

Albrecht, Thorben; Kellermann, Christian

**Working Paper**

## Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft: Konturen einer ganzheitlichen Technikfolgenabschätzung

Working Paper Forschungsförderung, No. 200

**Provided in Cooperation with:**

The Hans Böckler Foundation

*Suggested Citation:* Albrecht, Thorben; Kellermann, Christian (2020) : Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft: Konturen einer ganzheitlichen Technikfolgenabschätzung, Working Paper Forschungsförderung, No. 200, Hans-Böckler-Stiftung, Düsseldorf

This Version is available at:

<https://hdl.handle.net/10419/228963>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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



<https://creativecommons.org/licenses/by/4.0/de/legalcode>

# WORKING PAPER FORSCHUNGSFÖRDERUNG

---

Nummer 200, November 2020

## Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft

Konturen einer ganzheitlichen  
Technikfolgenabschätzung

Thorben Albrecht und Christian Kellermann

---

© 2020 by Hans-Böckler-Stiftung  
Georg-Glock-Straße 18, 40474 Düsseldorf  
[www.boeckler.de](http://www.boeckler.de)



„Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft“  
von Thorben Albrecht und Christian Kellermann ist lizenziert unter

**Creative Commons Attribution 4.0 (BY).**

Diese Lizenz erlaubt unter Voraussetzung der Namensnennung des Urhebers die Bearbeitung, Vervielfältigung und Verbreitung des Materials in jedem Format oder Medium für beliebige Zwecke, auch kommerziell.  
(Lizenztext: <https://creativecommons.org/licenses/by/4.0/de/legalcode>)

Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z. B. von Schaubildern, Abbildungen, Fotos und Textauszügen erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

**ISSN 2509-2359**

# Inhalt

Zusammenfassung.....	4
1. Einleitung .....	6
2. KI und die technologischen Grenzen ihrer Anwendbarkeit.....	8
3. Betrieblicher Einsatz im Spannungsverhältnis aus Rentabilität und Regulierung.....	11
4. Verschärfte Polarisierung am Arbeitsmarkt.....	14
5. Konturen einer ganzheitlichen Technikfolgenabschätzung .....	18
Literatur.....	24
Autoren .....	27

## Zusammenfassung

Die Digitalisierung verändert die Arbeitsgesellschaft seit mehreren Jahrzehnten. Mit neuen Technologien wie maschinellen Lernmethoden der Künstlichen Intelligenz kann die Digitalisierung in weitere Bereiche der Arbeitswelt vordringen – und tut dies bereits. Das grundlegende disruptive Potenzial, das vielfach in die KI hineinprojiziert wird, steht heute allerdings noch nicht in einem breit beobachtbaren Verhältnis konkreter Anwendungen in der betrieblichen Praxis.

Zwischen den hohen Erwartungen an KI und der Organisation von Arbeit besteht eine offene Leerstelle, die zu füllen der Anspruch des digitalen Werkzeugs KI und der dahinerstehenden Entwicklungsindustrie ist. Die Veränderungen, die auf unsere Arbeitsgesellschaft zukommen, stehen in einem Zusammenhang mit der technischen Entwicklung. Wie eng dieser Zusammenhang ist, bestimmt aber nicht (nur) die Geschwindigkeit der Technikentwicklung, sondern eine Reihe von Faktoren, die Einfluss auf die Gestaltung der Zukunft der digitalen Arbeitsgesellschaft haben.

Die Gestaltung dieser Zukunft setzt eine umfassende Technikfolgenabschätzung voraus. Sie muss die qualitativen Veränderungen von Arbeit und eine belastbare quantitative Schätzung der Auswirkungen auf die Gesamtbeschäftigung zusammenführen. Isolierte Betrachtungen von Technikfolgen eines bestimmten Phänomens der Digitalisierung im Allgemeinen und der KI im Besonderen bleiben zweifelsohne weiterhin notwendig, sie reichen aber nicht aus, um die digitale Arbeitsgesellschaft umfassend zu konturieren.

Für eine breit angelegte, arbeitsfokussierte Technikfolgenabschätzung ist es notwendig, die verschiedenen Einsatzformen von KI im Arbeitskontext zu kategorisieren und kontinuierlich zu erfassen. Dabei kommt es nicht primär auf die technischen Unterschiede zwischen verschiedenen KI-Systemen oder deren technischen Reifegrad an, sondern auf die Rolle, die die KI im Unternehmen spielt oder spielen soll. Je nach Rolle ergeben sich unterschiedliche Anforderungen an die Regulierung und Transparenz von KI sowie an die Rahmenbedingungen für die Beteiligung und Befähigung von Mitarbeitern und ihren Vertretungen in einem Betrieb.

Eine solche Technikfolgenabschätzung braucht geeignete Kategorisierungen und *Tools*, die innerbetrieblich sowie überbetrieblich eingesetzt werden können. Wir entwickeln in dem Papier vier Einsatzmodelle der KI auf betrieblicher Ebene. Die Technologie kann ein Werkzeug sein, das die Arbeitnehmer unterstützt oder teilweise ersetzt. Und sie

kann (zusätzlich) Entscheidungen in der Personalsteuerung unterstützen oder selbständig übernehmen.

Die betrieblichen Gestaltungsebenen des Technikeinsatzes fügen sich in eine makroökonomische Diskussion der wahrscheinlichen Szenarien für die Entwicklung der Gesamtbeschäftigung im Zuge der fortschreitenden Digitalisierung. Wir entwickeln auf diese Weise einerseits das (gesellschaftliche) Ziel, die Kontrolle über die Technik nicht zu verlieren, bzw. der KI und der Digitalisierung den notwendigen Entwicklungs- und Anwendungsspielraum zu geben. Andererseits beschreiben wir einen Weg, die Auswirkungen der Technologie am einzelnen Arbeitsplatz, auf betroffene Berufsgruppen und auf die Gesamtwirtschaft kalkulieren zu können und verlässliche Überprüfungsmechanismen auf verschiedenen Ebenen zu etablieren.

Ohne die Verschränkung der Analyse- und Handlungsebenen – so unsere Schlussfolgerung – droht die Gefahr einer verschärften Polarisierung in der digitalen Arbeitsgesellschaft, der Einkommen und damit auch eine verstärkte Polarisierung der Gesellschaft als Ganze.

# 1. Einleitung

Die Digitalisierung beschleunigt den Wandel der Arbeit. Auch wenn Arbeit immer Veränderungen unterworfen ist, gibt es bestimmte Phasen, in denen die Veränderung schneller verläuft. Meist spielt die Einführung neuer Technologien dabei eine wichtige Rolle, aber immer im Kontext von und in Wechselwirkung mit anderen gesellschaftlichen und kulturellen Veränderungen und Gegebenheiten. Auch die Digitalisierung muss im Kontext von neuen Formen der Globalisierung, demografischen Entwicklungen, Veränderungen in der Bildung und nicht zuletzt einem Wertewandel bei den arbeitenden Menschen gesehen werden.

Unter dem *Label* Künstliche Intelligenz (KI) findet mit der Einführung von bestimmten digitalen Techniken im Arbeitskontext und in der Gesellschaft ein qualitativ neuer technologischer Schritt statt. Definitionsgemäß ist KI ein Sammelbegriff für Algorithmen, die große Datenmengen verarbeiten können, bedingt lernfähig sowie in der Lage sind, selbstständig komplexe Lösungen zu finden. KI verfügt über die Fähigkeit zur Selbstoptimierung und zeichnet sich durch eine immanente Komplexität und intransparente Lösungswege aus („*Black Box*“). Das unterscheidet sie von bisher genutzten, auch digitalen Werkzeugen. Auf der betrieblichen Ebene können KI-Systeme Arbeit organisieren, steuern und kontrollieren. Sie können auch eine sich selbst organisierende und optimierende Struktur von Geschäfts- und Arbeitsbeziehungen schaffen, wie dies beispielsweise auf Arbeitsplattformen erfolgt. Damit führt der Einsatz von KI-Systemen im Arbeitskontext zu qualitativen und quantitativen Veränderungen, die eine Anpassung des regulativen Rahmens und in den Arbeitsbeziehungen erfordern.

