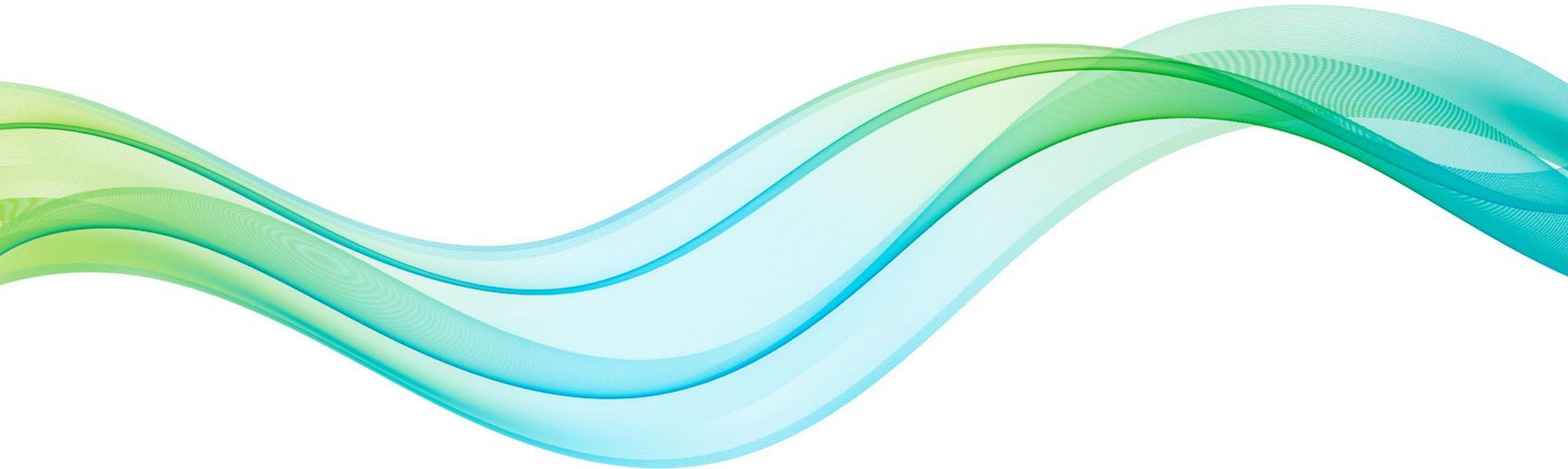


Glossar Künstliche Intelligenz

für die interdisziplinär vernetzte Arbeitsforschung

Dennis Richter, Christian Bernhard-Skala und Steffen Kinkel (Hrsg.)



Glossar Künstliche Intelligenz für die interdisziplinär vernetzte Arbeitsforschung

Ein gemeinsames Produkt Regionaler Kompetenzzentren der Arbeitsforschung und des Projekts CoCo – Connect & Collect

Herausgegeben von Dennis Richter, Christian Bernhard-Skala und Steffen Kinkel unter Mitarbeit Regionaler Kompetenzzentren der Arbeitsforschung und des Begleitforschungsprojekts CoCo – Connect & Collect.

Bei den BMBF-geförderten Regionalen Kompetenzzentren der Arbeitsforschung handelt es sich um Forschungsverbünde, die Innovationen für die Arbeitswelt im digitalen und nachhaltigen Wandel entwickeln und gestalten. Dies geschieht nicht allein aus der Wissenschaft heraus, sondern in einem Verbund aus Wissenschaft, Wirtschaft, Sozial- und Transferpartnern entlang unterschiedlicher Schwerpunktthemen wie etwa Künstliche Intelligenz, Gesundheit, Führung und Kreislaufwirtschaft.

Am vorliegenden Glossar wirkten mit seitens der Kompetenzzentren



➔ akzente40.de



➔ kompetenzzentrum-karl.de



➔ kmi-leipzig.de



➔ kompaki.de



➔ pal.webspace.tu-dresden.de



➔ wirksam.nrw

sowie darüber hinaus das Begleitforschungsprojekt CoCo- Connect & Collect



➔ coco-projekt.de

Lizenzierung:

Der Inhalt dieses Werkes steht unter einer Creative-Commons-Lizenz (Lizenztyp: Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International).



Mehr Informationen unter

➔ <https://creativecommons.org/licenses/by-sa/4.0/deed.de>

doi: 10.5281/zenodo.10779298

Zitationsvorschlag:

Richter D, Bernhard-Skala C, Kinkel S (Hrsg.) Glossar Künstliche Intelligenz für die interdisziplinär vernetzte Arbeitsforschung. Bonn und Karlsruhe. doi: 10.5281/zenodo.10779298

Schumann S, Kopp T (2024) Menschzentrierte KI. In: Richter D, Bernhard-Skala C, Kinkel S (Hrsg.) Glossar Künstliche Intelligenz für die interdisziplinär vernetzte Arbeitsforschung, Bonn und Karlsruhe, S 48-49. doi: 10.5281/zenodo.10779298



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Vorwort: Ein weiteres KI-Glossar?

Liebe Leserinnen und Leser,

Glossare ähneln Fachwörterbüchern. Sie ermöglichen das Nachschlagen von Begriffen zu einem bestimmten Thema. Ein Glossar im Kontext von Künstlicher Intelligenz (KI) soll demnach dazu beitragen, ein grundlegendes Verständnis über KI-bezogene Technologien und themenverwandte Begriffe in konkreten Anwendungsfeldern zu erlangen.

Es existieren bereits einige Glossare, die das KI-Vokabular für die interessierte Öffentlichkeit, für die Politik oder für Studierende ausdifferenzieren und in den Kontext einzelner wissenschaftlicher Disziplinen setzen. Alle diese Glossare haben einen großen Wert für die Verständigung über zentrale Begriffe zwischen Menschen unterschiedlicher Disziplinen und Anwendungskontexte, doch in der täglichen Arbeit der Regionalen Kompetenzzentren der Arbeitsforschung stießen wir damit an Grenzen. Die BMBF-geförderten Regionalen Kompetenzzentren der Arbeitsforschung sind Forschungsverbünde, die Innovationen für die Arbeitswelt im digitalen und nachhaltigen Wandel in einem Verbund aus Wissenschaft, Wirtschaft, Sozial- und Transferpartnern entwickeln und gestalten – nicht nur im Bereich der KI, allerdings mit einem starken Fokus darauf. Für einen erfolgreichen Diskurs zwischen den Regionalen Kompetenzzentren der Arbeitsforschung bedarf es daher eines gemeinsamen Vokabulars, das sowohl den Forscher:innen, den Praktiker:innen sowie den Sozial- und Transferpartner:innen gerecht wird.

Das vorliegende KI-Glossar ist ein gemeinsames Produkt aus der Mitte der Kompetenzzentren. Es wurde von zahlreichen Akteuren in den regionalen Kompetenzzentren der Arbeitsforschung und dem Begleitforschungsprojekt CoCo – Connect & Collect erstellt. Es verfolgt den Anspruch, technische Begriffe aus der Welt der Künstlichen

Intelligenz mit den Dimensionen der Arbeitsforschung Mensch, Organisation und Technik in Verbindung zu bringen und schafft so eine Verständigungsgrundlage für die notwendige Kommunikation zwischen Wissenschaftler:innen, Praktiker:innen, Sozial- und Transferpartner:innen sowie Politiker:innen bei der Entwicklung und Implementierung von KI in Betrieben.

Das Herausgeberteam hat unter der Leitung des regionalen Kompetenzzentrums KARL einen Verständigungsprozess über Begriffe der KI-basierten Arbeitswelt und deren Relevanz initiiert. Gemeinsam wurden Begriffe vorgeschlagen, priorisiert und an Autorinnen und Autoren unterschiedlicher disziplinärer Herkunft vergeben sowie in einem internen Double-Blind-Peer-Review begutachtet und überarbeitet.

Das Ergebnis können Sie nun in der Hoffnung lesen, dass es Ihnen zukünftig in Ihrer Arbeit mit KI weiterhilft. Sollten Sie es in Ihrem Projekt bei der Einführung oder Entwicklung von KI erfolgreich verwenden – an welcher Stelle auch immer – so war unser Projekt ein Erfolg – auch über den Kontext der Regionalen Kompetenzzentren der Arbeitsforschung hinaus.

Ein großer Dank der Herausgebenden geht an alle beteiligten Autorinnen und Autoren sowie Reviewenden.

