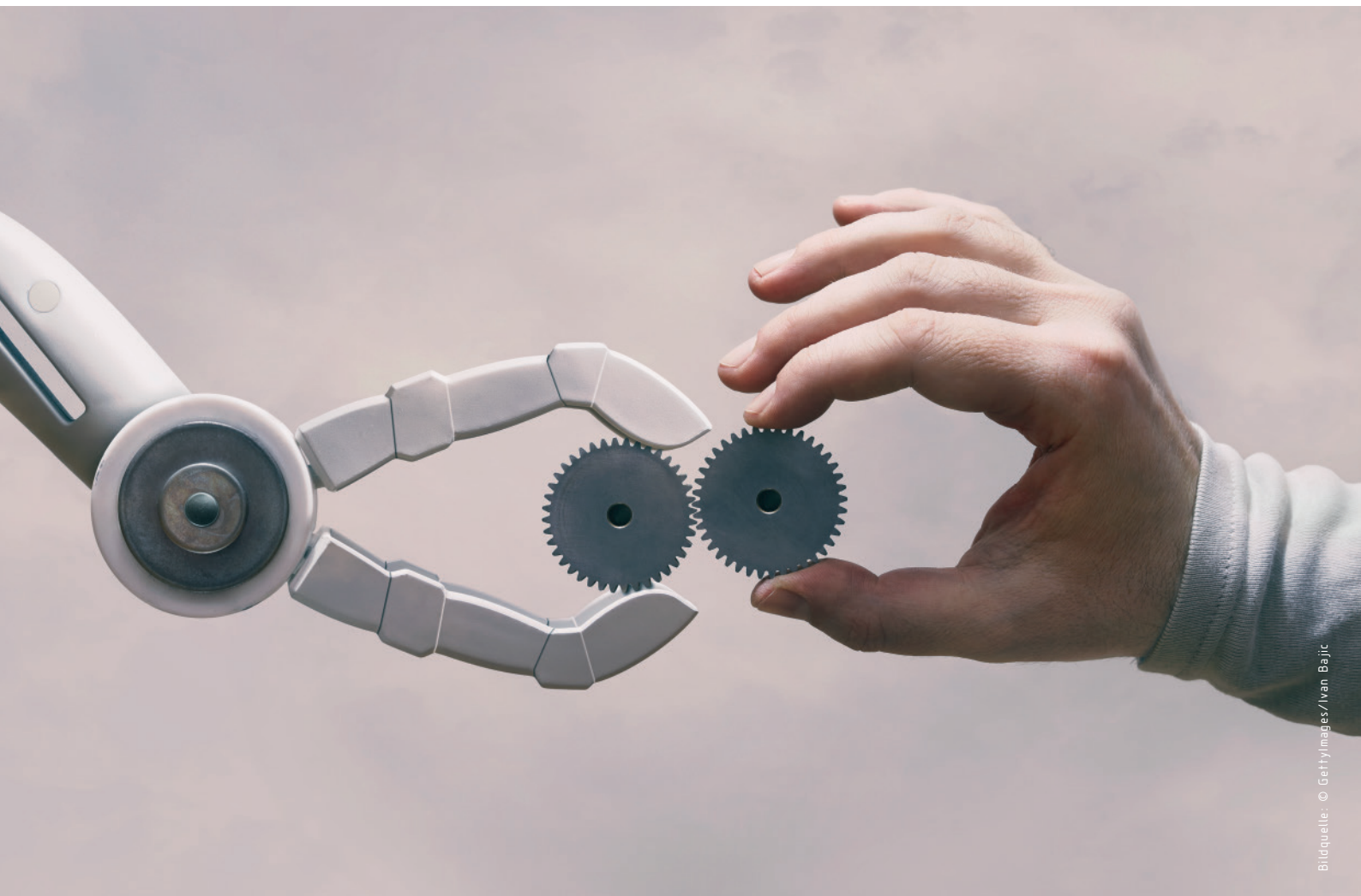


HANDLUNGSLEITFADEN

FÜR DIE SOZIALPARTNER

HUMANZENTRIERTE KÜNSTLICHE INTELLIGENZ
IN DER CHEMISCHEN INDUSTRIE



Ergebnisse des Verbundprojekts
„Humanzentrierte KI in der chemischen
Industrie (hKI-Chemie)“
gefördert vom Bundesministerium für Bildung
und Forschung (BMBF)

(Kontakt: malanowski@vdi.de)

Inhalt

| | |
|--|----|
| Hintergrund | 5 |
| Auf einen Blick: Welche Fragen gilt es zu klären? | 7 |
| Die Kriterien einer humanzentrierte KI in der chemischen Industrie | 7 |
| Welche Handlungsoptionen bieten sich der betrieblichen Ebene? | 11 |
| Der Blick nach vorn | 12 |
| Weiterführende Informationen | 13 |
| Literatur | 13 |

Vorwort

Die Fähigkeiten von Systemen, die auf Künstlicher Intelligenz (KI) basieren, konnten in den letzten Jahren beachtliche Fortschritte erzielen. Gleichwohl besteht nach wie vor eine große Diskrepanz zwischen der Leistungsfähigkeit von KI und erfahrenen Mitarbeiter*innen. KI-Systeme sind datengetrieben und lassen sich gut auf Aufgaben trainieren, die einen gleichbleibenden Bezugsrahmen haben. Plötzliche oder auch schleichende Veränderungen der äußeren Rahmenbedingungen stellen KI-Systeme hingegen vor große Probleme: Zu den veränderten Bedingungen gibt es nicht sofort ausreichend Daten, um die KI mit Deep-Learning Ansätzen neu zu trainieren. Menschen hingegen sind sehr gut darin, Gelerntes schnell auf neue Problemstellungen zu adaptieren. Um für die jeweilige Situation das beste Ergebnis zu erzielen, ist es daher sinnvoll, die unterschiedlichen Stärken von Mensch und KI zusammenzubringen. Es geht also nicht darum, Menschen durch KI zu ersetzen, sondern Menschen durch KI so zu unterstützen, dass die Symbiose aus Mensch und KI die bestmöglichen Ergebnisse liefert.

Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Verbundprojekt „Humanzentrierte Künstliche Intelligenz in der Chemischen Industrie (hKI-Chemie)“ erforscht, entwickelt und erprobt Mechanismen, die geeignet sind, Mensch und KI gemeinsam an einer Problemlösung arbeiten zu lassen. Einerseits soll die KI mit Hilfe des Menschen lernen, sich schneller auf neue Situationen einzustellen. Andererseits soll der Mensch von dem Wissen profitieren,

das KI-Systeme aus großen Datenmengen extrahieren können. Die Arbeiten im Verbundprojekt werden an konkreten Anwendungsbeispielen (Use Cases) aus der chemischen Industrie durchgeführt. Sie zeigen, wie KI-Systeme und ihre Interaktion mit dem Menschen so entworfen werden können, dass am Ende sowohl die Qualität der Entscheidungen als auch die Wirksamkeit sowie der Ideenreichtum der Mitarbeiter*innen gesteigert werden kann.