Insbesondere auf der gesamtwirtschaftlichen, überbetrieblichen Ebene stellt sich neben der Frage nach der neuen Qualität von Arbeit die Frage nach quantitativen Beschäftigungseffekten, sollten KI und andere digitale Technologien immer stärker Einzug in die Betriebe halten – wovon auszugehen ist. Die Debatte über mögliche Auswirkungen auf die Gesamtbeschäftigung ist jedoch nicht nur in Deutschland, sondern in vielen Ländern sehr uneinheitlich. Die Prognosen über Beschäftigungswirkungen variieren stark, was mit der jeweiligen Perspektive auf den Einsatz von KI und deren Potenzial zusammenhängt.

Wir wollen in diesem Paper die zentralen Dimensionen einer arbeitsorientierten Bewertung von fortschrittlichen digitalen Technologien im Allgemeinen und KI im Besonderen in der digitalen Arbeitsgesellschaft bestimmen. Das Ziel dabei ist, die Notwendigkeit und die Handlungsfelder einer ganzheitlichen Technikfolgenabschätzung zu definieren. Die erste

Dimension ist die Technik selbst, bzw. ihr technisches-wirtschaftliches Potenzial. KI ist Teil der Digitalisierung und beides ist nicht immer sauber zu trennen. Wir betrachten deswegen zunächst die Entwicklungen der KI und die technologischen Grenzen ihrer Anwendbarkeit mit dem Ziel, sie greifbarer für weitergehende Analysen im Arbeitskontext zu machen. Die zweite Dimension ist der Arbeitsplatz selbst. Wie wird KI konkret eingesetzt und welche arbeitspolitischen und -organisatorischen Veränderungen sind damit verbunden? Die dritte Dimension ist die gesamtwirtschaftliche. Kommt es durch das Automatisierungspotenzial zu technologischer Arbeitslosigkeit oder bringt die KI Wachstum und Beschäftigung für alle auf einem höheren Niveau – oder verschärft sich die Polarisierung, die bereits in den ersten Wellen der Automatisierung und Digitalisierung begann? Diese drei Dimensionen sind aus unserer Sicht elementar für die Konturierung der Technikfolgenabschätzung im Zuge der weitergehenden Digitalisierung und leiten die politischen Handlungsfelder an, die wir im letzten Abschnitt ausführen.

## 2. KI und die technologischen Grenzen ihrer Anwendbarkeit

Methoden Künstlicher Intelligenz sind aus der Verbindung von hochleistungsfähiger Hardware, großen Datensätzen (*Big Data*) und maschinellen Lernverfahren (basierend auf Algorithmen) möglich geworden. Während der Begriff der Künstlichen Intelligenz bereits in den 1950ern geprägt wurde, haben sich die heute maßgeblichen Verfahren erst in den letzten Jahrzehnten entwickelt. Dies hängt sowohl mit den Entwicklungen im Hardware-Bereich zusammen (immer kleinere und schnellere Chips und damit mehr Rechenleistung), als auch im Software-Bereich (immer bessere Algorithmen). In der Anwendung geht KI über die bisherigen Formen der maschinellen Informationsverarbeitung hinaus, indem immer komplexere Umgebungen und Prozesse per Computing analysier- und steuerbar werden. Die Entwicklungen in den Bereichen Speichertechnik, Rechenkapazität und -geschwindigkeit markieren somit einen *qualitativen* Schritt von der vormals überwiegend theoretischen hin zur anwendungsorientierten KI.

Ungeachtet der vielen (populär-kulturellen) Debattenbeiträge zur sog. „*General Purpose KI*“, stößt KI bislang immer noch dann an ihre Grenzen, wenn es um autonomes Verstehen und Begreifen geht, also um „Intelligenz“ über die Mustererkennung hinaus. Selbst trivialste Zusammenhänge können ein KI-System schnell überfordern. Vor allem die enormen Kombinationsmöglichkeiten von Daten führen auch bei überschaubaren Datensätzen schnell zu unbeherrschbaren Größenordnungen. Diesem Größenphänomen bei Daten ist vorerst nur mit Hilfe von Heuristik und der Berechnung von Wahrscheinlichkeiten beizukommen. Das bedeutet aber trotzdem, dass Ergebnisse aus solchen Prozessen gut genug sein können, um die Leistung eines Menschen zu übertreffen; sie sind jedoch nicht perfekt und stets mit einem Rest an Unsicherheit behaftet. Das widerspricht nicht intelligentem Handeln *im Prinzip*, weil auch menschliches Handeln in der Regel vor dem Hintergrund unsicherer, also nicht-vollständiger Informationslage geschieht. Aber damit verbunden ist eine begrenzte Einsetzbarkeit einzelner KI-Systeme *in der Praxis*.

Unsicherheit bestimmt auch jede Prognose, wie sich KI weiterentwickeln wird, was einige nicht daran hindert, KI sehr viel Potenzial in der Arbeitswelt „blanko“ zu bescheinigen. Dem heutigen Stand der Dinge nach sind genuine KI-Anwendungen in den allermeisten Betrieben entweder sehr eingegrenzt oder noch Zukunftsprojektionen bzw. finden erst in seltenen, z. T. experimentellen Fällen Anwendung. Uns interessieren deshalb zunächst die Entwicklung konkreter Anwendungsfelder der KI in

den letzten Jahren, um das aktuell praktische Technikpotenzial eingrenzen zu können.

Ein bereits relativ reifes, wesentliches Feld für den Einsatz von KI in verschiedensten Produktionsbereichen ist das maschinelle Lernen. Mithilfe des maschinellen Lernens werden zwei Querschnittsaufgaben in Unternehmen mittels KI beherrschbarer: (1) das Erkennen von Aktionen auf höherer Ebene und somit abstrakteres Erkennen von vergleichbaren Situationen (Mustererkennung) und (2) der Umgang mit extrem großen Datenmengen. Dies bedeutet, dass maschinelles Lernen dazu beiträgt, anders als beispielsweise Auswendiglernen (gleich Speichern), Schlussfolgerungen und Generalisierungen aus Beispielen und Beobachtungen zu ziehen und Lösungsvorschläge zwischen verschiedenen Situationen übertragen zu können. Typische praktische Anwendungsbereiche sind zielgenaue Werbung und Marketing, die Logistik, *Predictive Maintenance* ebenso wie das *Customer Relationship Management* und *People Analytics*. Beispielsweise aus Nutzerdaten wie dem Kauf- oder Suchverhalten auf Plattformen werden ähnliche Angebote automatisch erstellt. Die Vertriebsphase der Wertschöpfungskette (Werbung, Internet, Customer Relations) und das maschinelle Lernen stehen somit (nicht nur) aus diesem Grund in einem interdependenten Verhältnis.

In vielen anderen Kontexten markieren aber Methoden der KI derzeit die technologische Grenze der Digitalisierung der Betriebs- und Arbeitswelt. Zwischen der Potenzialzuschreibung der Technologie und dem konkreten KI-Einsatz am Arbeitsplatz klafft momentan noch eine große Lücke. Trotzdem ist es absehbar, dass KI und mit ihr auch weniger „intelligente“ Entwicklungen der Digitalisierung die Arbeitswelt in absehbarer Zeit zunehmend und grundlegend verändern werden (bzw. dies teilweise bereits tun). Das Entwicklungs- und Anwendungspotenzial zunehmender Berechenbarkeit von Prozessen in Arbeitszusammenhängen von Produktion, Dienstleistungen, Verwaltung oder Landwirtschaft ist offensichtlich. Der Präzisionsgrad leidet jedoch vor allem darunter, dass sowohl KI als solche schwer erfassbar und somit messbar ist, als auch dass viele Betriebs- und Arbeitsprozesse, in denen KI potenziell integriert werden kann, ähnlich wenig erfass- und messbar – und somit formalisierbar – sind.