Die Herausgebenden
Dennis Richter, Dr. Christian Bernhard-Skala, Prof. Dr. Steffen Kinkel

Inhaltsverzeichnis

Vorwort: Ein weiteres KI-Glossar?	3	KI-Methoden	35
Adaptives Lernsystem	5	KI-System	37
Algorithmic Management	6	Komplexitätskompetenz	38
ALTAI-Kriterien	7	Künstliche Intelligenz (KI)	39
Angewandte Ethik	8	Lernförderliche Arbeitsgestaltung	41
Arbeitsforschung	9	Maschinelles Lernen	42
Arbeitsgestaltung	11	Menschenbild	43
Arbeitsorganisation	12	Menschengerechte Arbeitsgestaltung	44
Arbeitswissenschaft	14	Mensch-Maschine-Interaktion	45
Assistenzsystem & KI-basiertes Werker-Assistenzsystem	15	Mensch-Roboter Interaktion	47
Assistenztechnik	16	Menschenzentrierte KI	48
Autonomie	17	Mitbestimmung/ Partizipation	50
Deep Learning	18	Moderierter Spezifikationsdialog	51
Delegationstechnik	19	Nachhaltigkeit (in der KI)	52
Der mitbestimmte Algorithmus	20	Nachvollziehbare KI	53
Emanzipation	21	Psychische Belastung und Beanspruchung	55
Erklärbare KI (XAI)/Erklärbarkeit	22	Reliabilität (reliability)	57
Expertensystem	24	Selbstbestimmung	58
Human In Command (Control)	25	Sprachmodelle (wie etwa ChatGPT)	59
Human In / On / Out of the Loop	26	Sustainable Development Goals (SDGs)	60
Humanisierung/Menschzentrierung	27	Transparente KI/Transparentes KI-System	61
Industrie 4.0	29	UX Design	63
Industrie 5.0	30	Wissensbasiertes System	64
Intelligenz	31	Wissensgraphen	65
Interpretierbare KI	32	Wissensrepräsentation und Schlussfolgern	66
KI-Akzeptanz	33	Danksagung	67
KI-getriebenes Geschäftsmodell	34	Impressum	70

Adaptives Lernsystem

Definition

Adaptive Lernsysteme (ALS) sind digitale Lernplattformen, welche – zum Teil auch durch [Künstliche Intelligenz](#) unterstützt – bedarfsgerechtes, an die Merkmale und Voraussetzungen der Lernenden angepasstes Lernen ermöglichen. Anstelle eines „one-size-fits-all“-Lehr- und Lernansatzes kann ein ALS gezielt auf Bedarfe des Lernenden eingehen, wie etwa dessen Fähigkeiten und Vorkenntnisse, Lernstil, Lernziele und arbeitsplatzbezogene Qualifikationsbedarfe. In diesem Kontext ist das Stichwort learning analytics zu nennen, was die Sammlung von Daten bedeutet, mit deren Hilfe das Lernen und die Lernumgebung verbessert werden sollen (Ifenthaler und Drachslers 2020).

Auf Basis statisch oder datenbasiert erhobener Merkmale der Lernenden können auch Lerninhalte sowie deren Aufbau und Schwierigkeitsgrad, die verwendeten Lernmaterialien, wie Texte, Videos, Audio usw., sowie deren Präsentation und die Navigation im System individualisiert werden.

ALS zielen darauf ab, Lernergebnisse technologiegestützt durch Individualisierung zu verbessern (Capuano und Caballé 2020). Dieser Ansatz birgt auch große Vorteile für Lehrende, welchen es ohne diese Systeme nicht möglich ist, den Lernstand der Lernenden detailliert zu überblicken und individualisiert darauf zu reagieren (Weich et al. 2021). Dennoch werden Lehrende in der Forschung zu ALS derzeit häufig noch nicht ausreichend adressiert (Özyurt und Özyurt 2015).

Obwohl ALS fraglos Stärken aufweisen, wie etwa den Gewinn an Flexibilität (Kabudi et al. 2021), lässt sich ein eindeutiger Vorteil adaptiver Lernsysteme gegenüber anderen, nicht adaptiven Systemen und Lernsettings aktuell noch nicht eindeutig nachweisen (Graf und Kinshuk 2008; Kersebaum et al. 2015; Kerres et al. 2023).

Quellen

- Capuano N, Caballé S (2020) Adaptive learning technologies. *AI Magazine* (41):96–98
- Graf S, Kinshuk R (2008) Adaptivität in Lernplattformen unter Berücksichtigung von Lernstilen. *Zeitschrift für E-Learning: Lernkultur und Bildungstechnologie* (3):40–51
- Ifenthaler D, Drachslers H (2020) Learning analytics; Spezielle Forschungsmethoden in der Bildungstechnologie. In: Niegemann HM, Weinberger A (Hrsg) *Handbuch Bildungstechnologie. Konzeption und Einsatz digitaler Lernumgebungen*. Springer, Wiesbaden
- Kabudi T, Pappas I, Olsen DH (2021) AI-enabled adaptive learning systems; A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*:1–12
- Kerres M, Buntins K, Buchner J, Drachslers H, Zawacki-Richter O (2023) Interaktive, adaptive und künstlich-intelligente Lernprogramme: Potenziale für das digitale Lernen ausloten. In: de Witt C, Gloerfeld C, Wrede SE (Hrsg) *Künstliche Intelligenz in der Bildung*. Springer Fachmedien, Wiesbaden, S 109–131
- Kersebaum A, Gillen J, Rath T (2015) Influence of an Adaptive Learning System on Exam Performance and Learning Success. *Journal of Information Technology and Application in Education* (4):1–9
- Özyurt Ö, Özyurt H (2015) Learning style based individualized adaptive e-learning environments; Content analysis of the articles published from 2005 to 2014. *Computers in Human Behavior* (52):349–358
- Weich A, Deny P, Priedigkeit M, Troeger J (2021) Adaptive Lernsysteme zwischen Optimierung und Kritik; Eine Analyse der Medienkonstellationen bettermarks aus informatischer und medienwissenschaftlicher Perspektive. *MedienPädagogik* (44):22–51

Autorin

- Ina Schiedermaier, M.A. (Sozialwissenschaften, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Algorithmic Management

Definition

Der Begriff Algorithmic Management (AM) bezeichnet das Management von Arbeitspersonen durch Algorithmen, die Entscheidungs-, Steuerungs- und Kontrollprozesse teilweise oder vollständig automatisiert ausführen (siehe auch in Teilen ähnliche, aber auch abweichende Definitionen in Bucher et al. 2021; Duggan et al. 2020; Mateescu und Nguyen 2019).

Bisherige Anwendungsbereiche sind insbesondere die Plattformarbeit (unter anderem Crowdwork, Gigwork) und die Logistik (Wood 2021; Jarrahi et al. 2021), wobei AM zunehmend auch in anderen Branchen und Arbeitsbereichen wie dem Marketing, dem Einzelhandel, der Beratung oder der Produktion eingesetzt wird (Bucher et al. 2021; Jarrahi et al. 2021).

Wenngleich der Einsatz von AM eine höhere Produktivität, Effizienz und Objektivität von Management-Entscheidungen verspricht (Briône 2020), deuten verschiedene Studien auf negative Auswirkungen auf die Arbeitspersonen hin, wie zum Beispiel einen höheren Leistungsdruck, eine geringere [Autonomie](#) oder eine geringere Arbeitsmoral und -zufriedenheit (Kellogg et al. 2020; Wood 2021; Bucher et al. 2021; Parent-Rocheleau und Parker 2022). Bei der Entwicklung und Einführung von AM-Systemen sollten daher ethische, rechtliche und soziale Implikationen (Brandl et al. 2020) berücksichtigt werden.

Welche Potentiale der Einsatz von AM für die [menschengerechte Arbeitsgestaltung](#) in der industriellen Produktion bietet, ist jedoch noch weitgehend unerforscht und wird deshalb im Kompetenzzentrum AKzentE4.0 untersucht.