Der vorliegende Handlungsleitfaden für eine humanzentrierte Künstliche Intelligenz in der chemischen Industrie richtet sich an das Technische Management, Manager*innen aus dem Bereich Human Resources und Betriebsräte, die sich mit Künstlicher Intelligenz bereits beschäftigen bzw. planen, hier in naher Zukunft aktiv zu werden. Damit werden hier ca. 1.900 Unternehmen der chemischen Industrie mit ca. 17.000 Betriebsräten angesprochen.

Ein besonderer Dank gilt dem Bundesarbeitgeberverband Chemie (BACV) und der Industriegewerkschaft Bergbau Chemie Energie (IGBCE). Beide Organisationen haben die Ausarbeitung des vorliegenden Handlungsleitfadens sehr engagiert unterstützt.

Prof. Dr.-Ing. Torben Weis
Universität Duisburg-Essen
Fachgebiet Verteilte Systeme

Hintergrund

Zunehmende Relevanz von KI für Arbeitswelten

Das Thema Künstliche Intelligenz (KI) erfährt gegenwärtig nicht nur eine große Aufmerksamkeit in der Wissenschaft, sondern auch in Wirtschaft, Politik und Gesellschaft. So arbeiten diverse industrielle Größen aus der chemischen Industrie in Asien, Nordamerika und Europa intensiv am Einsatz Künstlicher Intelligenz in Bereichen wie Forschung und Entwicklung, Wartung und Instandhaltung, Logistik oder Produktnachverfolgung. Die dazu notwendigen Technologien unterliegen einer rapiden Entwicklung und erreichen laufend Verbesserungen. Durch die breiter werdende Anwendung von KI dringt das Thema immer stärker in gegenwärtige und zukünftige Arbeitswelten ein.

„Schwache“ vs. „starke KI“

Bei einer Beobachtung der gegenwärtigen öffentlichen Debatte drängt sich der Gedanke auf, dass Künstliche Intelligenz schon in absehbarer Zukunft den Menschen deutlich übertrumpfen könnte (Abbildung 1). Auch wenn gute Filme und Serien aus dem Bereich Science Fiction dazu scheinbar klare Bilder bieten, ist diese

universelle, sogenannte starke KI jedoch selbst in Ansätzen in unserer gegenwärtigen Welt nicht existent. Praktische Anwendungen von KI sind heute hochspezialisiert und ihre Fähigkeiten nicht verallgemeinerbar („schwache KI“). Jede neue Aufgabe benötigt ein umfassendes Training mit enormen Datenmengen, das in der Regel in der Nutzungsphase andauert. Ein KI-System ist also niemals wirklich fertiggestellt.

Begriffsklärung Künstliche Intelligenz

Zusammenfassend gilt Künstliche Intelligenz als eine begriffliche Beschreibung für diverse Computersysteme, die Aufgaben nicht streng nach einprogrammierten Algorithmen, also Regeln und Rechenvorgängen, abarbeiten, sondern die mit großen Datenmengen trainiert wurden und nun „Entscheidungen“ treffen, die sie aus ihren bereits bekannten, smarten Daten ableiten. Der Begriff Künstliche Intelligenz kann dabei als Oberbegriff gesehen werden (Abbildung 2), der mehrere Ansätze umfasst wie z. B. Maschinelles Lernen (Machine Learning), Neuronale Netze oder mehrschichtiges Lernen (Deep Learning). Weitere Begriffe und Erläuterungen, die unter den Oberbegriff KI fallen, finden sich z. B. im Periodensystem Künstliche Intelligenz (siehe den Link unter „Literatur“).



Abbildung 1: „Schwache KI“ und „starke KI“
Quelle: eigene Darstellung VDI TZ

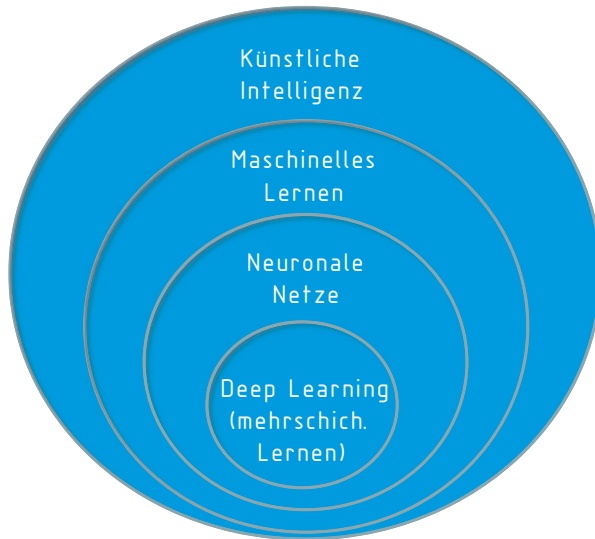


Abbildung 2: KI – Ein Oberbegriff, der mehrere Ansätze umfasst
Quelle: eigene Darstellung VDI TZ auf der Basis von Capgemini

Voraussetzung für KI: Steigende Datenmengen

Das Vorhandensein smarter Daten ist eine wichtige Voraussetzung für Künstliche Intelligenz. Durch das Erschließen von sehr großen Datenmengen, die als „Big Data“ bezeichnet werden, sollen in der Regel Geschäftsabläufe und Entscheidungsprozesse in Unternehmen und öffentlichen Organisationen optimiert oder automatisiert werden. Grundlage für die massenhafte Generierung von Daten sind unter anderem vernetzte Sensoren in der Industrie, das Internet der Dinge, unstrukturierte Daten aus dem World Wide Web oder die Abwicklung digitaler Geschäftsprozesse. In der chemischen Industrie werden zudem Anlagenzustände statt wie früher im Viertelstundentakt inzwischen sekundlich gemessen. Entscheidend ist damit heute, dass diese Daten nicht nur für die jeweiligen direkten Prozesse genutzt werden, sondern auch durch Kombination unterschiedlicher Datenquellen neue Erkenntnisse gewonnen werden. Vereinfacht wird das sogenannte Veredeln von Daten als „Smart Data“ bezeichnet. In unstrukturierten Daten können mittels der inzwischen verfügbaren enormen Rechenleistung durch statistische Analysen verborgene Muster erkannt („Mustererkennung“) und scheinbar kausale Beziehungen hergestellt werden, die Antworten außerhalb des Erwarteten geben können.