Trotz der regen Debatte über die Potenziale von KI, unsere Gesellschaft und Arbeitswelt grundlegend zu verändern, sind die konkreten Auswirkungen von KI auf die Arbeitswelt noch untererforscht, und der Diskurs darüber verläuft häufig anekdotenhaft. Getrieben werden so einige Anekdoten vom Ziel einiger KI-Entwickler, die menschliche Intelligenz vollständig zu mechanisieren – was für den einen ein ambitioniertes, erstrebenswertes Ziel, für die andere aber ein ambivalentes, dysto-

pisches erscheint. Der KI- und Robotik-Pionier Nils Nilsson drückte dieses Ziel so aus: „Die vollständige Automatisierung ökonomisch wichtiger Arbeit“ (Nilsson 2005: 69). Mittels eines „*Employment Test*“, schlug er vor, könne bemessen werden, welchen Anteil von menschlicher Arbeit ein KI-System akzeptabel ausführen könne. Man müsse somit KI-Systeme nur die gleichen Qualifikationstests durchlaufen, wie sie von Menschen für eine spezielle Tätigkeit verlangt werden. Aktuelle Versionen solcher Tests kommen zum Ergebnis, dass KI den Wissensberufen immer mehr Konkurrenz macht und entsprechend immer vertikaler in Betrieben ausgreift (Webb 2019; Muro et al. 2019).

Diese und andere Tests leiden aber unter einer ganzen Reihe methodischer Beschränkungen, durch die ihre Aussagekraft stark reduziert werden. Vor allem werden häufig die Tätigkeiten selbst nur schlagwortartig erfasst und dann in der Kontrastierung mit den angeblichen Fertigkeiten der KI vorschnell für redundant erklärt. Dazu kommt, dass nicht alles betrieblich umgesetzt wird, weil es technisch möglich ist. Die Einführung technischer Innovationen in der Arbeitspraxis, auch im Bereich der KI-Technologie, unterliegt besonderen Bedenken und realen Beschränkungen. Im nächsten Abschnitt systematisieren wir deshalb den möglichen Einsatz (und deren Grenzen) von KI am konkreten Arbeitsplatz. Das ist ein wesentlicher Schritt zum differenzierten Verständnis vom Einsatzzweck der KI in der Arbeitswelt insgesamt, der nur in ausgewählten Fällen überhaupt auf die Substitution von Arbeit abzielt (Kapitel 4).

### **3. Betrieblicher Einsatz im Spannungsverhältnis aus Rentabilität und Regulierung**

Ob und in welchem Umfang KI im einzelnen Unternehmen tatsächlich eingesetzt wird, hängt von mehreren Faktoren ab. Was kann KI technisch leisten, ist der erste. Was darf KI, also welchen regulatorischen Rahmenbedingungen und damit die Frage, unter welchen Bedingungen und für welche Aufgaben bzw. mit welchen Auflagen KI eingesetzt werden darf, ist der zweite. Zu dieser zweiten Frage zählen nicht nur ethische und regulatorische Beschränkungen, sondern auch den Aspekt, inwiefern sich KI überhaupt in den Produktions- und Arbeitsprozessen integrieren lässt, also ihre „Integrationsfähigkeit“. Diese beiden Faktoren bilden den Rahmen für die betriebswirtschaftliche Entscheidung, ob sich eine Investition in KI im Einzelfall tatsächlich lohnt. Dieser dritte Faktor, bzw. die Frage wie die betriebswirtschaftliche Berechnung von Kosten und Ertrag im Einzelfall aussieht, ist im Falle von KI besonders schwer anzustellen. Die Entscheidung über den tatsächlichen Einsatz von KI im Unternehmen hängt also davon ab, was KI kann, was KI darf, und was sie dem Unternehmen konkret bringt.

Auch ungeachtet der Rahmensetzung ist in vielen betrieblichen Kontexten schwer abzuschätzen, was Investitionen in KI tatsächlich für den Produktions- und Arbeitsprozess bedeuten und ob sie die erstrebten Produktivitätsfortschritte erzielen. Vor dem Hintergrund der Tatsache, dass sich die Technologie selbst sowie ihren Einsatz verändert, dürfte diese Entscheidungen für viele Unternehmen schwerfallen. Die Geschwindigkeit sowie das Ausmaß der Weiterentwicklung der KI in den kommenden Jahren, ist offen. Auch bei einzelnen Anwendungen lässt sich teilweise nur schwer einschätzen, wie schnell sich die betriebliche Anpassung an den KI-Einsatz einstellen wird und wie lange sie trägt. Darüber hinaus ist der regulative Rahmen eine schnell veränderliche Größe, da sie in vielen Bereichen erst am Anfang steht. Schließlich ist schwer einzuschätzen, ob die Technologie erfolgreich in die Arbeitsorganisation eines Unternehmens integriert werden kann und damit die Gesamtabläufe der Produktion oder Dienstleistungserbringung tatsächlich verbessert werden. Denn dies hängt stark von „weichen“ Faktoren ab, nämlich davon, ob die Interaktion der KI mit dem Menschen (Mitarbeiter, Kunden etc.) gelingt und produktiv genug ist. Misslingt dies, sind auch negative Effekte nicht auszuschließen.

Es ist dieses Spannungsverhältnis aus technologisch beweglichen Grenzen, unklarem regulativen Rahmen und betrieblicher Funktionalität,

das eine betriebswirtschaftliche Investitionsentscheidung mit vielen Unsicherheiten prägt. Wird eine positive Entscheidung trotzdem getroffen und eine KI-Investition getätigt, dann vielleicht eher getrieben von einem allgemeinen technologischen Optimismus, einem Wunsch, sich als „*Front Runner*“ darzustellen oder schlicht durch die Ratschläge der die technische Zukunft häufig überzeichnenden Beratungsinstitute. Somit sind auch teure Investitionsruinen nicht ausgeschlossen (wie in den 1970er/1980er bereits erlebt im Fall der sog. „CIM-Ruinen“: *Computer Integrated Manufacturing*). Im Umkehrschluss kann es eben passieren, dass produktive Potenziale von KI aufgrund von Unsicherheit und genereller Skepsis gar nicht genutzt werden. In beiden Fällen könnten eine fundiertere Debatte bzw. Folgenabschätzung der KI-Technologie helfen, die Unsicherheiten der betrieblichen Entscheidungen zu reduzieren.

Was sind zu erwartende Auswirkungen von KI-Anwendungen in der Arbeitswelt oder direkt am Arbeitsplatz? In welchen Bereichen spielt KI oder könnte KI eine Rolle spielen? Diese Fragen sind nicht zuletzt relevant, wenn man beachtet, dass der Einsatz von KI sowohl die betriebliche Funktionalität, die „Funktionalität der Arbeit“, sowie die qualitativen Bedingungen und die Machtverhältnisse am Arbeitsmarkt und im Unternehmen verschieben (kann).

Aus der jetzigen (qualitativen) Potenzialanalyse ist zu erwarten, dass ein KI-Einsatz vor allem Auswirkungen auf kognitive Routinetätigkeiten haben wird, wobei „Routine“ bei fortschreitenden KI-Fähigkeiten weit gefasst werden muss (mehr dazu im Kapitel 4). Sie umfasst z. B. schon heute die Bearbeitung von Standardfällen, beispielsweise im Finanz- und Versicherungsbereich, im medizinischen- und Gesundheitssystem sowie in vielen juristischen Bereichen. Hier spielt KI aber auch nicht mehr nur eine Rolle für Datenauswertungen, Diagnosen und Prognosen, sondern auch in der Forschung. Gerade in der Diskussion um Mustererkennung im Bereich der Medizin hat es sich aber gezeigt, dass ein vollständiger Ersatz menschlicher Entscheidungen (noch) nicht möglich ist, sondern viel dafür spricht, dass KI – zumindest in sensiblen und Grenzbereichen – eher als Assistenzsystem eingesetzt werden wird.

Ein anderes Feld des KI-Einsatzes, das auch die Grenzen deren Einsatz zeigt, sind betreuende und umsorgende Dienstleistungen am Menschen. Roboter, häufig verstanden als physische Agenten der KI, sind nicht selten der Gradmesser für die Automatisierung in einem Betrieb oder einer Dienstleistung in Folge der Digitalisierung und den damit verbundenen Effekten für Arbeit und Beschäftigung (z. B. Dauth et al. 2017; Bessen 2018). KI kann zwar ein Robotersystem steuern, das theoretisch alle möglichen Assistenzfunktionen in einem Unternehmen oder einem Haushalt ausführen kann. Bilderkennung, Sensoren, Aktoren (also an-

triebstechnische Baueinheiten) sind somit zunehmend in der Lage, diverse menschnahe Aufgaben zu erfüllen.