Quellen

Brandl C, Wille M, Nelles J, Rasche P, Schäfer K, Flemisch FO, Frenz M, Nitsch V, Mertens A (2020) AMICA: A Method Based on Risk Analysis to Integrate Responsible Research and Innovation into the Work of Research and Innovation Practitioners. *Sci Eng Ethics* 26 (2):667–689. doi:10.1007/s11948-019-00114-2

- Briône P (2020) My boss the algorithm. An ethical look at algorithms in the workplace. Involvement & Participation Association, London
- Bucher EL, Schou PK, Waldkirch M (2021) Pacifying the algorithm – Anticipatory compliance in the face of algorithmic management in the gig economy. *Organization* 28 (1):44–67. doi:10.1177/1350508420961531
- Duggan J, Sherman U, Carbery R, McDonnell A (2020) Algorithmic management and app work in the gig economy: A research agenda for employment relations and HRM. *Hum Resour Manag J* 30 (1):114–132. doi:10.1111/1748-8583.12258
- Jarrahi MH, Newlands G, Lee MK, Wolf CT, Kinder E, Sutherland W (2021) Algorithmic management in a work context. *Big Data Soc* 8 (2). doi:10.1177/20539517211020332
- Kellogg KC, Valentine MA, Christin A (2020) Algorithms at Work: The New Contested Terrain of Control. *Acad Manag Ann* 14 (1):366–410. doi:10.5465/annals.2018.0174
- Mateescu A, Nguyen A (2019) Explainer: Algorithmic Management in the Workplace. Data & Society, New York
- Parent-Rocheleau X, Parker SK (2022) Algorithms as work designers: How algorithmic management influences the design of jobs. *Hum Resour Manag Rev* 32 (3):100838. doi:10.1016/j.hrmr.2021.100838
- Wood AJ (2021) Algorithmic Management. Consequences for Work Organisation and Working Conditions (JRC Working Papers Series on Labour, Education and Technology, No. 2021/07). Joint Research Centre, European Commission, Seville

Autor:innen

- Florens Burgert, M.Sc. (Ingenieurwissenschaften, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Prof. Dr.-Ing. Susanne Mütze-Niewöhner (Arbeitswissenschaft, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Univ.-Prof. Dr.-Ing. Verena Nitsch (Arbeitswissenschaft, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)

ALTAI-Kriterien

Definition

Die ALTAI-Kriterien sind Kriterien für die Gestaltung vertrauenswürdiger Künstlicher Intelligenz (KI). Sie wurden von der durch die Europäische Kommission gegründete High Level Expert Group on Artificial Intelligence entwickelt. Die Buchstaben ALTAI stehen für Assessment List for Trustworthy AI und behandeln sieben Kriterien. Diese sieben Kriterien sind: menschliches Handeln und [Autonomie](#); Sicherheit und Widerstandsfähigkeit; Datenschutz und Datenverwaltung; [Transparenz](#); Diversität, Nichtdiskriminierung und Fairness; Ökologisches und gesellschaftliches Wohlergehen; Rechenschaftspflicht.

Das Kriterium menschliches Handeln und Autonomie bezieht sich auf die menschliche Autonomie, welche durch die KI nicht eingeschränkt werden sollte. Sicherheit und Widerstandsfähigkeit zielen hingegen auf das [KI-System](#) ab. Sicherheit meint, dass die KI dazu in der Lage ist, Dienste zu erbringen, denen der Mensch berechtigterweise vertrauen kann. Widerstandsfähigkeit meint die technische Robustheit des Systems in Bezug auf Veränderungen. Durch die datenbasierte Arbeitsweise Künstlicher Intelligenz müssen Datenschutz und Datenverwaltung bei der Entwicklung vertrauenswürdiger KI eine große Bedeutung zugeschrieben werden. Das Kriterium Transparenz kann nochmals in drei Dimensionen aufgeteilt werden: [Nachvollziehbarkeit](#) (wie etwa durch die Dokumentation der Entwicklung des Systems), Erklärbarkeit (sowohl bezogen auf technische Prozesse als auch auf Hintergründe der Entscheidungsfindung) und offene Kommunikation über die Grenzen des KI-Systems. Diversität, Nichtdiskriminierung und Fairness beziehen sich auf die Sicherstellung der Vermeidung unbeabsichtigter (un)direkter Vorurteile gegenüber bestimmten Gruppen oder Menschen durch das KI-System. Durch die Kriterien ökologisches und gesellschaftliches Wohlergehen soll sichergestellt sein, dass die Auswirkungen von KI-Systemen auf die Umwelt und Menschen sorgfältig überwacht und berücksichtigt und Schäden

ausgeschlossen werden. Das letzte Kriterium Rechenschaftspflicht erfordert die Einrichtung von Mechanismen, die Verantwortung für die Entwicklung, den Einsatz und/oder die Nutzung von KI-Systemen gewährleisten.

Die Kriterien wurden in ein Instrument übersetzt (Assessmentlist for Trustworthy Artificial Intelligence), welches KI-Entwicklerinnen und Entwicklern dabei helfen soll, vertrauenswürdige KI zu entwickeln. Das Instrument kann durch die Vorgabe einer Reihe konkreter Schritte für die Selbstbeurteilung im Entwicklungsprozess dazu beitragen, dass die Nutzenden von KI profitieren, ohne sich vermeidbaren Risiken auszusetzen.

Weiterführende Informationen zu den ALTAI-Kriterien: <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

Weiterführende Informationen zum ALTAI-Instrument: <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>

Quelle

High-Level Expert Group on Artificial Intelligence, Stix, C. (2020) The assessment list for trustworthy artificial intelligence (ALTAI). <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>. Zugriffen: 24.04.2023

Autorin

- Lena Hintzen, M.Sc. (Psychologie, MA&T Sell & Partner GmbH, AKzentE4.0)

Angewandte Ethik

Definition

Die Angewandte Ethik ist eine Teildisziplin der praktischen Philosophie, die Theorien der normativen Ethik auf unterschiedliche Bereiche des menschlichen Handelns und Zusammenlebens anwendet. Als normative Ethik befasst sie sich beispielsweise mit der Frage, wie das Zusammenleben der Menschen idealerweise geregelt werden kann. So geht es hier nicht um Tatsachenaussagen, sondern um Sollens-Aussagen, die für weitreichende Abwägungsprozesse gesellschaftlicher Fragen entwickelt werden. Während die theoretischen Wurzeln der praktischen Philosophie bis in die griechische Antike zurückreichen, hat sich die Angewandte Ethik vor der Entwicklung risikoreicher Phänomene in der Forschung, in der Technik und in der Medizin erst in den 1970er Jahren herausgebildet. In der Folge hat die Angewandte Ethik ihrerseits unterschiedliche Bereichsethiken mit eigenen theoretischen Schwerpunkten ausgebildet wie etwa die Medizin- oder die Umweltethik. Die Bearbeitung (neuer) gesellschaftlicher Risiken wie beispielsweise das Entstehen bedrohlicher Umweltverschmutzungen oder neuer ethischer Fragestellungen – etwa ob [Künstliche Intelligenz \(KI\)](#) bei der Bewilligung von Sozialleistungen eingesetzt werden kann und soll – setzt ethisches und das entsprechende Fachwissen voraus. So haben sich die Bereiche stark ausdifferenziert. Die Methoden der Angewandten Ethik sind an die vier klassischen Theorien der praktischen Philosophie angelehnt und umfassen die Deontologie, den Konsequentialismus, die Tugendethik und die Vertragstheorie (Grunwald und Hillerbrand 2021). Allerdings beginnen hier auch die methodischen Probleme der

Angewandten Ethik, da nicht immer klar ist, wie diese abstrakten Theorien auf die konkreten Anwendungsfälle anzuwenden sind. So müssen die methodologischen Prämissen aus jedem Einzelfall entwickelt werden, damit normative Abwägungsprozesse im Hinblick auf konkrete Fragestellungen abgeleitet werden können. Die Angewandte Ethik hat eine hohe Aktualität im Feld der KI entwickelt, wo in den letzten Jahren eine Vielzahl von Guidelines und Berichte etc. entstanden sind (etwa Kirchsclaeger 2021). Auch hier jedoch gilt, dass die Bewertungen auf konkrete Einzelfälle angewendet werden.