KI in der chemischen Industrie

Die chemische Industrie in Deutschland zeichnet sich durch eine hohe technologische Affinität aus. Laut einer Branchenanalyse zur chemischen Industrie liegt die Qualifikation der Beschäftigten auf einem überdurchschnittlich hohen Niveau: 2019 lag der Anteil der

Spezialist*innen und Expert*innen in der chemischen Industrie bei 33 Prozent, während er im gesamten verarbeitenden Gewerbe bei 26,7 Prozent lag.

Eine Befragung von Führungskräften, durchgeführt von Tata Consultancy Services und Bitkom Research, kommt zu dem Ergebnis, dass 43 Prozent der Unternehmen der chemischen Industrie die Einführung von KI-getriebenen Prozessen erwägen. Während auf der betrieblichen Ebene das Engagement zusehends wächst, KI-Anwendungsmöglichkeiten in den eigenen Prozessen zu identifizieren und einzuleiten, werden zugleich auch Vorbehalte geäußert. Durch voranschreitende Automatisierung könnten bestimmte Beschäftigungsprofile zukünftig nicht mehr passen. Betroffene Beschäftigte könnten ihren Arbeitsplatz verlieren. Vorbehalte bezüglich möglicher Intransparenz und Unsicherheiten können über eine frühzeitige Diskussion zu einer humanzentrierten Künstlichen Intelligenz (hKI) angegangen werden.

Leitbild Humanzentrierung

Eine humanzentrierte KI ist eine ergänzende Intelligenz, die die Beschäftigten nicht ersetzt. Es geht demnach um eine Erweiterung der menschlichen Fähigkeiten: die Kombination aus Mensch und KI erfüllt Aufgaben besser als einer der beiden allein. Das Leitbild einer humanzentrierten bzw. menschengerechten KI findet sich bereits in zahlreichen Schriften auf politischer Ebene: u. a. in der Strategie Künstliche Intelligenz der Bundesregierung, im Abschlussbericht der Enquete-Kommission Künstliche Intelligenz des Deutschen Bundestages und im gegenwärtigen Entwurf einer europäischen KI-Regulierung der Europäischen Kommission (AI-Act) (vgl. Abbildung 3).

Wie dieses Leitbild auf betrieblicher Ebene umgesetzt werden kann, zeigt der vorliegende Leitfaden auf.

An wen richtet sich der Leitfaden?

Der vorliegende Handlungsleitfaden richtet sich gleichermaßen an Technolog*innen, Human Resources, Management, Betriebsräte und Gewerkschaften. Anhand konkreter Beispiele aus der betrieblichen Praxis veranschaulicht er den Mehrwert von humanzentrierter Künstlicher Intelligenz und bietet Umsetzungshilfen an.

Wie wurde der Leitfaden erstellt?

Der Leitfaden basiert zum einen auf Ergebnissen einer umfangreichen Literaturanalyse. Zum anderen fließen Ergebnisse aus sozialpartnerschaftlichen Workshops mit Branchenvertreter*innen ein: Workshops auf betrieblicher Ebene sowie einem überbetrieblichen Workshop. Bei den Workshops wurden die Ergebnisse der Literaturrecherche offen diskutiert und gemeinsam eingeordnet.

Auf einen Blick: Welche Fragen gilt es zu klären?

Die Entwicklung und Anwendung von Künstlicher Intelligenz im Unternehmen wirft Fragen auf. Die folgenden Fragen sollten Akteure auf betrieblicher Ebene sozialpartnerschaftlich adressieren und lösen, wenn Künstliche Intelligenz im Unternehmen eingesetzt wird oder zukünftig eingesetzt werden soll:

- Wie kann ein systematisches Vorgehen aussehen? In welchem Rahmen tauschen sich die involvierten Akteure aus? Wer im Betrieb ist wie zu beteiligen?
- Wer sind die betrieblichen Ansprechpersonen bei Fragen, Problemen und Diskussionsbedarf in Bezug auf KI?
- Bleibt das Entgelt gleich oder ändert es sich, wenn KI eingesetzt wird? Ist eine Weiterqualifizierung sinnvoll?
- Schafft Künstliche Intelligenz die Arbeit für Mitarbeiter*innen im Betrieb ab?
- Können die betroffenen Mitarbeiter*innen beim Einsatz von KI noch selber Entscheidungen treffen?
- Wie können Mitarbeiter*innen befähigt werden, die zugrunde liegenden Prozesse der KI zu verstehen?
- Was bedeutet der Einsatz von KI für Leistungskontrollen? Können und dürfen mit den erhobenen Daten Leistungskontrollen durchgeführt werden? Falls ja: in welchem Umfang? Falls nein: wie sehen Garantien für Beschäftigte aus (Stichwort Betriebsvereinbarung)?
- Benachteiligt bzw. diskriminiert die KI einzelne Mitarbeiter*innen im Betrieb? Wie kann der Einsatz von KI zu Diversität und Fairness beitragen?
- Wie kann der Einsatz von KI zur Nachhaltigkeit beitragen?

Die Kriterien einer humanzentrierte KI in der chemischen Industrie

Die Kriterien einer humanzentrierte Künstliche Intelligenz lassen sich in drei übergeordneten Themen zusammenfassen:

- Interaktion von Mensch und Technik
- Technische und rechtliche Aspekte
- Orientierung an ökologischen und sozialen Werten

Wie lässt sich die Interaktion von Mensch und Technik aktiv gestalten?



Erklärbarkeit und Transparenz von KI

Worum geht es? Einer der zentralen Aspekte, die in existierenden Kriterienkatalogen an eine humanzentrierte KI genannt

werden, ist die nötige Erklärbarkeit der mit Algorithmen automatisierten Vorgänge. Die immer komplexer werdenden Entscheidungen, die die KI potenziell ohne menschliche Intervention durchführen kann, können sonst nur mit stetig wachsendem Aufwand von den Beschäftigten nachvollzogen werden. Laut der hochrangigen KI-Expertengruppe der Europäischen Kommission wird unter Nachvollziehbarkeit die Eigenschaft verstanden, die Wirkweise verstehen und Verhalten vorausschauen zu können. Eine transparente KI bietet den Mitarbeiter*innen einen Überblick über die Schritte, die sie durchführt, und legt zudem Gründe für Entscheidungsvorschläge offen. Im Gegensatz zu einer nicht oder kaum nachvollziehbaren „Black Box“, bei der nur Input und Output bekannt sind, ist bei einer humanzentrierten KI auch der sogenannte Throughput,

| | | |
|---|---------------------------------------|---------------------------------------|
| OECD | Principles on Artificial Intelligence | 2019 |
| Enquete-Kommission KI des Deutschen Bundestages | Abschlussbericht | 2020 |
| Europäische Kommission | Entwurf einer KI-Regulierung | 2021 |
| Europäische Kommission | KI-Regulierung | voraussichtlich im Herbst/Winter 2022 |

Abbildung 3 - Zentrale Veröffentlichungen und Richtlinien, die Kriterien einer humanzentrierten KI enthalten

also die Verarbeitungsvorgänge der Daten, nachvollziehbar. Erklärbarkeit gilt als grundsätzlich für das Entstehen von dauerhaftem Vertrauen, und somit auch von Akzeptanz. Wenn Entscheidungen, die von der KI getroffen werden, nicht nachvollziehbar und vor allem erklärbar sind, kann das Verhalten der Künstlichen Intelligenz von den Mitarbeiter*innen nicht hinterfragt oder angefochten werden.