Trotzdem wird vor allem im persönlichen Dienstleistungsbereich immer deutlicher, dass sensible Entscheidungen und emotionale Aufgaben zu einem hohen Grad weiterhin von Menschen ausgeübt werden. Zum einen, weil KI keine menschlichen Emotionen entwickeln kann (Simulationen sind möglich, aber bislang qualitativ nur sehr begrenzt mit menschlichen Emotionsausdrücken zu vergleichen). Zum anderen, weil der Mensch als emotionales Wesen sensible Informationen über sich selbst vor allem von Menschen akzeptieren dürfte, nicht alleine von einer Maschine. Beispielsweise im Versicherungsbereich, in denen die Bewertung der Versicherungsfälle vollautomatisiert abläuft, wird die Kommunikation bei abschlägig entschiedenen Entscheidungen in der Regel weiterhin von Menschen gemacht. Insgesamt müssen wir davon ausgehen, dass Tätigkeiten, die emotionale Intelligenz und Empathie voraussetzen, sowie jene, die ethische Entscheidungen erfordern, noch lange allein oder überwiegend den Menschen vorbehalten bleiben.

Je nachdem, welche konkreten Entscheidungen zum KI-Einsatz im Betrieb und am Arbeitsplatz tatsächlich getroffen werden, beeinflusst das auch die möglichen quantitativen Auswirkungen auf Beschäftigung in einem Betrieb, einer Branche, einem Sektor und schließlich gesamtwirtschaftlich. Wir wollen im nächsten Abschnitt die quantitativen Beschäftigungswirkungen vor dem Hintergrund unterschiedlicher Annahmen betrachten. Wir kommen zu dem Schluss, dass weder das Szenario einer drohenden technologischen Arbeitslosigkeit, noch das eines neuen *Pareto Optimums* überzeugend sind. Vielmehr gehen wir davon aus, dass KI und Digitalisierung zu einer weiteren Polarisierung von Arbeit, Beschäftigung, Einkommen und Status führen, je unkontrollierter sie in die Arbeitswelt Einzug halten.

## 4. Verschärfte Polarisierung am Arbeitsmarkt

Die Debatte über quantitative Beschäftigungseffekte des Einsatzes digitaler Technologien läuft bereits seit einigen Jahren. Etliche namhafte Studien sind zum Ergebnis gekommen, dass die Digitalisierung der Arbeitswelt zu massiven Verwerfungen in Form von Arbeitsplatzverlusten führen wird. Dies gilt auch – und gerade – für KI und den vermehrten Einsatz von maschinellen Lernsystemen.

Methodologisch steht hinter solchen Ergebnissen meistens ein Abgleich von Berufsprofilen in Arbeitsmarktstatistiken und einer Schätzung des Technikpotenzials durch KI-Entwickler oder durch das Erfassen von Patentanmeldungen. So wurde argumentiert, dass bis zur Hälfte aller Jobs in den nächsten zwei Jahrzehnten durch die Digitalisierung, vor allem aber durch die KI, wegrationalisiert werden könnten. Besonders gefährdet seien Berufe in Transport, Logistik und Produktion, aber auch im Dienstleistungssektor. Niedrige und mittlere Einkommensgruppen stehen im Fokus des Arbeitsplatzabbaus, weshalb die digitalisierungsgetriebene technologische Arbeitslosigkeit auch die Polarisierung vorantreiben würde. Allerdings könnten maschinelles Lernen und mobile (Leichtbau-)Robotik nicht nur manuelle Routine-Tätigkeiten, sondern auch kognitive Aufgaben ohne feste Vorgaben erledigen und auch am mittleren und oberen Ende der Skala für Verwerfungen sorgen (Frey/Osborne 2013; 2017; Muro et al. 2019).

Im Prinzip ist Konsens, dass ein Substitutionspotenzial von Arbeit durch die Digitalisierung besteht. Die Frage ist nur, wie groß ist dieses Potenzial, und welche Gegenkräfte wirken gegebenenfalls? Den Unterschied zwischen direkten Verwerfungen auf dem Arbeitsmarkt und abmildernden Effekten macht in der Regel eine „Nettokalkulation“ für die Gesamtgesellschaft (Arntz et al. 2017; Arntz et al. 2018; Dauth et al. 2017; Fuchs et al. 2018). Beschäftigung oder Tätigkeiten, die bedingt durch die Digitalisierung an der einen Stelle wegfallen, steht neue Beschäftigung in anderen Bereichen gegenüber. Steigende Produktivität in Folge der Digitalisierung zieht zudem gesamtwirtschaftliche Wettbewerbs-, Wertschöpfungs- und letztlich auch Beschäftigungseffekte nach sich (McKinsey Global Institute 2018; World Economic Forum 2018).

Eine gängige Annahme des Szenarios technologischer Arbeitslosigkeit ist, dass vor allem einfache Tätigkeiten automatisierbar sind. Je höher der Routineanteil einer Tätigkeit ist, desto größer ist ihr Substitutionspotenzial. Vertiefte Forschung hat aber deutlich gezeigt, dass Routine nicht gleich Routine ist, denn auch einfachste Routinetätigkeiten sind in Arbeitsprozesse und Arbeitsorganisationen eingegliedert, die nicht

ohne weiteres aufgebrochen und neu strukturiert werden können. Und selbst viele einfache Routinetätigkeiten sind deshalb so wertvoll, weil dafür schwer übertragbares und formalisierbares Erfahrungswissen gebraucht wird – etwas, das die KI bislang in vielen Fällen nicht hinreichend abbilden kann. Routinearbeit statisch, also isoliert und isolierbar zu begreifen, wird den Tätigkeiten deswegen häufig nicht gerecht und überzeichnet im Gegenzug das realistische Vermögen der KI. Menschliches Arbeitsvermögen dagegen ist vor allem qualitativ und kontextabhängig erfassbar und umfasst ein breites Spektrum (Pfeiffer/Suphan 2015). Wissen, das nicht formalisiert ist, das „stillschweigend“ ist (Autor 2015) geht über die formale Qualifikation hinaus und umfasst alle menschlichen Sinne wie Intuition, Bauchgefühl und Emotion – und es besteht aus Allgemeinwissen, also *Common Sense*, also genau dem Level des Verstehens, von dem KI und maschinelles Lernen noch sehr weit entfernt sind. Misst man das Arbeitsvermögen in diesem Verständnis (soweit das eben möglich ist), wird schnell klar, dass das Pendel noch nicht auf die Seite der KI geschwungen ist und auch noch eine lange Strecke vor sich hat, um an einfache informelle Fähigkeiten eines Menschen im Arbeitskontext heranzukommen.

Die binäre Gegenüberstellung von Routine- und Nicht-Routine-Tätigkeiten greift also zu kurz und überhöht verfrüht die digitalen Technologien einschließlich des Werkzeugs KI zum Ersatzmenschen. Offensichtlich gibt es aber die Effizienzsteigerungen bestehender Prozesse und von Arbeitsorganisationen: Fahrerlose Transportsysteme, Mensch-Roboter-Kollaboration (*Cobots*), Datenbrillen (*Smart Glasses*), 3D-Druck und additive Fertigung, digitale Assistenzsysteme, *Enterprise Resource Planning*, digitale Zwillinge und vieles mehr sind zunehmend Teil der betrieblichen oder auch überbetrieblichen Arbeitsteilung. Diese haben weitreichende Auswirkungen auf einzelne Arbeitsplätze, mit der in vielen Fällen eine Verdichtung von Arbeitsabläufen und Arbeitsbelastung einhergeht (Dispan/Schwarz-Kocher 2018). Neben den qualitativen Aspekten dieser Veränderungen für die Arbeit, führt das auch zu *möglichen* quantitativen Veränderungen. Aber daraus direkt auf konkrete (negative) quantitative Beschäftigungseffekte Schlussfolgerungen ziehen zu wollen, wird der Komplexität auch von Routinetätigkeiten nicht gerecht.