Quellen

Grunwald A, Hillerbrand R (Hrsg.) (2021) Handbuch Technikethik. J.B. Metzler, Stuttgart
Kirchsclaeger PG (2021) Digital Transformation and Ethics. Ethical Considerations on the Robotization and Automation of Society and the Economy and the Use of Artificial Intelligence. Nomos, Baden-Baden

Autorin

- Dr. Bettina-Johanna Krings (Technik- und Arbeitssoziologie sowie Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Arbeitsforschung

Definition

Erkenntnisgegenstand der Arbeitsforschung ist die menschliche Arbeit. Die Arbeitsforschung strebt nach einem erweiterten Verständnis von Arbeit im betrieblichen und gesellschaftlichen Kontext und zugleich nach der Ableitung von menschengerechten Gestaltungs- und Interventionsmöglichkeiten. Arbeitsforschung ist eine anwendungsorientierte und empirische Wissenschaft. Sie setzt sich mit dem Verhältnis des Menschen zu seiner Arbeitstätigkeit, seinem Arbeitsplatz, seinen Arbeitsmitteln und seiner Arbeitsumgebung auseinander. Hierzu eingesetzte (und zunehmend digitalisierte) Verfahren der Arbeitsanalyse und -bewertung beziehen sich vor allem auf situative Belastungs- und Beanspruchungsgrößen des arbeitenden Menschen (Schlick et al. 2018). Die Arbeitsforschung stellt betrieblichen, sozialpartnerschaftlichen und arbeitspolitischen Akteuren systematisch ermittelte Erkenntnisse als Entscheidungsgrundlage für eine menschengerechte und produktive Gestaltung von Arbeitstätigkeiten bereit. Dadurch soll der Mensch seine Fähigkeitenpotenziale entwickeln und gleichzeitig wirksam, motiviert und gesund arbeiten.

Schwerpunktt Themen der Arbeitsforschung sind technik- und arbeitsorientierte Produktivitäts-, Innovations- und Beschäftigungsstrategien unter den Bedingungen von Arbeitsteilung und Fremdversorgung. Hierzu setzt sie sich mit Zukunftsentwicklungen auseinander (Zink 2022), wie zum Beispiel:

1. Automatisierung und [↗ Mensch-Maschine-Interaktion](#): Die fortschreitende Digitalisierung sowie die zunehmende Verbreitung von Künstlicher Intelligenz rückt die Zusammenarbeit von Menschen und Maschinen in den Mittelpunkt des Forschungsinteresses (Rothe et al. 2019).
2. Sozio-demografischer Wandel: In einer alternden Bevölkerung sind altersdifferenzierte und entwicklungsförderliche Arbeitsbedingungen zu schaffen, die sich vorteilhaft auf das Arbeitsvermögen und die Gesundheit des Menschen auswirken.
3. Flexibles Arbeiten: Digitale Technologien erweitern die Möglichkeiten für eine raumzeitliche Arbeitsflexibilisierung wie Remote-Arbeit und die sogenannte Gig-Economy. Hier gilt es, soziale Faktoren wie Wertschätzung, Fehlerkultur, Rückmeldung und Persönlichkeit zu untersuchen.

Menschliche Arbeitstätigkeiten vollziehen sich im Allgemeinen in Arbeitssystemen, die aus einem sozialen und einem technischen Teilsystem bestehen (Sydow 2002). Diese Teilsysteme sind differenziert zu analysieren und integrativ zu gestalten. Geeignete Ansätze zeigt das MTO-Konzept auf (Ulich 2011), das die Arbeitsaufgabe in den Mittelpunkt der Betrachtung rückt. Sie verknüpft einerseits das soziale mit dem technischen Teilsystem; sie verbindet andererseits den Menschen mit den organisationalen Strukturen. Die Aufgabenverteilung (unter anderem Funktionsteilung) zwischen Mensch und Technik prägt die menschlichen Erfahrungs-, Entscheidungs- und Handlungsspielräume.

Richtet sich das Interesse der Arbeitsforschung unter den Bedingungen einer routinisierten Massenfertigung vor allem auf eine schützende Wirkung präventiver Gestaltungsmaßnahmen (zum Beispiel Ergonomie, Arbeitsschutz), gewinnt in komplexen, widersprüchlichen Arbeitsverhältnissen eine entwicklungsförderliche und befähigende Dimension der menschlichen Arbeit an Bedeutung. Aus individueller Einsicht und Verpflichtung gegenüber den Kundenbedürfnissen erbrachte Leistungen prägen hier wesentlich den Arbeitsprozess und das Arbeitsergebnis – und damit Qualität und Wert von Waren und Dienstleistungen.

Quellen

- Rothe I, Wischniewski S, Tegtmeier P, Tisch A (2019) Arbeiten in der digitalen Transformation – Chancen und Risiken für die menschengerechte Arbeitsgestaltung. Zeitschrift für Arbeitswissenschaft (73):246-251. doi:10.1007/s41449-019-00162-1
- Schlick C, Bruder R, Luczak H (2018) Arbeitswissenschaft. 4. Auflage. Springer Vieweg, Berlin, Heidelberg. doi:10.1007/978-3-662-56037-2
- Sydow J (2002) Strategische Netzwerke: Evolution und Organisation. 5. Auflage. Gabler, Wiesbaden
- Ulich E (2011) Arbeitspsychologie. 7. Auflage. Schäffer-Poeschel, Stuttgart
- Zink KJ (2000) Die Zukunft der Arbeit erforschen. Ein Memorandum der Gesellschaft für Arbeitswissenschaft e. V. zum Strukturwandel der Arbeit. Dortmund

Autor

- Dr. Martin Braun (Arbeitswissenschaft, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Projekt CoCo – Collect & Connect)

Arbeitsgestaltung

Definition

Die Arbeitsgestaltung (auch Arbeitssystemgestaltung) wendet Erkenntnisse von Arbeitswissenschaft und -forschung in betriebspraktischen Anwendungen an, um die Arbeit an den Menschen oder den Menschen an die Arbeit anzupassen. Insofern leistet jede absichtsvolle Maßnahme zur Festlegung oder Veränderung der menschlichen Arbeit einen Beitrag zur Arbeitsgestaltung. Die Arbeitsgestaltung zielt darauf, günstige Bedingungen für ein aufgabengerechtes Zusammenwirken von Menschen, Betriebs- oder Arbeitsmitteln, Arbeitsmethoden und Arbeitsplätzen zu schaffen. Sie berücksichtigt hierbei die menschlichen Leistungsvoraussetzungen (einschließlich möglicher individueller Leistungseinschränkungen) sowie technologische, wirtschaftliche und rechtliche Rahmenbedingungen. Arbeitsgestaltung orientiert sich an den Kriterien der Ausführbarkeit, Erträglichkeit, Zumutbarkeit, Persönlichkeitsförderlichkeit sowie Sozialverträglichkeit. Sie zielt darauf, einseitige Belastungen auszugleichen sowie auf die menschliche Lern- und Leistungsbereitschaft positiv einzuwirken. Grundlegende Leitlinien zur Arbeitssystemgestaltung sowie relevante Begriffe sind in der DIN EN ISO 6385 definiert; eine allgemeinere Beschreibung der ergonomischen Grundsätze findet sich in DIN EN ISO 26800. Maßnahmen der Arbeitsgestaltung beziehen sich auf ergonomische Bedingungen (zum Beispiel Lärm, Beleuchtung, Klima, Möblierung) sowie auf organisatorische und inhaltliche Kriterien der Arbeitstätigkeit (in Sachen Arbeits- und Funktionsteilung, Prozessgestaltung und Informationsbeziehungen). Gestaltungsaufgaben betreffen unter anderem die Funktionsteilung und Interaktion von Mensch und Maschine im Kontext der fortschreitenden Digitalisierung der Arbeitsprozesse (einschließlich nutzerfreundlicher, intuitiver Schnittstellen) und

die Einrichtung zunehmend mobiler Arbeitsplätze; ferner die Gestaltung lern- und gesundheitsförderlicher Arbeitsformen angesichts heterogener Beschäftigungsformen sowie der Auswirkungen des sozio-demografischen Wandels. Derartige Gestaltungsmaßnahmen erhöhen im Allgemeinen die Arbeitssicherheit, -leistung oder -produktivität; indirekt beeinflussen sie die Wirtschaftlichkeit von Anwendungen. Um Merkmale menschzentrierter Gestaltung unter anderem bei informationsverarbeitenden Tätigkeiten zu erfüllen, werden präventive und partizipative Methoden angewandt.

Quellen

- DIN EN ISO 6385:2016 Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen
- DIN EN ISO 26800:2011 Ergonomie – Genereller Ansatz, Prinzipien und Konzepte
- Hacker W (2022) Arbeitsgestaltung bei Digitalisierung. Zeitschrift für Arbeitswissenschaft 76 (1):90-98
- REFA - Verband für Arbeitsstudien und Betriebsorganisation (1991)
Grundlagen der Arbeitsgestaltung. Hanser, München
- Spath D, Braun M (2021) Human Factors and Ergonomics in Digital Manufacturing.
In: Salvendy G, Karwowski W (Hrsg) Handbook of Human Factors and Ergonomics. 5. Auflage. Wiley, New York, S 1438-1459

Autor

- Dr. Martin Braun (Arbeitswissenschaft, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Projekt CoCo – Connect & Collect)

Arbeitsorganisation

Definition

Die Arbeitsorganisation (AO) ist neben der Funktionsteilung zwischen Mensch und Technik Teil der [menschgerechten Arbeitsgestaltung](#). Die AO gestaltet hauptsächlich die Verteilung oder Kombination der Aufträge und Anforderungen auf die Arbeitenden nach inhaltlichen, zeitlichen und örtlichen Merkmalen.