Wie kann es umgesetzt werden? Eine transparente KI hängt somit von der Art der Kommunikation ihrer Mechanismen ab – dies schließt Nutzeroberflächen, passende Visualisierungen und die Angabe des „Warums“ mit ein. Die passende Ansprache und Interaktion sind entscheidend für die Akzeptanz der Nutzer*innen, beispielsweise in konkret auf bestimmte Rollen im Unternehmen zugeschnittenen Weiterbildungen. Außerdem muss bei der Bereitstellung von Daten, Informationen oder Schlussfolgerungen deutlich werden, ob sie durch menschliche Leistung bereitgestellt oder durch eine KI entstanden bzw. gesammelt worden sind.



Entscheidung durch den Menschen

Worum geht es? Als weiterer wichtiger Aspekt wird die konsequente Kontrollbefähigung der Beschäftigten gesehen.

Demnach sollen KI-Systeme möglichst in einer Weise arbeiten, durch die Mitarbeiter*innen in ihren Handlungen unterstützt werden, selbstbestimmte Entscheidungen zu treffen. Durch den Einsatz einer solchen KI soll letztlich keine Autonomie abgebaut, sondern mithilfe der KI ein Unterstützungssystem für die Beschäftigten bereitgestellt und Entscheidungsspielräume erweitert werden.

Wie kann es umgesetzt werden? Die Beschäftigten müssen in Zusammenarbeit mit einer humanzentrierten KI weiterhin die volle Kontrolle über die teils automatisch ablaufenden Prozesse haben. Dazu zählt die permanente Option, bestimmte Abläufe zu stoppen, falls Fehlfunktionen erkannt werden. Auch sollten zu jeder Zeit die Prozesse transparent dargestellt werden, beispielsweise über eine geeignete Bedienoberfläche,

und die völlige Autorität über die verarbeiteten Daten bestehen. Ein behilfliches Werkzeug kann eine visuell aufbereitete Vorschau von Abläufen sein, wenn bestimmte Parameter bzw. Aktionen eingestellt sind.



Mitbestimmung und Akzeptanz

Worum geht es? Um die Akzeptanz der Beschäftigten bei der Einführung von zum Teil disruptiv wirkenden KI-Technologien

zu sichern, ist eine frühe Einbindung der Mitarbeiter*innen über formale Mitbestimmungsstrukturen und informelle Mitgestaltungskanäle wichtig. So können bereits in der Konzipierung neuer Anwendungen unternehmens- und arbeitsplatzspezifische Eigenheiten mitgedacht werden, die in Pauschallösungen Gefahr laufen unterzugehen. Die Beteiligung der Beschäftigten bei der Einführung neuer Technologien in den Arbeitsalltag hängt stark mit der spezifischen Unternehmens- und Mitbestimmungskultur zusammen.

Wie kann es umgesetzt werden? Kontinuierliche Feedbackschleifen können während der gesamten Laufzeit des KI-Systems eingesetzt werden. So lässt es sich vermeiden, nicht nur eine anfängliche Vorabprüfung des KI-Systems anzustellen, sondern den Dialog zwischen Arbeitnehmer- und Arbeitgeberseite bzgl. Gestaltung und Optimierung als fortwährenden Prozess zu verstehen (Deutscher Bundestag 2020).

Wichtig für die Möglichkeit der Gestaltungsteilhabe und Akzeptanz ist die Kommunikation mit den Beschäftigten, Endnutzer*innen und Kund*innen. Somit zeigt sich, dass eine von Anfang an mitgedachte Erklärbarkeit eine Grundvoraussetzung von Mitbestimmung und Akzeptanz sein kann, wie auch die offene Diskussion rund um Themen wie die Verwendung personenbezogener Daten, Arbeitsplatzsicherheit und Diskriminierungsprävention.

Mitbestimmungsstrukturen hängen darüber hinaus auch mit regulatorischen Rahmenbedingungen und bestehenden Vernetzungsstrukturen auf nationaler, europäischer und internationaler Ebene zusammen.

Infobox - Mensch und Technik

Einschätzungen von Praktiker*innen

- **Erklärbarkeit** und **Transparenz** der KI sind Voraussetzung für Vertrauen und Akzeptanz durch die Beschäftigten. Erklärbarkeit und Transparenz müssen aber auch technisch umgesetzt werden.
- Damit Entscheidung durch die Beschäftigten gelingt, sind enge **Interaktionsschleifen** und **stetige Kontrollbefähigung** wichtige Voraussetzungen. Sie sollten bereits in der Planungsphase berücksichtigt werden.
- KI-Anwendungen sollten **die Mitarbeiter*innen unterstützen** und nicht ersetzen. Hilfreich sind hierbei frühzeitige **Betriebsvereinbarungen** und die Identifizierung von zukünftig benötigten Qualifikationen („**Future Skills**“).
- **Mitbestimmung** der Beschäftigten und das **Einbringen von Gestaltungsvorschlägen** haben eine hohe Bedeutung. **Offen ist der ideale Zeitpunkt**, zu dem Mitarbeiter*innen in den Prozess eingebunden werden sollten. Es gilt jedoch die allgemeine Regel: Je früher, desto besser. Auf diese Weise entsteht frühzeitig Vertrauen bei allen Beteiligten.

Welche technischen und rechtlichen Aspekte gilt es zu berücksichtigen?



Technische Robustheit und Risikomanagement

Worum geht es? Eine humanzentrierte KI-Infrastruktur sollte technisch robust sein und soweit möglich auch mit unbekanntem und neuen Situationen umgehen können. Gerade bei unvorhergesehenen Anwendungsfällen ist die Bedeutung der Beschäftigten hoch, denn die KI kann nur aus historischen Daten lernen und damit aus Situationen, die bereits in der Vergangenheit aufgetreten sind.