Andererseits gehen mit dem Substitutionspotenzial auch positive quantitative Beschäftigungserwartungen einher. Denn mit der Verdichtung von Arbeit ist – auf jeden Fall maximal bis zur individuellen „*Burn-out-Grenze*“ – ein Produktivitätszuwachs verbunden. Digitalisierung und der Einsatz von KI bedeuten Investitionen und Investitionen können entweder mehr oder weniger Beschäftigung bedeuten. Im optimistischen Szenario führt der digitale Technikeinsatz zu einem technologischen *Up-*

*grade* und verändert die Kapitalzusammensetzung in einem Betrieb. Die steigende Gesamtbeschäftigung ist dann eine Folge der gestiegenen Nachfrage nach einem bestimmten Kapitaltyp. Der Anstieg der Produktnachfrage für Branchen, welche die *Inputs* für diesen Kapitaltyp bereitstellen, führt zu steigender Beschäftigung in der Wirtschaft als Ganzes.

Darüber hinaus verändern die Investitionen in Digitalisierungstechniken die Kostenstruktur und damit die relative Wettbewerbsfähigkeit eines Unternehmens. Betriebe, die durch Digitalisierung ihre Kosten senken, können ihre Preise senken und die Nachfrage nach ihren Produkten und Dienstleistungen entsprechend steigern (vorausgesetzt ist dabei eine konstante Nachfrage aus anderen Teilen der Wirtschaft) (Arntz et al. 2018). In der Folge steigt der *Output* des investierenden Unternehmens und es entsteht in diesen Unternehmen neues Einkommen in Form von Löhnen, Gewinnen und Kapitaleinkommen. Wesentlich für dieses Szenario ist der Wettbewerbseffekt sowie die Aufteilung des Ertrags in Kapital- und Arbeitseinkommen. Die steigende Produktivität senkt die Produktionskosten automatisierter Tätigkeiten, was entweder zu Wachstum der Gewinne und/oder einer gesteigerten Nachfrage nach Arbeit nicht-automatisierter Tätigkeiten führen kann. In entsprechenden Simulationen kommt es demnach häufig zu einer langfristig steigenden Nachfrage nach Arbeitskräften in der Gesamtwirtschaft (Fuchs et al. 2018).

Betrachtet man diese und ähnliche Berechnungen näher, beruhen sie stark auf den *Spill-over*-Effekten von einem Sektor zum anderen. Dieser ist aber tatsächlich voraussetzungsvoller, als das häufig angenommen wird. Mehrere Probleme tun sich auf: Steigt die Produktivität von Kapital relativ zu Arbeit, ist technologische Innovation im Grundsatz arbeitensparend und es ist eher unwahrscheinlich, dass steigende Produktivität zu steigenden Durchschnittslöhnen führt. Und falls Arbeit schneller durch Technik ersetzt als Neue geschaffen wird, ersetzt Technik Arbeit, ohne dass es zu steigender Arbeitsnachfrage in anderen Sektoren kommen muss. Digitalisierungsentscheidungen sind am Ende des Tages für das einzelne Unternehmen eine Folge der Kalkulation von relativen Faktorpreisen, also der relativen Preise *aller* notwendigen Produktionsfaktoren. Somit ist es nicht unwahrscheinlich, dass Digitalisierung eher mit sinkenden Durchschnittslöhnen einhergeht, weil der Substitutionseffekt zur sinkenden Nachfrage nach Arbeit führt (Acemoglu/Restrepo 2018). Im Ergebnis entsteht dann zunächst einmal wachsende Ungleichheit am Arbeitsmarkt und in der Gesamtwirtschaft (Korinek/Stiglitz 2017). In diesem Kontext wird auch der Bedarf nach neuer Regulierung oder Umverteilung von Innovationsgewinnen deutlich, weil sonst das Gesamtergebnis technologischer Innovation schlechter für die Ge-

sellschaft als Ganzes sein kann, als die Situation vor der Innovation (Acemoglu 2019).

Selbst wenn höhere Löhne in Folge der Digitalisierung in einer Branche oder der Gesamtwirtschaft Realität werden, heißt das noch nicht, dass sich die *potenzielle* Kaufkraft in neuen *tatsächlichen* Kaufaktivitäten – in demselben oder anderen Wirtschaftsbereichen – realisiert. Viel hängt von anderen Faktoren ab, z. B. wie die Digitalisierung die Nachfrage neu strukturiert (durch neue Kaufmuster etc.), ob Hoch- oder Geringverdiener davon profitieren oder welche Sparquote eine Wirtschaft aktuell prägt. Bis heute ist empirisch beobachtbar, dass höher qualifizierte Arbeit von der Digitalisierung profitiert, wobei die menschliche Tätigkeit überwiegend komplementär ergänzt wird. Hier gibt es eine *elastische* Nachfrage für die entsprechenden Produkte und Dienstleistungen dieser Arbeitskräfte; es gibt aber ein *unelastisches* Arbeitskräfteangebot in diesen Bereichen (Fachkräftemangel). Anders sieht das Bild bei Tätigkeiten mit niedrigeren Qualifikationsprofilen aus. Gerade die Nachfrage nach manuellen Tätigkeiten ist relativ preisunelastisch: Fällt der Preis manueller Tätigkeiten durch Digitalisierung, steigt ihre Nachfrage nicht in gleichem Maße. Wir erleben also bereits die Vorstufe der „Polarisierung 4.0“ (vgl. Autor et al. 2017; Autor/Salmons 2017).

Der von Seiten etlicher Akteure unterstellte Automatismus zwischen Digitalisierung, Produktivitätssteigerungen und einer steigenden (gesamtwirtschaftlichen) Nachfrage nach Arbeit erweist sich somit als sehr zweifelhaft. Ob die Digitalisierung und der zunehmende Einsatz von KI zu positiven oder negativen Effekten für Arbeit und Beschäftigung führen wird, bleibt weitgehend offen. Viel hängt von den Weichenstellungen und dem Regulierungsrahmen ab. Eine Laissez-faire-Politik wäre demnach kein guter Weg in die digitalisierte Arbeitsgesellschaft. Im Gegenteil: Eine unkontrollierte Digitalisierung mit KI an der Spitze birgt das Potenzial, unsere Arbeitsgesellschaft aus dem „Gleichgewicht“, in dem sie sich befindet, zu bringen, mit dem Ergebnis (zu) weniger Gewinner und (zu) vieler Verlierer.

## 5. Konturen einer ganzheitlichen Technikfolgenabschätzung

Es stellt sich auf absehbare Zeit nicht die grundsätzliche Frage, ob in Zukunft noch gearbeitet werden muss oder soll. Technologische Arbeitslosigkeit als potenziell umfassendes und dauerhaftes Szenario ist vor allem ein Szenario, das aus der Technikentwicklung selbst stammt, ebenso wie das eines neues *Pareto Optimum* bei Arbeit und Beschäftigung.

Um eine sozial, ökonomisch und ökologisch ausgewogene Konturierung der Arbeitsgesellschaft in einem hochdynamischen Prozess wie der Digitalisierung erreichen zu wollen, reicht es aber nicht aus, nur den Technologieeinsatz zu bewerten und entsprechende Maßnahmen regulativ feinzustimmen. Denn tatsächlich verlangt die schnelle Weiterentwicklung der digitalen Kerntechnologien, mit KI und maschinellem Lernen an der aktuellen Technikgrenze, nach einer *kontinuierlichen und ganzheitlichen* Betrachtung von Arbeit und der sozioökonomischen Dimensionen der Technik (Kellermann/Obermaier 2020).

Wie eingangs erwähnt, müssen die relevanten Dimensionen für eine solche Konturierung erst noch bestimmt werden. Dies ist schwieriger als es auf den ersten Blick vielleicht aussieht, da die Entwicklungen – technologisch, betrieblich, gesamtgesellschaftlich – hochdynamisch sind. Die Bestimmung der ersten diskutierten Dimension – die Technik und die Vorhersage ihrer Potenziale – ist selbst für KI-Entwickler schwer bis unmöglich. Sowohl Soft- als auch Hardware stehen stets an harten Grenzen der Weiterentwicklung. Eine Entwicklung über diese Grenzen hinaus ist nicht unmöglich, aber sie ist nach wie vor eine *Wildcard*, in deren Hintergrund wiederum viele Detaillösungen für unterschiedlichste Bereiche und Arbeitskontexte betreffend vorangetrieben werden. Es ist gleichzeitig nicht auszuschließen, dass aus vielen Einzel-Intelligenzen eine KI wird, die zunehmend das kann, was mit ihr von Beginn an verbunden wird, nämlich eine Form von *General Purpose Intelligence* (Russel 2019) – aber das ist Stand heute ein eher unwahrscheinliches Szenario.