Funktionsteilung und AO müssen, um [menschzentrierte Arbeitsgestaltung](#) (insbesondere DIN EN ISO 6385 (2016-12)) zu verwirklichen, die ausführbare, schädigungslose, beeinträchtigungsfreie sowie lern- und gesundheitsförderliche Gestaltung der Arbeitsprozesse sichern.

Hauptaktivitäten der AO betreffen:

- das Verhältnis von Anforderungsteilung und -kombination bei Arbeitenden zum Erzeugen von vollständigen/ganzheitlichen Aufgaben hinsichtlich ihrer Teiltätigkeiten (Vorbereiten, Organisieren, Ausführen, Kontrollieren) und ihrer geistigen Anforderungen (Automatismen, Wissens-, Denkanforderungen)
- den Grad der Festlegung der Vorgehensweisen (Arbeitsmethoden) im Interesse von Entscheidungsmöglichkeiten (Handlungsspielraum mit Verantwortung) der Arbeitenden
- das Ausmaß der zeitlichen und örtlichen Bindung oder Entkoppelung von Arbeitsprozessen;
- Vorhandensein und Art der Gliederung von Arbeitsprozessen (Mengen-, Zeit-, inhaltliche Gliederung) mit ihrer Wirkung auf Zielentwicklung und Motivierung der Arbeitenden.

Die sogenannten Neuen Formen der AO entwickelten komplexe Leistungs- und beanspruchungsgünstige Arbeitsorganisationsformen insbesondere als Auftragswechsel, Auftragsserweiterung und -bereicherung sowie Selbstorganisation.

Da Arbeitsteilung oder -kombination für Individuen, Dyaden oder Gruppen möglich ist, sind den Aufträgen und Arbeitsmitteln angemessene Formen der Zusammenarbeit (Kooperation, agile Kollaboration) als Arbeit im Raumverband, als Sukzessivverband (wie etwa Fließband) oder in verschiedenen Arten von Integrativverbänden (wie etwa zeit- und ortsgleich versus getrennt) im Unterschied zur Einzelarbeit in hohem Maße leistungs- und beanspruchungsbestimmend (Grap 1992; Hacker und Sachse 2023; Heeg 1988; Schlick et al. 2018; Frost et al. 2020).

Die Gestaltungsaufgaben und -gegenstände der AO sind zum Teil, eng miteinander verbunden oder beeinflussen sich gegenseitig (siehe Schalenmodell in Mütze-Niewöhner und Nitsch 2020).

Die Ziele der AO müssen mit den strategischen Unternehmenszielen in Einklang stehen. Es kann davon ausgegangen werden, dass eine effektive, effiziente und menschengerechte AO sowohl zur ökonomischen und ökologischen (wie etwa durch Vermeidung von Ressourcenverschwendung) als auch zur sozialen Nachhaltigkeit beiträgt (wie etwa durch Maßnahmen zur Reduzierung [psychischer Belastung](#), siehe hierzu auch <https://www.gda-psyche.de/>).

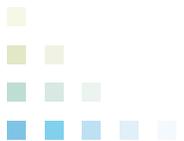
Eine menschenzentrierte Gestaltung der AO setzt eine angemessene Beteiligung der Beschäftigten und ihrer Interessensvertretungen voraus (siehe hierzu auch Mütze-Niewöhner und Nitsch 2020).

Quellen

- DIN EN ISO 6385 (2016-12) Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen, Berlin
- Frost M, Guhlemann K, Cordes A, Zittlau K, Hasselmann O (2020) Produktive, sichere und gesunde Arbeitsgestaltung mit digitalen Technologien und Künstlicher Intelligenz – Hintergrundwissen und Gestaltungsempfehlungen. Z. Arb. Wiss. 74:76–88. doi:10.1007/s41449-020-00200-3
- Grap R (1992) Neue Formen der Arbeitsorganisation für die Stahlindustrie. Verl. der Augustinus-Buchh, Aachen
- Hacker W, Sachse P (2023) Allgemeine Arbeitspsychologie; Psychische Regulation von Arbeitstätigkeiten. vdf, Zürich
- Heeg FJ (1988) Moderne Arbeitsorganisation; Grundlagen der organisatorischen Gestaltung von Arbeitssystemen bei Einsatz neuer Technologien. Hanser, München
- Mütze-Niewöhner S, Nitsch V (2020) Arbeitswelt 4.0. In: Frenz W (Hrsg) Handbuch Industrie 4.0: Recht, Technik, Gesellschaft. Springer, Berlin, Heidelberg
- Schlick C, Bruder R, Luczak H (2018) Arbeitswissenschaft. Springer Berlin Heidelberg, Berlin, Heidelberg

Autor:innen

- Prof. Dr. rer.nat.habil. Dr. phil.h.c. Winfried Hacker (Arbeits- und Organisationspsychologie, TU Dresden, Fakultät Psychologie, Arbeitsgruppe Wissen-Denken-Handeln, PAL)
- Prof. Dr.-Ing. Susanne Mütze-Niewöhner (Arbeitswissenschaft, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Dr. Ulrike Pietrzyk (Arbeits- und Organisationspsychologie, Fakultät Psychologie - Arbeitsgruppe Wissen-Denken-Handeln TU Dresden, PAL)



Arbeitswissenschaft

Definition

Die Arbeitswissenschaft ist eine ursprünglich ingenieurwissenschaftlich geprägte Fachdisziplin. Sie integriert Einzelwissenschaften unter anderem der Arbeitstechnologie, der Arbeitsmedizin, der Arbeits- und Organisationspsychologie, der Pädagogik und der Wirtschaftswissenschaft. Begrifflich ist der Übergang zur [Arbeitsforschung](#) fließend. Die Arbeitswissenschaft misst der Verbindung von Forschung und Praxis einen großen Wert bei. Frühe Ansätze der Arbeitswissenschaft werden Taylor (1911) zugeschrieben; seine produktivitätsorientierten Arbeitsstudien begründeten das Prinzip detailliert geplanter, hoch arbeitsteiliger und zentral kontrollierter Arbeitsabläufe. Dabei auftretende, einseitige Belastungen offenbarten die Notwendigkeit einer menschengerechten [Arbeitsgestaltung](#) durch ausgewogene Arbeitsverhältnisse.

Arbeitswissenschaft charakterisiert grundsätzlich die Analyse und Gestaltung von Arbeitssystemen. Sie untersucht die Anpassung der (Erwerbs-)Arbeit an den Menschen (zum Beispiel Technikgestaltung) und des Menschen an die Arbeit (zum Beispiel Qualifizierung, Motivation). Sie betrachtet die Arbeit unter Aspekten der menschlichen Zusammenarbeit und des Zusammenwirkens von Mensch, Technik und Organisation (Luczak und Volpert 1987). Im Zuge der Digitalisierung von Arbeitsprozessen gewinnt die informatorische Dimension von Arbeit als Erkenntnisobjekt an Bedeutung. Hierbei werden die Wirkungen vernetzter Informationstechnologien und -verfahren (wie [Künstliche Intelligenz](#), Cyber-Physische Systeme oder Robotersteuerungen) betrachtet und deren Wirkungen auf den Menschen untersucht.

Ziel von Arbeitswissenschaft ist die systematische Generierung von theoretischen und empirischen Erkenntnissen in Bezug auf die Analyse, Planung, Ordnung,

Gestaltung, Leistung und Durchführung menschlicher Arbeit. Sie folgt gleichermaßen den Prämissen der wirtschaftlichen und menschengerechten Prozessgestaltung. Die Erkenntnisse der Arbeitswissenschaft zielen unter anderem auf eine sichere, zumutbare, lernförderliche und sozialverträgliche Gestaltung von Arbeitssystem und Arbeitsaufgabe (Ulich 2011). Hierzu wendet sie unter anderem organisationstheoretische, technologische und psychologische Methoden an. Der Mensch wird mehrheitlich anhand objektiver Parameter beschrieben, die sich unter anderem auf seine Fähigkeiten, Fertigkeiten, Eigenschaften und Bedürfnisse beziehen (Schlick et al. 2018).