Wie kann es umgesetzt werden? Die technische Robustheit der KI wird erhöht, wenn die KI aus möglichst verschiedenen historischen Daten lernt. Dabei kann es sinnvoll sein, innerhalb des KI-Systems Ausnahmesituationen als solche zu kennzeichnen und Mitarbeiter*innen, die mit der KI zusammenarbeiten, darauf hinzuweisen, dass die KI eine Ausnahmesituation identifiziert hat. Zur Eingrenzung eines vorhandenen Gefahrenpotentials sollte ein Konzept zum Risikomanagement vorhanden sein, das auf den jeweiligen Anwendungskontext abgestimmt ist.



Datensicherheit und Privatsphäre

Worum geht es? Daten fallen an unterschiedlichen Stellen innerhalb des Prozesses an: als Input bei der Entwicklung der KI, als Throughput im laufenden Prozess sowie als Output der KI selbst. Bei allen drei Arten von Daten gilt es, Datensicherheit und Privatsphäre zu beachten. Bei technischen Daten kann die Datensicherheit zentralen Einfluss auf den Unternehmenserfolg haben (Industriespionage, wettbewerbsrelevante Informationen etc.). Bei personenbezogenen Daten sind sowohl Datensicherheit als auch Privatsphäre von großer Bedeutung. Damit verbunden ist die Frage nach Überwachung am Arbeitsplatz, auch in unbeabsichtigter Form.

Wie kann es umgesetzt werden? Bei der Umsetzung sollte ein Inventar der verwendeten Daten (inkl. Input-, Throughput- und Output-Daten) erstellt werden, wobei

sensible Daten besonders gekennzeichnet werden sollten. Bei sensiblen Daten sollte die technische Infrastruktur auf mögliche Schwachstellen untersucht werden. Die Zweckbindung der Daten sollte sichergestellt, personenbezogene Daten vor der Weiterverarbeitung anonymisiert werden.



Verantwortungsübernahme, Haftung und Rechtsstaatlichkeit

Worum geht es? Die Verantwortlichkeit für den Einsatz eines KI-Systems und dessen Folgen muss immer bei einer (Rechts-)Person verbleiben und darf nicht an die Technik delegiert werden. Haftung versteht sich hierbei im juristischen, Verantwortungsübernahme im übergeordneten ethischen Sinn. Rechtsstaatlichkeit im vorliegenden Kontext bedeutet, dass geltendes Recht gewahrt wird – im europäischen Kontext besonders relevant ist dabei der im April 2021 vorgelegte Entwurf der EU-Kommission einer KI-Regulierung (die Verabschiedung des AI-Act wird für den Herbst/Winter 2022 erwartet). Zudem umfasst das Prinzip der Verantwortungsübernahme und Haftung das Konzept der Überprüfbarkeit im Sinne eines Audits („auditability“). Wichtige Voraussetzung für die Überprüfbarkeit ist dabei die Erklärbarkeit und Transparenz des KI-Systems.

Wie kann es umgesetzt werden? Bereits bei der Planung eines KI-Systems sollten Fragen der Verantwortungsübernahme und Haftung geklärt werden. Sinnvoll ist dabei das Hinzuziehen von unternehmensinternen oder externen Jurist*innen zur Klärung der Haftung und des anwendbaren Rechts sowie das Hinzuziehen des Managements zur Klärung der Verantwortungsübernahme. Mögliche negative Effekte eines KI-Systems sollten identifiziert, evaluiert, dokumentiert und minimalisiert werden.

Mit der Frage nach geltendem Recht stellt sich auch die Frage nach dem Geltungs- und Einsatzbereich der KI: häufig werden KI-Systeme über Ländergrenzen hinweg eingesetzt, geltendes Recht in Bezug auf den Umgang mit Daten hat aber meist nationalen Bezug. Als erster Schritt sollte deshalb der (mögliche) Geltungs- und Einsatzbereich der KI evaluiert werden. In einem zweiten

Infobox - Technische und rechtliche Aspekte

Einschätzungen von Praktiker*innen

- Das **Risikomanagement** sollte auf den jeweiligen Anwendungsfall angepasst sein. Häufig kann die KI zumindest teilweise in **bestehende Risikobewertungskonzepte** eingegliedert werden. Die endgültige Bewertung eines Risikos sollte weiterhin bei den Beschäftigten liegen.
- Den **Schutz personenbezogener Daten** und die **Zweckbindung** erhobener Daten gilt es zu achten, um Arbeitsüberwachung und Leistungskontrollen vorzubeugen. Personenbezogene Daten sollten nur dann erhoben werden, wenn es z. B. aufgrund von Arbeitsabläufen und Verantwortungsübernahme notwendig ist.
- **Haftungsaspekte** sollten bereits in der Designphase beachtet werden und die **Verantwortung stets bei den Beschäftigten** liegen.
- Digital unterstützende Hilfsmittel müssen so konzipiert sein, dass sie sich in bestehende Arbeitsprozesse eingliedern.

Schritt kann ein Inventar geltenden Rechts innerhalb des Einsatzbereichs erstellt werden. Bei einem transnationalen Einsatz sollte sich das KI-System am juristisch größten gemeinsamen Nenner orientieren, so dass geltendes Recht im gesamten Einsatzbereich eingehalten wird.

Wie gelingt eine Orientierung an ökologischen und sozialen Werten?



Diversität, Fairness und Nicht-Diskriminierung

Worum geht es? Bilden historische Daten und Echtzeitdaten eine Wirklichkeit ab, in der Diskriminierung oder systematische

Ungleichheit vorherrscht, ist es wahrscheinlich, dass die KI diese abbildet und reproduziert. Ein Beispiel hierfür ist die Tatsache, dass gängige Gesichtserkennungssoftware das Geschlecht dunkelhäutiger Frauen deutlich schlechter erkennt als das weißer Männer, was u. a. darauf zurückzuführen ist, dass in den Daten, mit denen die Gesichtserkennungssoftware trainiert wurde, weiße Männer überproportional vertreten sind.

Wie kann es umgesetzt werden? Diversität, Fairness und Nicht-Diskriminierung müssen von Beginn der Entwicklung an mitgedacht werden. Es bedarf aber gleichzeitig eines laufenden Monitorings bzw. einer laufenden Evaluation, da Verzerrungen häufig unerwartet erst im Laufe des Entwicklungsprozesses oder auch der Go-Live-Phase sichtbar werden. Diverse Design-Teams sowie ein partizipativer Entwicklungsprozess und eine partizipative Einführungsphase können helfen, Verzerrungen frühzeitig entgegenzuwirken.