Es gilt also die technologische Entwicklung kontinuierlich und sehr genau unter (gesellschaftlicher) Beobachtung zu halten. Eine öffentliche, unabhängige, auf den einzelnen Anwendungsfall wie auch auf gesellschaftliche Effekte ausgerichtete Technikfolgenabschätzung ist insofern eine wesentliche Grundvoraussetzung, das Technikpotenzial sowie mögliche gesellschaftliche Effekte konkret zu bewerten. Hierfür ist z. B. das KI-Observatorium des deutschen Bundesministeriums für Arbeit und Soziales ein wichtiger Schritt in die richtige Richtung (vgl. [www.denkfabrik-bmas.de/projekte/ki-observatorium](http://www.denkfabrik-bmas.de/projekte/ki-observatorium)).

Neben den technologischen Fortschritten müssen die laufenden Transformationen von Arbeit, Branchen und einzelnen Arbeitsplätzen und -kontexten kontinuierlich erfasst werden. Repräsentative Auswertungen des KI-Einsatzes im Arbeitskontext sind bisher noch Mangelware, sowohl in Deutschland als auch im europäischen Ausland. Eine Ausnahme ist das gemeinsam von dem amerikanischen IT-Unternehmen IBM und der deutschen Dienstleistungsgewerkschaft Ver.di in Auftrag gegebene Forschungsprojekt zum Einsatz von Watson-KI bei IBM Kunden (IBM 2019). In diesem Projekt wird konkret der Frage nachgegangen, welche Effekte für Dienstleistungstätigkeiten der KI-Einsatz nach sich zieht.

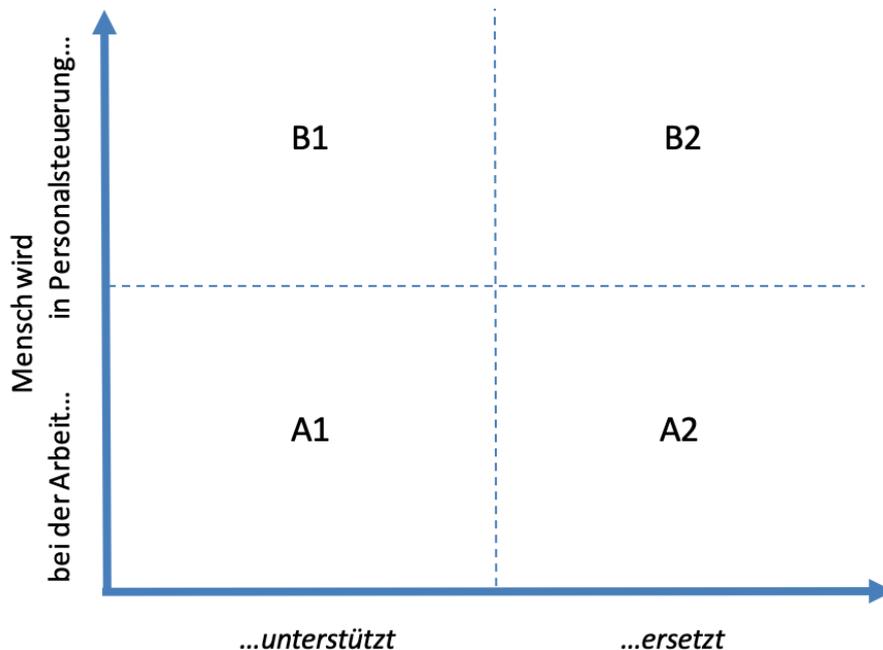
Für eine breit angelegte, arbeitsfokussierte Technikfolgenabschätzung ist es notwendig, die verschiedenen Einsatzformen von KI im Arbeitskontext zu kategorisieren und kontinuierlich zu erfassen. Dabei kommt es nicht primär auf die technischen Unterschiede zwischen verschiedenen KI-Systemen oder deren technischen Reifegrad an, sondern auf die Rolle, die die KI im Unternehmen spielt oder spielen soll. Je nach Rolle ergeben sich unterschiedliche Anforderungen an die Regulierung und Transparenz von KI sowie an die Rahmenbedingungen für die Beteiligung und Befähigung von Mitarbeitern und ihren Vertretungen in einem Betrieb.

Eine solche Technikfolgenabschätzung braucht geeignete Kategorisierungen und *Tools*, die innerbetrieblich sowie überbetrieblich eingesetzt werden können. Für eine solche Kategorisierung können zunächst zwei Dimensionen der KI-Anwendung am Arbeitsplatz ausreichen, um eine erste Matrix über die arbeitsorganisatorischen Folgen des KI-Einsatzes im Betrieb zu bilden (Albrecht 2020). Dabei ist zum einen „horizontal“ zu differenzieren, ob die KI als Werkzeug dient, das die Arbeit von Mitarbeiterinnen und Mitarbeitern unterstützt oder ersetzt (A) oder ob sie als Instrument der Personalsteuerung eingesetzt wird (B): Beispiele für Kategorie A sind z. B. Chatbots in der Kommunikation mit Kunden oder Bewertungen eines Versicherungsfalls. Beispiele für Kategorie B ist z. B. der Einsatz von KI um Schichtpläne zu erstellen, Teams zusammenzusetzen, Leistungsbeurteilungen oder eine Bewerberauswahl zu bestimmen. Zum anderen geht es „vertikal“ um die Frage nach der Letztentscheidung: Hier ist danach zu differenzieren, ob die Entscheidungskompetenz beim Menschen bleibt (1) oder von der KI übernommen wird (2).

Im Hinblick auf die Rolle, die KI in einem Unternehmen übernehmen kann, sind daher folgende vier Einsatzmodelle zu unterscheiden: Die Technologie kann ein Werkzeug sein, das die Arbeitnehmer unterstützt (A1) oder teilweise ersetzt (A2). Und sie kann (zusätzlich) Entscheidun-

gen in der Personalsteuerung unterstützen (B1) oder selbständig übernehmen (B2).

Abbildung 1: Schematische Darstellung unterschiedlicher Grade der Eingriffstiefe der KI in den Arbeitsprozess



Aus den unterschiedlichen Kategorien können die unterschiedlichen Anforderungen an die Rahmenbedingungen klarer definiert werden, unter denen die KI so eingesetzt werden kann, dass sie einer sozial und ökonomisch nachhaltigen Gestaltung von Arbeit im Sinne Guter Arbeit gerecht wird. Dabei ist davon auszugehen, dass der (technische) Regulierungsbedarf von A1 eher gering, während er bei B2 eher hoch ist. So stellen beispielsweise Arbeitsplattformen einen Extremfall der Kategorie B2 dar, weil mehr oder weniger die gesamte „Betriebsorganisation“ eigenständig von letztentscheidender KI übernommen wird.

Diese Kategorisierung kann dazu beitragen, eine Technikfolgenabschätzung für den KI-Einsatz am Arbeitsplatz zu etablieren. Dies erscheint notwendig, weil die Entwicklung der gesellschaftlichen Institutionen einschließlich des einzelnen Arbeitsplatzes nicht allein von den technologisch-disruptiven Entwicklungen getrieben werden darf. Es müssen vielmehr Institutionen geschaffen werden, mit denen der Einsatz der neuen Technologien auf die gesellschaftlichen Normen und Bedarfe abstimmt werden kann, sowohl am einzelnen Arbeitsplatz als auch auf die Rolle der Arbeit in der Arbeitsgesellschaft insgesamt.

Dafür sind zunächst vier Handlungsebenen von besonderer Bedeutung: es braucht (1) einen geeigneten politisch-regulativen Rahmen, (2) Transparenz über Ziele und Funktionsweise des KI- oder Technologie-Einsatzes im konkreten Anwendungsfall, (3) die Beteiligung der Beschäftigten beim Einsatz und der Nutzung der Technologien und (4) die Befähigung der Beschäftigten, mit KI konstruktiv-kritisch umzugehen.