Als wissenschaftliche Disziplin organisiert sich die Arbeitswissenschaft in der Gesellschaft für Arbeitswissenschaft (GfA), das zentrale Diskursorgan ist die Zeitschrift für Arbeitswissenschaft.

Quellen

- Luczak H, Volpert W (1987) Arbeitswissenschaft: Kerndefinition – Gegenstandskatalog – Forschungsgebiete. RKW, Eschborn
- Schlick C, Bruder R, Luczak H (2018) Arbeitswissenschaft. 4. Auflage. Springer Vieweg, Wiesbaden
- Taylor FW (1911) The principles of scientific management. Harper & Brothers, London
- Ulich E (2011) Arbeitspsychologie. 7. Auflage. Schäffer-Poeschel, Stuttgart

Autor

- Dr. Martin Braun (Arbeitswissenschaft, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Projekt CoCo – Connect & Collect)

Assistenzsystem & KI-basiertes Werker-Assistenzsystem

Definition

1. Assistenzsysteme (AS) sind eine Form von Unterstützungssystemen, das heißt, technologische Systeme, die Menschen bei einer Aktivität unterstützen (Karafillidis et al. 2015). Innerhalb der Unterstützungssysteme zeichnen sich AS dadurch aus, dass sie die Ausführung einer Aktivität, die der Mensch auch grundsätzlich ohne AS durchführen könnte, erleichtern oder die damit verbundene Leistung fördern (Karafillidis et al. 2015). Die physischen und kognitiven Fähigkeiten der Nutzenden werden durch das AS lediglich erweitert, nicht ersetzt. Ein AS kann entsprechend sowohl als digitales (zum Beispiel Arbeitsanleitung auf Tablet) wie auch als physisches System (zum Beispiel Exoskelett) ausgestaltet werden (zur Unterscheidung physischer und kognitiver AS, Apt et al. 2018). AS sind Teil eines soziotechnischen Systems, beruhend auf dem Zusammenspiel zwischen Mensch, Tätigkeit und Technologie (Karafillidis et al. 2015). Bei der Gestaltung von AS müssen diese drei Schlüsselemente berücksichtigt werden, wobei ein besonderer Schwerpunkt auf der Gestaltung der Mensch-Maschine-Schnittstelle liegt.
2. KI-basierte Assistenzsysteme (KI-AS) sind AS, deren Funktionalität durch mindestens eine KI-Komponente ermöglicht wird (Anslinger et al. 2022) → [Künstliche Intelligenz \(KI\)](#) „als zentrale Technologie eingesetzt“ (Anslinger et al. 2022) oder die insbesondere „auf umfassende Datenbanken zurückgreifen“. KI kann an den einzelnen Schlüsselementen Mensch, Tätigkeit und Technologie ansetzen oder die Schnittstelle zwischen ihnen ergänzen. Als Implikationen für den Aktivitätsrahmen eines KI-AS nennen Mittelstand digital (o.J.) die Fähigkeit, „selbstständig auf Situationen zu reagieren“ und Kraus et al. (2020, S.14) beispielsweise die „Lernfähigkeit und Problemlösungs-Kompetenz für gemeinsame Ziele“.
3. (KI-basierte) Werker-Assistenzsysteme sind (KI-)AS, die im Produktionskontext zur Unterstützung von Arbeiterinnen und Arbeitern (Werkerinnen und Werkern) genutzt werden.

Quellen

- Anslinger J, Huber J, Haslgrübler M, Thaler A (Hrsg.) (2022) Verantwortungsvolle Einbindung von KI-Assistenzsystemen am Arbeitsplatz. doi:10.17605/OSF.IO/98B4H
- Apt W, Schubert M, Wischmann S (2018) Digitale Assistenzsysteme. Perspektiven und Herausforderungen. iit, Berlin
- Karafillidis A, Weidner R (2015) Grundlagen einer Theorie und Klassifikation technischer Unterstützung. In: Weidner R, Redlich T, Wulfsberg JP (Hrsg.) Technische Unterstützungssysteme. Springer, Berlin, Heidelberg, S 66-89
- Kraus M, Ludwig B, Minker W, Wagner N (2020) Assistenzsysteme. In: Görz G, Schmid U, Schneeberger J, Braun T (Hrsg.) Handbuch der Künstlichen Intelligenz. De Gruyter, Berlin, Boston, S 859-906
- Lensing K (2021) KI-basierte Assistenzsysteme für die Industrie 4.0. In: ten Hompel M, Vogel-Heuser B, Bauernhansl T (Hrsg.): Handbuch Industrie 4.0. Springer Vieweg, Berlin, Heidelberg. doi:10.1007/978-3-662-45537-1_163-1
- Mittelstand Digital: Assistenzsysteme (o. J.). <https://www.mittelstand-digital.de/MD/Navigation/DE/Themen/Mensch-Digitalisierung/Assistenzsysteme/assistenzsysteme.html>. Zugegriffen: 13. September 2023

Autor:innen

- Leon Pfenning, M.Sc. (Industrie- und mobile Robotik, Assisted Working and Automation (AWA), Hochschule Darmstadt, KompAKI)
- Sophie Sandner, M.Sc. (Assistenzsysteme in der Produktion, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW), Technische Universität Darmstadt, KompAKI)
- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)
- Mandy Wölke (Human-Computer Interaction und Innovationswissenschaften, Institut für Angewandte Informatik (InfAI), KMI)
- Dr. Andrea Altepost (Arbeitswissenschaft, Institut für Textiltechnik der RWTH Aachen University, WIRKsam)

Assistenztechnik

Definition

Der Begriff der Assistenztechnik entstammt schon der analogen, vordigitalen Zeit. Insbesondere in Fertigungs- und Produktionsverfahren wurde Technik gewünscht, die den Menschen im Arbeitsvorgang im Maschinen-, Anlagen-, Werkzeug- und Automobilbau entlastet. Das Heben, Drehen und Wenden schwerer Bauteile sollte auf assistierende Geräte und mechanisch-automatisierte Systeme übergeben werden. Mit der zunehmenden Einführung von digitalen Lösungen wurde der Begriff Assistenztechnik im Bereich des Wohnens und der Unterstützung von älteren Menschen in ihrem Zuhause bei der Gesundheitsprävention neu aufgegriffen. Im Smart Home sollte Assistenztechnik mehr Sicherheit und verbesserte Betreuung erbringen. Ein digitaler Notrufknopf und technische Lösungen bei Bewegungen wie zum Beispiel im Bad oder beim Treppensteigen gaben dem Menschen mehr Wohnqualität. In der Digitalisierung des Büro- und Verwaltungswesens trugen assistierende Software-Werkzeuge zur Erleichterung des Arbeitens bei (Weiss et al. 2003). Kennzeichen jeder Assistenztechnik ist es, dass die Entscheidung über die Nutzung und Handhabung der Technik beim handelnden Menschen liegt (Offensive Mittelstand 2018). Es gilt die Handlungsträgerschaft Mensch. Gut und einvernehmlich gestaltete Assistenztechnik fördert die Humanisierung der Arbeit (Cernavin et al. 2018).

Siehe auch → [Delegationstechnik](#)

Quellen

Cernavin O, Schröter W, Stowasser S (Hrsg.) (2018) Prävention 4.0 – Neue Perspektiven für Führung, Organisation, Sicherheit und Gesundheit im Betrieb. Springer Fachmedien, Wiesbaden
„Offensive Mittelstand – Gut für Deutschland“, Stiftung „Mittelstand – Gesellschaft – Verantwortung“ (Hrsg.) (2018) Potenzialanalyse Arbeit 4.0. Künstliche Intelligenz für die produktive und präventive Arbeitsgestaltung nutzen: Ein Selbstbewertungscheck zur Einführung der neuen 4.0-Technologien. Heidelberg
Weiss W, Busch C, Schröter W (Hrsg.) (2003) Multimedia Arbeitsplatz der Zukunft. Assistenz und Delegation mit mobilen Softwareagenten. Mit einem Vorwort des [damaligen] Bundeswirtschaftsministers Wolfgang Clement. Talheimer Verlag, Mössingen

Autor

- Welf Schröter (Arbeitsgestaltung, Forum Soziale Technikgestaltung, KARL)

Autonomie

Definition

Der Begriff Autonomie (altgr. Selbstgesetzgebung) bezeichnet [↗ selbstbestimmte](#) Handlungsentscheidungen, die von nicht unter Zwang stehenden, freien Urheberinnen und Urhebern getroffen werden (Heidbrink 2017). Freiheit beinhaltet nach Kant (1870) die selbst auferlegte Begrenzung (Gräb-Schmidt 2018). Darunter ist die Fähigkeit, sich selbst in Freiheit ein Gesetz aufzuerlegen, auch genannt Selbstgesetzgebung, zu verstehen (Thimm und Bächle 2019). Dieses Gesetz weist eine moralische Rechtfertigung auf, wenn es ohne weiteres für die gesamte Gesellschaft gelten kann (Kant 1870). Entscheidend für die Autonomie ist ebendiese Willensfreiheit im Einklang mit moralischen Vorgaben (Thimm und Bächle 2019).