Ökologische und soziale Verantwortung

Worum geht es? Der Einsatz einer humanzentrierten Künstlichen Intelligenz sollte positive ökologische und soziale

Effekte haben. Mindestvoraussetzung sollte sein, dass ökologische und soziale Faktoren mit dem Einsatz der KI nicht schlechter gestellt sind als ohne ihren Einsatz.

Ökologie und Soziales sollten in diesem Kontext bewusst weit gedacht werden, etwa im Sinne der Nachhaltigkeitsziele der Vereinten Nationen (Sustainable Development Goals, SDGs). Künstliche Intelligenz kann hierbei sowohl fördernder als auch hemmender Faktor bei der Zielerreichung sein und bedarf einer aktiven und bewussten Gestaltung und klarer Rahmenbedingungen, um vornehmlich fördernder Faktor zu sein.

Wie kann es umgesetzt werden? Ökologische und soziale Effekte sollten von Beginn der Entwicklung an mitgedacht werden, gleichzeitig bedarf es eines laufenden Monitorings bzw. einer laufenden Evaluation. Stakeholder, die auf dem Fachgebiet der ökologischen und sozialen Verantwortung spezialisiert sind, sollten hinzugezogen werden (etwa als Teil des partizipativen Entwicklungs- und Monitoringprozesses). Dies kann unternehmensinterne Stakeholder (z. B. Corporate Social Responsibility-Abteilungen, die Beschäftigten und ihre Interessenvertretungen), aber auch externe Stakeholder umfassen (Kund*innen, Berater*innen, Vertreter*innen von Nichtregierungsorganisationen).

Infobox - Nachhaltigkeit und Werte

Einschätzungen von Praktiker*innen

- **Diversität** und **Fairness** sollten in der Konzeptionsphase in jedem Fall mitgedacht werden. Bezugspunkte können dabei auch weniger offensichtlich sein – z. B. kann es bedeuten, Schriftgrößen für ältere Mitarbeitende anzupassen. **Positive Effekte** können auch durch ein **niedrigschwellig gestaltetes KI-System** entstehen, dessen Nutzung wenig vom Qualifikationsniveau abhängt.
- **Ökologisch-nachhaltige Effekte** erzielt ein KI-System zum Beispiel durch eine Verringerung notwendiger Reisen oder durch passgenaueren Einkauf (Ressourcenschonung). Digitale Prozessoptimierung und weitere KI-unterstützte Effizienzsteigerungen sollten mit **internen Nachhaltigkeitsstrategien** verknüpft werden.

Welche Handlungsoptionen bieten sich der betrieblichen Ebene?

Jedes der oben genannten Kriterien einer humanzentrierten KI erfordert spezifische Schritte für seine Umsetzung. Gleichzeitig sind die folgenden übergeordneten Schritte auf betrieblicher Ebene sinnvoll und notwendig, um eine Humanzentrierung zu ermöglichen.



Eine kooperative Unternehmenskultur etablieren

Voraussetzung für Humanzentrierung und einen erfolgreichen KI-Change-Prozess ist eine strukturierte Zusammenarbeit aller relevanter betrieblicher Akteure in einer kooperativ geprägten Unternehmenskultur. Dies fördert sowohl die Passgenauigkeit der eingeführten KI-Anwendung als auch die Transparenz der Ziele des neuen Systems und damit die Akzeptanz der Beschäftigten.



KI-Einführung als Change-Prozess verstehen

Die Einführung von KI-Systemen bringt für Unternehmen und ihre Beschäftigten Veränderungen mit sich. Deshalb sollte die Einführung von KI-Systemen als Change-Prozess verstanden werden. Change-Management bedeutet in diesem Kontext die gemeinsame menschengerechte Gestaltung des Einsatzes von Mensch-Maschine-Kollaborationen. Zentrales Kriterium innerhalb des KI-Change-Managements ist die Einbeziehung aller Betroffenen.



Kriterien und Vorteile von Humanzentrierung kennen und diskutieren

Um KI-Systeme humanzentriert auszurichten, müssen Entscheider*innen die Kriterien und Vorteile einer Humanzentrierung kennen. Der vorliegende Leitfaden möchte hier zentrale Hilfestellung sein. Sinnvoll ist es aber auch, in der Breite zu sensibilisieren und humanzentrierte Künstliche Intelligenz als Konzept in Unternehmen und Betrieben zu etablieren.



Humanzentrierung als Ziel etablieren und einfordern

In einem nächsten Schritt sollten Management und HR die Humanzentrierung von

Künstlicher Intelligenz als Ziel etablieren. Arbeitnehmervertreter*innen kommt parallel die Rolle zu, die Humanzentrierung der KI aktiv einzufordern.



Ansprechpersonen für humanzentrierte KI frühzeitig benennen

In einem systematischen KI-Change-Prozess wird eine (inner-)betriebliche Anlaufstelle für Künstliche Intelligenz notwendig und ist frühzeitig zu benennen. Die Ansprechpersonen der Anlaufstelle benötigen sowohl technische als auch soziale Kompetenzen. So können sie sowohl rein technische Fragestellungen beantworten als auch eine realistische Perspektive auf den Einsatz von KI herstellen („was kann die KI bzw. was kann sie nicht“).



Betroffene frühzeitig und dauerhaft einbinden

Um Vertrauen in KI-Systeme zu schaffen, sollten früh im Planungsprozess die wichtigsten Fragen aus der Belegschaft diskutiert werden. Dazu gehören Fragen im Kontext der Inklusion („Kann eine KI darin unterstützen, z. B. ältere Beschäftigte stärker einzubinden oder zu befähigen?“) und möglicherweise bestehende Befürchtungen in Bezug auf mögliche Leistungskontrollen. Es ist von großer Bedeutung, die konkret geplanten Anwendungsfälle offen und transparent zu kommunizieren, um ein realitätsnahes Bild der jeweiligen Einsatzbereiche zu schaffen.



Ressourcen bereitstellen und Prozesse etablieren

Für eine effiziente KI-Einführung sind idealerweise bereits im Vorfeld klare Prozesse zu planen und Ressourcen für die Umsetzung zur Verfügung zu stellen.

Voraussetzung ist dabei eine Diskussion über die Bereitstellung der notwendigen Ressourcen. Humanzentrierung kann ein Mehr an notwendigen Ressourcen, an einbeziehenden Stakeholdern und zu beachtender Prinzipien bedeuten. Innerhalb dieses Prozesses ist es deshalb von großer Wichtigkeit, die Vorteile einer Humanzentrierung für die Beschäftigten, die technische Infrastruktur, das

Management sowie das Unternehmen und die Gesellschaft als Ganzes herauszustellen. Ressourcen, die bei der Entwicklung und Bereitstellung einer humanzentrierten Künstlichen Intelligenz zunächst zusätzlich notwendig sind, können an anderen Stellen eingespart oder durch positive Effekte aufgewogen werden. Auf diese Weise können durch eine humanzentrierte KI zum Beispiel Fehler, oder eine mögliche negative Berichterstattung vermieden und die Zufriedenheit der Mitarbeiter*innen bzw. der Unternehmenserfolg erhöht werden.