(1) Eine Überprüfung (und Weiterentwicklung) der bestehenden Regularien und Institutionen erscheint im Hinblick auf den KI-Einsatz in der Arbeitswelt in mehreren Bereichen notwendig. Ein zentrales Feld ist der Datenschutz. Die europäische Datenschutzgrundverordnung (DSGVO) beinhaltet wichtige Prinzipien, wie „*privacy by design*“ und „*privacy by default*“, und Regeln, die auch für KI-Systeme aller Kategorien gelten und umgesetzt werden müssen. Zudem eröffnet sie Spielraum für spezifischere Regelungen für den nationalen Gesetzgeber beim Arbeitnehmerschutz aufgrund des besonderen Abhängigkeitsverhältnisses der Arbeitnehmer. Solche besonderen Schutzrechte dürften besonders für einen KI-Einsatz nach B1 und B2 in der Matrix eine notwendige Handlungsoption sein. Offensichtlich betrifft die Regulierung aber auch weitere Bereiche, wie die Anpassung des Arbeitsschutzes, das Arbeitsrecht und die Rechte der Mitbestimmungsakteure. Ein Regulierungsbeispiel wäre ein Verbot vollautomatischer Entscheidungen mit Blick auf Kündigungen.

(2) Von zentraler Bedeutung für den Einsatz von KI ist auch die Transparenz über deren Ziele und Funktionsweisen, soweit dies bei selbstlernenden Systemen möglich ist. Sie herzustellen, bedeutet vor allem offenzulegen, welche Kriterien für die Datenerfassung und -Evaluation und welche Entscheidungen auf ihrer Grundlage getroffen (Kategorien A1 und B1) oder durch sie gefällt werden (A2 und B2). Dabei ist die wichtigste Frage, ob der Mensch die Technik kontrolliert oder die Technik den Menschen. Es geht dabei nicht darum, Entscheidungen von KI gänzlich auszuschließen, aber es ist notwendig, ein System zu schaffen in dem der Mensch nicht zum bloßen „Anhängsel“ des Maschinenbefehls wird, sondern im Gesamtprozess und kritischen Einzelsituationen die Entscheidungshoheit hat. Besonders sensibel ist diese Frage, wenn KI auch bei der Personalsteuerung eingesetzt wird (Kategorien B1 und B2). Selbst wenn diese Entscheidungen „nur“ auf Grundlage von KI-Empfehlungen im (Personal-)Management getroffen werden (B1), gibt es das Problem, dass es dem Management selbst häufig unmöglich ist, die KI-Empfehlung nachzuvollziehen. Das liegt zum einen daran, dass viele KI-Anbieter keine Einblicke in ihre Systeme zulassen. Zum anderen geht bei selbstlernenden Systemen irgendwann die Nachvollziehbarkeit verloren, solange das System keine „Erklärungsfunktion“ besitzt.

(3) Der Einsatz von KI-Systemen in spezifischen nationalen Betriebs- und Arbeitskontexten bedeutet auch, KI in den jeweiligen Mitbestimmungsstrukturen integrieren zu müssen. In Deutschland heißt das, dass Mitarbeiter und ihre Interessenvertretungen frühzeitig beteiligt werden müssen (in anderen europäischen Ländern gelten andere Mitbestimmungsregelungen rund um KI). Sicherzustellen, dass diese frühe Form der Beteiligung auch wirklich und transparent stattfindet, ist eine Herausforderung. In der Praxis verfügen häufig Arbeitgeber selbst nicht über die nötigen Informationen, die sie ihren Beschäftigten und deren Vertretungen zugänglich machen müssten, weil KI-Anbieter sie nicht transparent machen. Weiterer Regulierungsbedarf ergibt sich aus der Tatsache, dass KI-Systeme im Betrieb lernen und sich damit, ganz anders als bei nicht-lernenden IT-Systemen, permanent die technischen Grundlagen verändern. Diese dynamische Eigenschaft von lernenden Systemen ist für alle Beteiligten von wesentlicher Bedeutung, die bislang regulatorisch untererfasst ist. Vereinbarungen zwischen den Sozial- und Betriebspartnern zu KI-Systemen werden daher einen neuen Charakter bekommen müssen: Anstelle konstanter Regeln werden sie regelmäßige Konsultationen zwischen den Betriebsparteien und neue zentrale und dezentrale Konfliktlösungsmechanismen vorsehen müssen.

(4) Solche „Prozessvereinbarungen“ sind zwar für Gewerkschaften und Mitarbeitervertretungen nicht völlig neu, aber um die spezifischen Herausforderungen des KI-Einsatzes bewältigen zu können, wird es notwendig sein, die Beschäftigten in technischen, gesetzlichen und vor allem auch kulturellen Fragen weiterzubilden. Das eröffnet auch den Gewerkschaften ein neues Feld: Mit dem Aufbau eines betriebsübergreifenden Know-how zu KI-Fragen könnten sie Mitarbeitervertretungen sowie Beschäftigten schulen und unterstützen. Sie könnten sich z. B. an einer Zertifizierung von KI-Anwendungen im Hinblick auf die Verhaltens- und Leistungskontrollfunktionen beteiligen, jedenfalls bei Standardanwendungen. Eine solche Zertifizierung könnte dabei helfen, Blockaden bei der Einführung von KI-Systemen zu vermeiden, indem sie die gerade beschriebene prozesshafte Begleitung durch die Mitbestimmungsträger erleichtern, wenn auch, aufgrund der ebenfalls beschriebenen permanenten Veränderlichkeit der Systeme, nicht ersetzen.

Schließlich wird die Qualifizierung von Mitarbeitern entscheidend für die erfolgreiche Einführung und Nutzung von KI-Systemen sein. Es muss sichergestellt werden, dass Arbeitnehmer befähigt werden, mit den KI-Systemen umzugehen und ggf. neue Aufgaben zu übernehmen, wenn ein Teil ihrer bisherigen Aufgaben durch KI-Systeme substituiert wird. Die Qualifizierungsmaßnahmen müssen sowohl den Umgang mit dem betriebs- oder tätigkeitspezifischen System als auch ein grundle-

gendes Verständnis des Systems und deren dahinterliegende Logik selbst umfassen, sowie neue Fähigkeiten vermitteln, die durch veränderte Tätigkeiten bei der Nutzung von KI-Systemen entstehen.

Diese vier Handlungsebenen tragen zu einem neuen gesellschaftlichen Rahmen bei, der die disruptiv-dynamische Technologieentwicklung in einem vorhersehbaren und transparenten institutionellen Kontext verankern kann. Gesellschaftliches Ziel ist es, durch einen solchen institutionellen Rahmen, die Kontrolle über die Technik nicht zu verlieren, bzw. der KI und der Digitalisierung den notwendigen Entwicklungs- und Anwendungsspielraum zu geben aber gleichzeitig die Auswirkungen der Technologie am einzelnen Arbeitsplatz, auf betroffene Berufsgruppen und auf die Gesamtwirtschaft kalkulieren zu können und verlässliche Überprüfungsmechanismen auf verschiedenen Ebenen zu etablieren. Sonst droht die Gefahr einer verschärften Polarisierung in der Arbeitsgesellschaft, der Einkommen und damit auch eine verstärkte Polarisierung der Gesellschaft als Ganze. Schaffen wir es nicht, zu regulieren, ob und wie die „intelligente“ Maschine für und mit den Menschen arbeiten soll, hat das letztlich auch fatale Folgen für die normative Frage, wie, womit und warum wir arbeitsteilig füreinander weiter arbeiten wollen.