Im Kontext der [↗ Arbeitswissenschaft](#) wird Autonomie als entscheidendes Kriterium für menschengerechte Arbeit verstanden (Grote 1997), wobei sie sich auf verschiedene Aspekte beziehen kann, wie etwa die Arbeitsmethode, -zeiten, -einteilung, den Arbeitsplatz und andere (Spiegelaere et al. 2016). (siehe auch [↗ Emanzipation](#)).

Relevant ist auch der Bereich der Künstlichen Intelligenz (KI). Häufig wird KI-Autonomie attestiert, da diese Systeme ohne direktes menschliches Zutun gewünschte Tätigkeiten ausführen, wie im prominenten Beispiel des autonomen Fahrens. Der Eindruck, diese Form der Selbststeuerung eines Systems sei mit menschlicher Autonomie gleichzusetzen oder von ihr zu entkoppeln, kann als zweifelhaft und zudem als nicht wünschenswert erachtet werden. Ein System, das ohne menschliche Hilfe agiert, tut dies nicht autonom, sondern basierend auf dem menschlichen Willen, der durch Programmierung in ihm angelegt ist (Rath 2019). Statt selbstgesetzgebend zu handeln „fügt [es] sich einem Modus der Steuerung“ (Gräb-Schmidt 2018) und bedeutet somit reine Funktionalität, oder auch technische Machbarkeit (ebd.).

Es lässt sich zusammenfassen, dass die Verwendung des Begriffs Autonomie im KI-Kontext unzutreffend ist, da sich dieser ursprünglich auf den Menschen fokussierte Begriff nicht ohne weiteres auf KI übertragen lässt, ohne seine wahre Bedeutung zu verfehlen.

Quellen

- Gräb-Schmidt E (2018) Zwischen System und Verantwortung; ethische Überlegungen zum Begriff der Autonomie angesichts der Entwicklung autonomer Systeme. Analysen und Argumente:2–10
- Grote G (1997) Autonomie und Kontrolle; zur Gestaltung automatisierter und risikoreicher Systeme. vdf Hochschulverlag, Zürich
- Heidbrink L (2017) Definitionen und Voraussetzungen der Verantwortung. In: Heidbrink L, Langbehn C, Loh J (Hrsg) Handbuch Verantwortung. Springer VS, Wiesbaden, S 3–33
- Kant I (1870) Grundlegung zur Metaphysik der Sitten. L. Heimann
- Rath M (2019) Zur Verantwortungsfähigkeit künstlicher „moralischer Akteure“; Problemanzeige oder Ablenkungsmanöver? In: Rath M, Krotz F, Karmasin M (Hrsg) Maschinenethik. Normative Grenzen autonomer Systeme. Springer VS, Wiesbaden, S 223–242
- Spiegelaere S de, van Gyes G, van Hootegem G (2016) Not All Autonomy is the Same. Different Dimensions of Job Autonomy and Their Relation to Work Engagement & Innovative Work Behavior. Human Factors and Ergonomics in Manufacturing & Service Industries (26):515–527
- Thimm C, Bächle TC (2019) Autonomie der Technologie und autonome Systeme als ethische Herausforderung. In: Rath M, Krotz F, Karmasin M (Hrsg) Maschinenethik. Normative Grenzen autonomer Systeme. Springer VS, Wiesbaden, S 73–87

Autorin

- Ina Schiedermaier, M.A. (Sozialwissenschaften, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Deep Learning

Definition

Deep Learning (DL) oder auch Tiefes Lernen genannt, ist ein Teilgebiet des [maschinellen Lernens](#), das auf künstlichen neuronalen Netzen (KNN) mit mehreren verborgenen Schichten basiert. Diese KNN sind definiert als Algorithmen, die dem menschlichen Gehirn nachempfunden sind und zeichnen sich durch eine Eingabe-, eine Ausgabeschicht sowie zahlreiche verborgene Schichten dazwischen aus. Zwischen den vorhandenen Neuronen dieser Schichten können verschiedene Verbindungen vorkommen. Die Neuronen der Eingabeschicht nehmen die Eingabedaten von außen in das Netz auf, während die Neuronen der Ausgabeschicht die Ausgabedaten und somit das Ergebnis ausgeben (Tetzner et al. 2021). Die Neuronen eines KNN haben eine numerische Gewichtung und enthalten Teilinformationen, die sie an Neuronen der nächsten Schicht weitergeben. Durch jede Weitergabe wird die Berechnung der Ausgabe optimiert und das KNN lernt, von den Eingabedaten auf die richtigen Ausgabedaten zu schließen. Je höher die Anzahl solcher versteckter Schichten und damit je höher die Anzahl der Verbindungen ist, desto komplexere Muster in den Daten können erkannt werden, wie bei DL der Fall ist (Terstegen und Link 2022). Durch diese KNN mit mehreren verborgenen Schichten können beim DL komplexe Aufgabenstellungen unterschiedlicher Anwendungsbereiche, die umfassende kognitive Fähigkeiten erfordern, gelöst werden, indem sowohl strukturierte als auch unstrukturierte Daten abgebildet, Merkmale und Muster automatisch identifiziert und diese erlernten Zusammenhänge anschließend in unbekanntem zu untersuchenden Daten erkannt werden (Tetzner et al. 2021; Terstegen und Link 2022).

Während DL vor allem in der Sprach-, Text-, Bild- und Videoverarbeitung neue Möglichkeiten eröffnet, beispielweise durch automatische Sprachverarbeitung, Bildanalyse, medizinischen Diagnostik oder der Prozesssteuerung (Döbel et al. 2018; Mich 2020), werden diese immer komplexer werdenden Strukturen und automatisierten Prozesse für den Menschen immer weniger nachvollziehbar und entwickeln sich immer mehr zu einer Black Box. Aus diesem Grund spielen in diesem Zusammenhang

Modelle, die einzelne Ausgaben für Mitarbeitende besser nachvollziehbar machen, und somit [erklärbare KI](#) eine immer wichtiger werdende Rolle (Döbel et al. 2018). Außerdem sind einige DL-Modelle sehr rechenintensiv und nehmen viel Trainingszeit in Anspruch. Ein Modell, das sich zur Verarbeitung sequenzieller Daten etabliert hat, ist das Transformer-Modell, das einige dieser Herausforderungen vermeidet (Chandra et al. 2023).

Quellen

- Chandra A, Tünnermann L, Löfstedt T, Gratz R (2023) Transformer-based deep learning for predicting protein properties in the life sciences. *eLife* (12). doi:10.7554/eLife.82819
- Döbel I, Leis M, Molina Vogelsang M, Neustroev D, Petzka H, Riemer A, Rüping S, Voß A, Wegele M, Welz J (2018) Maschinelles Lernen. Eine Analyse zu Kompetenzen Forschung und Anwendung. München: Fraunhofer-Gesellschaft. <https://publica.fraunhofer.de/handle/publica/299081>
- Mich L (2020) Artificial Intelligence and Machine Learning. In: Xiang Z, Fuchs M, Gretzel U, Höpken W (Hrsg.) *Handbook of e-Tourism*. Springer International Publishing, Cham, S 1–21 doi:10.1007/978-3-030-05324-6_25-1
- Terstegen S, Link J (2022) Künstliche Intelligenz. Grundlagen, Anwendungen und Praxisbeispiele. Hg. v. ifaa — Institut für angewandte Arbeitswissenschaft e.V.
- Tetzner A, Kühne T, Gluchowski P, Pfoh M (2021) Künstliche Neuronale Netze – Aufbau, Funktion und Nutzen. In: Frick D, Gadatsch A, Kaufmann J, Lankes B, Quix C, Schmidt A, Schmitz U (Hrsg.) *Data Science*. Springer Fachmedien, Wiesbaden, S 225–239. doi:10.1007/978-3-658-33403-1_14

Autorin

- Elena Kick, M.Sc. (Wirtschaftswissenschaften, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Delegationstechnik

Definition

Kennzeichen der Delegationstechnik ist die softwaretechnisch entstandene Möglichkeit, einen rechtsverbindlichen Vorgang (Transaktion) auf ein IT-System zu übertragen, zu delegieren. Dabei wechselt die Handlungsträgerschaft Mensch hinüber in eine Handlungsträgerschaft autonomer Softwaresysteme. Das bevollmächtigte algorithmische System wird ermächtigt, jenseits des Menschen verbindliche Entscheidungen zu treffen. Dies unterscheidet Delegationstechnik von [↗ Assistenztechnik](#). Der Begriff Delegationstechnik dient heute in der Debatte über die Potentiale sogenannter [↗ Künstlicher Intelligenz](#) zur Differenzierung der algorithmischen Anwendungen (Schröter 2019). Delegationstechnik zielt auf die Automatisierung des virtuellen Raumes und der Wertschöpfungsketten (Cernavin et al. 2018).