Evaluation und Qualitätsmanagement

Auch ein KI-System, das bereits „läuft“, sollte mit Blick auf die Humanzentrierung weiterhin im Sinne eines Change-Management

begleitet werden. Das beinhaltet auch, dass die Wirksamkeit überprüft und auftretende Verbesserungspotentiale angegangen werden. Hierzu kann möglicherweise auf bestehende Strukturen aufgebaut werden: Je nach Unternehmensstruktur kann der IT-Bereich diese Aufgaben übernehmen und die langfristige Nutzung sicherstellen. Auch kann ein agiles oder konventionelles Qualitätsmanagement für einen kontinuierlichen Verbesserungsprozess eingebunden werden und die Erfolgskontrolle übernehmen. Dies stellt das ständige Feedback durch die Mitarbeiter*innen sicher und kann geeignete Schritte für Anpassungen sowohl auf software-, hardware- oder betrieblicher Ebene einleiten. Es wird empfohlen, Kennzahlen und Ziele in den Feedbackgesprächen (z. B. Audits) zu dokumentieren und neu zu eruieren. Kennzahlen und Ziele könnten zum Beispiel sein: Genauigkeit der KI, Nutzbarkeit, Interesse der Beschäftigten, Akzeptanz der Mitarbeiter*innen, wirtschaftlicher Mehrwert etc.

Der Blick nach vorn

Der vorliegende Leitfaden stellt Kriterien einer humanzentrierten Künstlichen Intelligenz zusammen und zeigt auf, wie Humanzentrierung im Betrieb umgesetzt werden kann. Der Leitfaden möchte damit sowohl Management als auch Arbeitnehmer*innen und ihren betrieblichen Vertretungen wertvolle Orientierung bieten.

Humanzentrierte Künstliche Intelligenz kann in Unternehmen der chemischen Industrie Prozesse effizienter und effektiver gestalten. Humanzentrierte Künstliche Intelligenz kann zudem Tätigkeiten von Mitarbeiter*innen erleichtern und verbessern. Um eine zu dem jeweiligen Betrieb und dem jeweiligen Unternehmen passende KI-Lösung zu entwickeln, ist eine transparente Diskussion und Umsetzung der Kriterien einer humanzentrierten KI sinnvoll.

Begleitend zur zukünftigen Nutzung von humanzentrierter künstlicher Intelligenz und zum Wandel

bestehender bzw. der Entwicklung gänzlich neuer Abläufe bedarf es eines angepassten KI-Change-Managements. Dies ist notwendig, um in Phasen der Veränderung Unsicherheiten abzubauen und eine kooperative Herangehensweise gezielt zu stärken. Ein frühzeitig kommunizierter und abgestimmter Prozess für Veränderungsmanagement bietet den Betrieben der chemischen Industrie und ihren Beschäftigten den Mehrwert, unterschiedliche Sichtweisen zu Zielen und Methoden vor Ort zusammenzubringen und mögliche Konflikte früh zu lösen.

Mit einer auf den Menschen ausgerichteten Künstlichen Intelligenz können sowohl die KI als auch der Mensch ihre jeweiligen Stärken in den Betrieben der chemischen Industrie deutlich besser ausspielen. So kann das künftig noch enger verzahnte Zusammenspiel menschlicher und maschineller Intelligenz gemeinsam und frühzeitig zum Nutzen aller im Betrieb gestaltet werden.

Weiterführende Informationen

Links zu nützlichen Plattformen

OECD-Observatorium Künstliche Intelligenz: Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung betreibt eine Plattform in englischer Sprache, die einen großen Überblick zu KI-relevanten Daten liefert, z. B. KI in der Forschung, Investitionen in KI, KI-Qualifizierung und Arbeitsmarkt. Die Daten sind ansprechend visuell aufbereitet.

<https://oecd.ai/en/trends-and-data>

Plattform Lernende Systeme: Eine große Plattform zum Thema KI in deutscher Sprache u. a. mit Veranstaltungstipps und Best-Practice-Beispielen für den deutschsprachigen Raum, gefördert vom Bundesministerium für Bildung und Forschung (BMBF)

<https://www.plattform-lernende-systeme.de/>

Das **KI Observatorium Künstliche Intelligenz** in Arbeit und Gesellschaft ist die Info-Plattform der Denkfabrik des Bundesministeriums für Arbeit und Soziales (BMAS).

<https://www.ki-observatorium.de/>

Das **Periodensystem der Künstlichen Intelligenz** von Bitkom e.V. bietet einen Überblick über die unterschiedlichen Arten von KI. Es will dabei helfen, systematisch über Einsatzzwecke, Chancen und Risiken von KI zu reflektieren.

<https://periodensystem-ki.de/>

Im Rahmen der Nachhaltigkeitsinitiative der deutschen Chemie: unter dem Dach von **Chemie³** setzen sich BAVC, IGBCE und VCI dafür ein, Nachhaltigkeit in der chemisch-pharmazeutischen Industrie als Leitbild zu verankern.

<https://www.chemiehoch3.de/>

Literatur

Buolamwini, Joy; Gebru, Timnit (2018): Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: Conference on Fairness, Accountability and Transparency, S. 77–91. Online verfügbar unter <http://proceedings.mlr.press/v81/buolamwini18a.html>, zuletzt geprüft am 23.09.2022.

CFI (2021): Data Anonymization. Online verfügbar unter <https://corporatefinanceinstitute.com/resources/knowledge/other/data-anonymization>, zuletzt geprüft am 23.09.2022.

Del Ponce Castillo, Aída (2021): The AI Regulation: entering an AI regulatory winter? – Why an ad hoc directive on AI in employment is required. European Trade Union Institute (ETUI). Online verfügbar unter https://www.etui.org/sites/default/files/2021-06/The%20AI%20Regulation.%20Entering%20an%20AI%20regulatory%20winter_2021.pdf, zuletzt geprüft am 23.09.2022.

Deutscher Bundestag (2020): Unterrichtung der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale*. Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale. Berlin (Drucksache 19/23700). Online verfügbar unter <https://dip21.bundestag.de/>

dip21/btd/19/237/1923700.pdf, zuletzt geprüft am 23.09.2022.