## Literatur

- Acemoglu, D. (2019): *Elizabeth Warren's Bold Ideas Don't Go Far Enough*. Project Syndicate. <https://www.project-syndicate.org/commentary/good-jobs-agenda-us-by-daron-acemoglu-2019-12> (zuletzt abgerufen am 31.10.2020).
- Acemoglu, D./Restrepo, P. (2018). *Artificial Intelligence, Automation and Work*. NBER Working Paper No. 24196. <https://www.nber.org/papers/w24196.pdf> (zuletzt abgerufen am 31.10.2020).
- Albrecht, T. (2020) Regulation, Transparenz, Beteiligung und Befähigung – Voraussetzungen für den erfolgreichen Einsatz Künstlicher Intelligenz am Arbeitsplatz. In: Jens Nachtwei & Antonia Sureth (2020). Sonderband „Zukunft der Arbeit“, HR Consulting Band 12, Berlin: Veröffentlichungsreihe für Qualitätssicherung in Personalauswahl und -entwicklung (i.E.).
- Arntz, M./Gregory, T./Zierahn, U. (2017). Revisiting the Risk of Automation. In: *Economics Letters*, Nr. 159, S. 157–160.
- Arntz, M./Gregory, T./Zierahn, U. (2018). *Digitalisierung und die Zukunft der Arbeit: Makroökonomische Auswirkungen auf Beschäftigung, Arbeitslosigkeit und Löhne von morgen*. Bundesministerium für Forschung und Entwicklung (BMBF), Mannheim. <http://ftp.zew.de/pub/zew-docs/gutachten/DigitalisierungundZukunftderArbeit2018.pdf> (zuletzt abgerufen am 31.10.2020).
- Autor, D. (2015). Why are There Still So Many Jobs? The History and Future of Workplace Automation. In: *Journal of Economic Perspectives*, Nr. 29(3), S. 3–30. <https://economics.mit.edu/files/11563> (zuletzt abgerufen am 31.10.2020).
- Autor, D./Dorn, D./Katz L.F./Patterson, C./Van Reenen, J. (2017). Concentrating on the Falling Labor Share. In: *American Economic Review Papers & Proceedings*, Nr. 107(5). S. 180–185. <https://economics.mit.edu/files/12544> (zuletzt abgerufen am 31.10.2020).
- Autor, D./Salomons, A. (2017). Robocalypse Now. Does Productivity Growth Threaten Employment? In: *ECB Forum on Central Banking – Investment and growth in advanced economies – Conference Proceedings*. Frankfurt am Main: European Central Bank, S. 119–128. <https://www.ecb.europa.eu/pub/pdf/other/ecb.ecbforumcentralbanking2017.en.pdf> (zuletzt abgerufen am 31.10.2020).

- Bessen, J. (2018). AI and Jobs: The Role of Demand. In: *NBER Working Paper Series*, Nr. 24235. <https://www.nber.org/papers/w24235.pdf> (zuletzt abgerufen am 31.10.2020).
- Dauth, W./Findeisen, S./Südekum, J./Wößner, N. (2017). German Robots – The Impact of Industrial Robots on Workers. In: *IAB-Discussion Paper*, 30/2017. <http://doku.iab.de/discussionpapers/2017/dp3017.pdf> (zuletzt abgerufen am 31.10.2020).
- Dispan, J./Schwarz-Kocher, M. (2018). Digitalisierung im Maschinenbau. Entwicklungstrends, Herausforderungen, Beschäftigungswirkungen, Gestaltungsfelder im Maschinen- und Anlagenbau. In: Working Paper Forschungsförderung, Nummer 094. [https://www.boeckler.de/pdf/p\\_fofoe\\_WP\\_094\\_2018.pdf](https://www.boeckler.de/pdf/p_fofoe_WP_094_2018.pdf) (zuletzt abgerufen am 31.10.2020).
- Frey, C.B./Osborne, M.A. (2013). *The Future of Employment: How Susceptible are Jobs to Computerisation?* Oxford Martin School Working Paper. University of Oxford. [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf) (zuletzt abgerufen am 31.10.2020).
- Frey, C.B./Osborne, M.A. (2017). The Future of Employment: How sSusceptible are Jobs to Computerisation? In: *Technological Forecasting & Social Change*. Nr. 114, S. 254–280.
- Fuchs, J./Kubis, A./Schneider, L. (2019): Zuwanderung und Digitalisierung. Wie viel Migration aus Drittstaaten benötigt der deutsche Arbeitsmarkt künftig? Bertelsmann Stiftung. [https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Migration\\_fair\\_gestalten/IB\\_Studie\\_Zuwanderung\\_und\\_Digitalisierung\\_2019.pdf](https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Migration_fair_gestalten/IB_Studie_Zuwanderung_und_Digitalisierung_2019.pdf) (zuletzt abgerufen am 31.10.2020).
- IBM (2019): KI-Studie über die Folgen für Arbeitnehmende und Arbeit. <https://www.ibm.com/de-de/blogs/think/2019/09/17/watson-ki-studie/> (zuletzt abgerufen am 31.10.2020).
- Kellermann, C./Obermayer, R. (2020). Von der Würde der Arbeit in digitaler und klimaneutraler Zukunft. spw 238. [https://www.spw.de/data/238\\_kellermann\\_obermayer.pdf](https://www.spw.de/data/238_kellermann_obermayer.pdf) (zuletzt abgerufen am 31.10.2020).
- Korinek, A./Stiglitz, J.E. (2017). Artificial Intelligence and Its Implications for Income Distribution and Unemployment. In: *National Bureau of Economic Research*. <https://www.nber.org/papers/w24174.pdf> (zuletzt abgerufen am 31.10.2020).

- McKinsey Global Institute [MGI] (2018). *Notes from the AI frontier. Modeling the Impact of AI on the World Economy*.  
<https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy> (zuletzt abgerufen am 31.10.2020).
- Muro, M./Maxim, R./Whiton, J. (2019). Automation and Artificial Intelligence. How Machines are Affecting People and Places. Metropolitan Policy Program. [https://www.brookings.edu/wp-content/uploads/2019/01/2019.01\\_BrookingsMetro\\_Automation-AI\\_Report\\_Muro-Maxim-Whiton-FINAL-version.pdf](https://www.brookings.edu/wp-content/uploads/2019/01/2019.01_BrookingsMetro_Automation-AI_Report_Muro-Maxim-Whiton-FINAL-version.pdf) (zuletzt abgerufen am 31.10.2020).
- Nilsson, N. (2005). Human-Level Artificial Intelligence? Be Serious! American Association for Artificial Intelligence.  
<http://ai.stanford.edu/~nilsson/OnlinePubs-Nils/General%20Essays/AIMag26-04-HLAI.pdf> (zuletzt abgerufen am 31.10.2020).
- Pfeiffer, S./Suphan, A. (2015). *Der AV-Index. Lebendiges Arbeitsvermögen und Erfahrung als Ressourcen auf dem Weg zu Industrie 4.0*. Working Paper 2015 #1 (Finale Fassung des ursprünglich am 13.04.2015 publizierten Drafts). <https://www.sabine-pfeiffer.de/files/downloads/2015-Pfeiffer-Suphan-final.pdf> (zuletzt abgerufen am 31.10.2020).
- Russel, S. (2019). *Human Compatible. AI and the Problem of Control*. Allen Lane.
- Webb, M. (2019). *The Impact of Artificial Intelligence on the Labour Market*. Stanford University Working Paper.
- World Economic Forum (2018). *The Future of Jobs Report (2018)*.  
[http://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs\\_2018.pdf](http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf) (zuletzt abgerufen am 31.10.2020).

## Autoren

**Thorben Albrecht** ist ehemaliger Staatssekretär im Bundesministerium für Arbeit und Soziales und war Mitglied in der „Global Commission on the Future of Work“ der ILO.

**Dr. Christian Kellermann** lehrt an der Hochschule für Technik und Wirtschaft in Berlin und ist Geschäftsführer des Denkwerk Demokratie e. V. Zuvor war er Geschäftsführer des Instituts für die Geschichte und Zukunft der Arbeit.

Wir danken Dr. Kajsa Borgnäs für wertvolle Kommentare und Hinweise.

---

Die Gestaltung der Zukunft der digitalen Arbeitsgesellschaft setzt eine umfassende Technikfolgenabschätzung voraus. Sie muss die qualitativen Veränderungen von Arbeit und eine belastbare quantitative Schätzung der Auswirkungen auf die Gesamtbeschäftigung zusammenführen. Dazu ist es notwendig, die verschiedenen Einsatzformen von KI im Arbeitskontext zu kategorisieren und kontinuierlich zu erfassen. Ohne die Verschränkung der Analyse- und Handlungsebenen droht der digitalen Arbeitsgesellschaft die Gefahr einer verschärften Polarisierung der Einkommen und damit auch der Gesellschaft insgesamt.

---