: Electronic Privacy Information Center.

SAP Deutschland SE & Co. KG (2020): Corona Warn-App. <https://www.coronawarn.app/de/>, letzter Zugriff: 30.09.2020.

Shoham, Yoav et al. (2018): Artificial Intelligence Index 2018 annual report. <http://cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf>, letzter Zugriff: 25.07.2019.

- Siemens (2019): White paper 2019 –Artificial Intelligence and Machine Learning. <https://new.siemens.com/global/en/company/topic-areas/digitalization/ai-ml-white-paper.html>, letzter Zugriff: 25.07.2019.
- Singer, Peter (2016): Can artificial intelligence be ethical? World Economic Forum. <https://www.weforum.org/agenda/2016/04/can-artificial-intelligence-be-ethical>, letzter Zugriff: 05.06.2020.
- Stiefel, Jörg / Brooks Daniel (2019): Why is There No Successful Whole Brain Simulation (Yet)?, In: Biol Theory 14, 122–130.
- Stiftung für die Rechte zukünftiger Generationen SRzG (2019): Nachwuchsquoten in Parteien und Parlamenten - Warum die Mitbestimmung junger Menschen durch Quoten sinnvoll ist. Positionspapier der Stiftung für die Rechte zukünftiger Generationen. http://generationengerechtigkeit.info/wp-content/uploads/2019/02/PP-Nachwuchsquoten_2019.pdf, letzter Zugriff: 23.12.2019.
- Stockholm Resilience Center (2019): The nine planetary boundaries. <https://www.stockholmresilience.org/research/planetary-boundaries/planetary-boundaries/about-the-research/the-nine-planetary-boundaries.html>, letzter Zugriff: 15.01.2019.
- Tremmel, Jörg (2019): Gerechtigkeit zwischen den Generationen. In: Schweiger, Gottfried / Drerup, Johannes (Hg.): Philosophie der Kindheit. Stuttgart: JB Metzler, 371-379.
- Tremmel, Jörg (2012): Eine Theorie der Generationengerechtigkeit. Münster: mentis.
- Trottenberg, Ulrich (2016): Kinder müssen lernen, Algorithmen zu lieben. In: Süddeutsche Zeitung. <https://www.sueddeutsche.de/digital/digitalisierung-kinder-muessen-lernen-algorithmen-zu-lieben-1.2840333>, letzter Zugriff: 05.03.2019.
- Ullrich, S. (2019): Moralische Maschinen. Was die Maschine über die Moral ihrer Schöpferinnen und Schöpfer verrät. In: Rath, M. / Krotz, F. / Karmasin, M. (Hrsg.): Maschinenethik: Normative Grenzen autonomer Systeme. Wiesbaden: Springer VS, 243–261.
- United Nations (2019): Big Data UN Global Network Group. <https://unstats.un.org/bigdata/>, letzter Zugriff: 15.01.2019.
- United Nations (1987): Brundtland Report: Our Common Future. Report of the World Commission on Environment and Development.
- Welcherling, Peter (2019): KI-Systeme im Militär. Deutschlandfunk. https://www.deutschlandfunk.de/autonome-waffen-ki-systeme-im-militaer.676.de.html?dram:article_id=459749, letzter Zugriff: 01.01.2020.
- WDR 5 (2019a): Das philosophische Radio – Auslaufmodell? – das Automobil. <https://www1.wdr.de/mediathek/audio/wdr5/wdr5-das-philosophische-radio/audio-auslaufmodell---das-automobil-100.html>, letzter Zugriff, 14.01.2020.
- WDR 5 (2019b): Das philosophische Radio - Was wollte Alan Turing? <https://www1.wdr.de/mediathek/audio/wdr5/wdr5-das-philosophische-radio/index.html>, letzter Zugriff, 01.12.2019.
- Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen WBGU (2019): Unsere gemeinsame digitale Zukunft, Berlin: WBGU.

Zeit Online (2017): De Maizière hält russische Einflussversuche für denkbar.
<https://www.zeit.de/politik/deutschland/2017-08/bundestagswahl-thomas-de-maiziere-russland-einflussnahme>, letzter Zugriff: 05.08.2019.

Zierer, Klaus (2018): Lernen 4.0 - Pädagogik vor Technik: Möglichkeiten und Grenzen einer Digitalisierung im Bildungsbereich. Schneider Verlag Hohengehren, ISBN 978-3834017703.

Zimmermann, Hendrik / Frank, David (2019) Künstliche Intelligenz für die Energiewende: Chancen und Risiken. Germanwatch, www.germanwatch.org/de/17095, letzter Zugriff: 01.12.2019.

Zou, James / Schiebinger, Londa (2018): AI can be sexist and racist — it's time to make it fair. In: Nature. <https://www.nature.com/articles/d41586-018-05707-8>, letzter Zugriff: 20.01.2020.

Über die Stiftung für die Rechte zukünftiger Generationen (SRzG)



Stiftung für die Rechte
zukünftiger Generationen

Die Stiftung für die Rechte zukünftiger Generationen (SRzG) ist eine advokatorische Denkfabrik an der Schnittstelle zwischen Wissenschaft und Politik und gilt als „bekanntester außerparlamentarischer Think Tank in Sachen Generationengerechtigkeit“ (Wirtschaftswoche). Sie wurde 1997 von einer überparteilichen Allianz fünf junger Menschen im Alter von 18 bis 27 Jahren ins Leben gerufen, wird von einem der jüngsten Stiftungsvorstände Deutschlands geleitet und verfolgt das Ziel, durch praxisnahe Forschung und Beratung das Wissen und das Bewusstsein für Generationengerechtigkeit und Nachhaltigkeit in Politik, Wirtschaft und Gesellschaft zu schärfen. Die Stiftung ist finanziell unabhängig und steht keiner politischen Partei nahe.

UNTERSTÜTZEN SIE UNS MIT IHRER SPENDE!

per Überweisung:

Stiftung für die Rechte zukünftiger Generationen

GLS Gemeinschaftsbank eG

IBAN: DE64 4306 0967 8039 5558 00

BIC (SWIFT-CODE): GENODEM1GLS

...oder auf generationengerechtigkeit.info/unterstuetzen/

IMPRESSUM

Herausgeberin: Stiftung für die Rechte zukünftiger Generationen
Mannspergerstr. 29, 70619 Stuttgart, Deutschland
Tel: +49 711 28052777
Fax: +49 3212 2805277
E-mail: kontakt@srzg.de
generationengerechtigkeit.info

Redaktion: Stiftung für die Rechte zukünftiger Generationen
Autor*innen: Thomas Betten, Christina Wunder (die Autor*innen haben gleichermaßen zu dieser Arbeit beigetragen)
Mitarbeit: Michelle Grebe, Jonathan Hoffmann, Dr. Dr. Jörg Tremmel, Carl-Georg Luft, Charlotte Unruh
Verantwortlich: Der Vorstand

Design: Stiftung für die Rechte zukünftiger Generationen
Bildnachweis: Titelseite: Pexels/ ThisIsEngineering

© Stiftung für die Rechte zukünftiger Generationen

Stand: 1. Auflage, September 2020.