Europäische Kommission (2019a): Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Building Trust in Human Centric Artificial Intelligence (COM(2019)168). Online verfügbar unter <https://digital-strategy.ec.europa.eu/en/library/communication-building-trust-human-centric-artificial-intelligence>, zuletzt geprüft am 23.09.2022.

Europäische Kommission (2019b): Ethik-Leitlinien für eine vertrauenswürdige KI. Online verfügbar unter <https://op.europa.eu/o/portal-service/download-handler?identifier=d3988569-0434-11ea-8c1f-01aa-75ed71a1&format=pdf&language=de&productionSystem=cellar&part=>, zuletzt geprüft am 23.09.2022.

Europäische Kommission (2021): Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union (COM/2021/206 final). Online verfügbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>, zuletzt geprüft am 23.09.2022.

- Hartmann, Leonie; Klook, Rebecca; Stöcker, Luisa (2021): These: „Das Change Management im Unternehmen steht im Wandel von Mensch-Mensch zu Mensch-Maschine-Kollaboration vor vielfältigen Herausforderungen.“. In: Corporate Communications Journal 2, 2021, S. 41–46. Online verfügbar unter https://opus.hs-osnabrueeck.de/files/2523/Hartmann_Klook_Stoecker_CCJ_2021.pdf, zuletzt geprüft am 23.09.2022.
- Horn, Nikolai (2021): Vertrauen, das Handlungsoptionen schafft, ohne naiv zu sein. In: Corporate Digital Responsibility, 06.12.2021. Online verfügbar unter <https://corporate-digital-responsibility.de/article/philosophischer-zwischenruf-vertrauen/>, zuletzt geprüft am 23.09.2022.
- Huchler, Norbert; Adolph, Lars; André, Elisabeth; Bauer, Wilhelm; Bender, Nadine; Müller, Nadine et al. (2020): Kriterien für die Mensch-Maschine-Interaktion bei KI. Ansätze für die menschengerechte Gestaltung in der Arbeitswelt. Hg. v. Plattform Lernende Systeme. Online verfügbar unter https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2_Whitepaper2_220620.pdf, zuletzt geprüft am 23.09.2022.
- IG BCE; BAVC (2021): Praxis-Leitfaden: WAI40-Dialog im Betrieb umsetzen. Online verfügbar unter https://work-industry40.de/fileadmin/docs/WAI40_Journey_Online-Version.pdf, zuletzt geprüft am 23.09.2022.
- Margetis, George; Ntoa, Stavroula; Antona, Margherita; Stephanidis, Constantine (2021): HUMAN-CENTERED DESIGN OF ARTIFICIAL INTELLIGENCE. In: Gavriel Salvendy (Hg.): Handbook of Human Factors and Ergonomics. Unter Mitarbeit von Waldemar Karwowski. 5th ed. Newark: John Wiley & Sons Incorporated, S. 1085–1106.
- Nisser, Annerose; Malanowski, Norbert (2019): Branchenanalyse chemische und pharmazeutische Industrie: Zukünftige Entwicklungen im Zuge Künstlicher Intelligenz (Working Paper Forschungsförderung, 166). Online verfügbar unter https://www.boeckler.de/pdf/p_fofoe_WP_166_2019.pdf, zuletzt geprüft am 23.09.2022.
- OECD (2019): OECD Principles on AI. Online verfügbar unter <https://www.oecd.org/going-digital/ai/principles/>, zuletzt geprüft am 23.09.2022.
- Palomares, Iván; Martínez-Cámara, Eugenio; Montes, Rosana; García-Moral, Pablo; Chiachio, Manuel; Chiachio, Juan et al. (2021): A panoramic view and swot analysis of artificial intelligence for achieving the sustainable development goals by 2030: progress and prospects. In: Appl Intell 51 (9), S. 6497–6527. DOI: 10.1007/s10489-021-02264-y.
- Riedl, Mark O. (2019): Human-centered artificial intelligence and machine learning. In: Hum Behav & Emerg Tech 1 (1), S. 33–36. DOI: 10.1002/hbe2.117.
- Shahroudnejad, Atefeh (2021): A Survey on Understanding, Visualizations, and Explanation of Deep Neural Networks. arXiv preprint. Online verfügbar unter <https://arxiv.org/pdf/2102.01792>, zuletzt geprüft am 23.09.2022.
- Shneiderman, Ben (2020): Bridging the Gap Between Ethics and Practice. In: ACM Trans. Interact. Intell. Syst. 10 (4), S. 1–31. DOI: 10.1145/3419764.
- Stowasser, Sascha; Suchy, Oliver; Huchler, Norbert; Müller, Nadine; Peissner, Matthias; Stich, Andrea et al. (2020): Einführung von KI-Systemen in Unternehmen: Gestaltungsansätze für das Change-Management. Plattform Lernende Systeme. Online verfügbar unter https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2_Whitepaper_Change_Management.pdf, Zuletzt geprüft am 23.09.2022.
- Tata Consultancy Services; Bitkom Research (2020): Deutschland lernt KI. Wie Unternehmen digitale Technologien einsetzen. Online verfügbar unter https://www.tcs.com/content/dam/tcs-germany/pdf/TrendstudieDigitalisierung/2020_Studie_Deutschlandlernt-KI_DE.pdf, zuletzt geprüft am 23.09.2022.
- Vinuesa, Ricardo; Azizpour, Hossein; Leite, Iolanda; Balaam, Madeline; Dignum, Virginia; Domisch, Sami et al. (2020): The role of artificial intelligence in achieving the Sustainable Development Goals. In: Nature communications 11 (1), S. 233. DOI: 10.1038/s41467-019-14108-y.
- Vogt, Roland (2021): Digitale Souveränität und Künstliche Intelligenz für den Menschen. In: Ernst Andreas Hartmann (Hg.): Digitalisierung souverän gestalten. Innovative Impulse im Maschinenbau. Berlin, Heidelberg: Springer Vieweg, S. 143–150.
- Waeﬂer, Toni; Schmid, Ute (2021): Explainability is not Enough: Requirements for Human-AI-Partnership in Complex Socio-Technical Systems. Online verfügbar unter <https://bit.ly/3rrv5kq>, zuletzt geprüft am 23.09.2022.

Ihre Ansprechpersonen

Dr. Norbert Malanowski
Dr. Annerose Nisser
Michael Auer
Simon Beesch
E-Mail: malanowski@vdi.de

VDI Technologiezentrum GmbH
VDI-Platz 1, 40468 Düsseldorf

www.vditz.de
[@technikzukunft](https://twitter.com/technikzukunft) · [in](https://www.linkedin.com/company/vdi-technologiezentrum-gmbh)