Der Begriff wurde in dem vom Bundeswirtschaftsministerium vor rund zwei Jahrzehnten geförderte Projekt Multimedia-Arbeitsplatz der Zukunft (Weiss et al. 2003) im Rahmen des Programms Mensch-Technik-Interaktion in der Wissensgesellschaft verwendet. Er bezog sich auf die damalige Entwicklung von Softwareagenten und mobilen Agentenplattformen. Der Begriff wird zunehmend Bestandteil der sozial-partnerschaftlichen Gestaltungsprozesse in Betrieben und Verwaltungen sowie der Bildung im Bauhandwerk. Er ist Teil der Präventionsforschung und neuerer industriesoziologischer Diskurse.

Siehe auch [↗ Der mitbestimmte Algorithmus](#)

Quellen

- Cernavin O, Schröter W, Stowasser S (Hrsg.) (2018) Prävention 4.0 – Neue Perspektiven für Führung, Organisation, Sicherheit und Gesundheit im Betrieb. Springer Fachmedien, Wiesbaden
- Schröter W (2019) Der mitbestimmte Algorithmus. Arbeitsweltliche Kriterien zur sozialen Gestaltung von Algorithmen und algorithmischen Entscheidungssystemen. Mössingen
- Weiss M, Busch C, Schröter W (Hrsg.) (2003) Multimedia Arbeitsplatz der Zukunft. Assistenz und Delegation mit mobilen Softwareagenten. Mit einem Vorwort des [damaligen] Bundeswirtschaftsministers Wolfgang Clement. Talheimer Verlag, Mössingen

Autor

- Welf Schröter (Arbeitsgestaltung, Forum Soziale Technikgestaltung, KARL)

Der mitbestimmte Algorithmus

Definition

Das Konzept Der mitbestimmte Algorithmus zielt auf die Gestaltung und adaptive Anpassung von [↗ Delegationstechniken](#) vor deren Beschaffung und Einführung in die Arbeitswelt (Schröter 2019; ifaa 2021). In mehrjährigem Findungsdialog zwischen Beschäftigtenvertretungen sowie im Gespräch mit Repräsentanten der Arbeitgeberseite wurde vom Forum Soziale Technikgestaltung eine mehrteilige konzeptionelle Handlungsfolge aufbereitet, die mehrere Eckpunkte umfasst. Dazu gehören:

- Begriff Nachholende Digitalisierung
- Differenzierung von [↗ Assistenztechnik](#) und [↗ Delegationstechnik](#)
- Begriff Algorithmisches Steuerungs- und Entscheidungssystem
- Begriff Sich selbst verändernde Software-Werkzeuge
- Begriff Antizipierende vorausschauende Arbeitsgestaltung
- Begriff [↗ Moderierter Spezifikationsdialog](#)
- Kriterien zur sozialen Gestaltung
- Erfahrungsdiskurse

Das Konzept stärkt die Gestaltungskompetenz der Sozialpartner. Es betont unter anderem die wachsende Bedeutung der algorithmischen Steuerung horizontaler Wertschöpfungsketten sowie deren Rückwirkungen auf die Betriebe. Das Konzept wird seit längerem in Kooperation mit verschiedenen BMBF-Projekten, in INQA, mit Teilen der Sozialwissenschaften und der [↗ Arbeitsforschung](#) diskutiert und weiterentwickelt. Das Konzept ist anwendungsorientiert.

Quellen

- ifaa — Institut für angewandte Arbeitswissenschaft e.V., Stowasser S (2021) KI Zusatzqualifikationen. Für eine produktive und menschengerechte Arbeitsgestaltung. Düsseldorf
- Schröter W (2019) Der mitbestimmte Algorithmus. Arbeitsweltliche Kriterien zur sozialen Gestaltung von Algorithmen und algorithmischen Entscheidungssystemen. Mössingen

Autor

- Welf Schröter (Arbeitsgestaltung, Forum Soziale Technikgestaltung, KARL)

Emanzipation

Definition

Menschliche [Autonomie](#) und Entfaltung stellen einen Leitwert freiheitlicher Gesellschaften dar (Demirovic et al. 2018). Die Folgen technologischer Entwicklung für den Menschen sind im Kontext der [Arbeitswissenschaft](#) entsprechend auch daraufhin zu befragen, inwieweit sie diese Leitwerte zu realisieren helfen (Frey et al. 2021; Jochum 2021). Der Anspruch, dass der Einsatz von Künstlicher Intelligenz (KI) den Spielraum für menschliche Selbstentfaltung erweitern und nicht einschränken soll, wurde in seiner Wichtigkeit in der jüngsten Stellungnahme des Deutschen Ethikrats zu den Herausforderungen durch KI unterstrichen (Deutscher Ethikrat 2023). Für den Einsatz von KI bedeutet dieser Anspruch konkret, dass dieser die Freiheiten der einzelnen Beschäftigten, ihre individuellen Arbeitsprozesse selbst zu gestalten, erhalten oder sogar ausbauen soll. Verantwortung und die Letztentscheidung müssen beim Menschen verbleiben. Auch eine Unterwerfung der Beschäftigten unter Formen algorithmischen Managements ist vor diesem Hintergrund abzulehnen (siehe zum Beispiel Raffetseder et al. 2017).

Quellen

- Demirovic A, Lettow S, Maihofer A, Bromberg S (Hrsg.) (2018) Emanzipation: Zu Geschichte und Aktualität eines politischen Begriffs. Westfälisches Dampfboot, Münster
- Deutscher Ethikrat (2023) Ethikrat: Künstliche Intelligenz darf menschliche Entfaltung nicht verhindern. <https://www.ethikrat.org/mitteilungen/mitteilungen/2023/ethikrat-kuenstliche-intelligenz-darf-menschliche-entfaltung-nicht-vermindern/?cookieLevel=not-set>. Zugegriffen: 19.01.2024
- Frey P, Schaupp S, Wenten KA (2021) Towards Emancipatory Technology Studies. *Nanoethics* 15 (1):19–27. doi:10.1007/s11569-021-00388-6
- Jochum G (2021) Dialectics of Technical Emancipation—Considerations on a Reflexive, Sustainable Technology Development. *Nanoethics* 15 (1):29–41. doi:10.1007/s11569-021-00387-7
- Raffetseder EM, Schaupp S, Staab P (2017) Kybernetik und Kontrolle. Algorithmische Arbeitssteuerung und betriebliche Herrschaft. *PROKLA* 187 (2):229–248. doi:10.32387/prokla.v47i187.143

Autor

- Philipp Frey, M.A. (Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Erklärbare KI (XAI)/Erklärbarkeit

Definition

→ **Künstliche Intelligenz (KI)** und insbesondere Methoden des → **maschinellen Lernens (ML)** werden von verschiedenen Nutzenden verwendet, deren Hintergrundwissen und Fähigkeiten unterschiedlich stark ausgeprägt sind. Oft werden diese Systeme aufgrund ihrer Komplexität als Black Box wahrgenommen, vor allem da ihre Prozesse und Algorithmen nicht ohne weiteres vom Endnutzenden nachvollzogen werden können (Eldrandaly et al. 2023; Meske et al. 2022). Es werden daher Methoden der erklärbaren KI (XAI, explainable artificial intelligence) entwickelt, um eine gerechte und gleichverteilte Einsicht in diese teils autonom laufenden Prozesse zu ermöglichen. XAI konzentriert sich dabei auf die Erklärbarkeit und Verständlichkeit der Prozesse für Endnutzende (Gunning und Aha 2019) und erfüllt damit wichtige Voraussetzungen für → **transparente KI-Systeme** (Gunning et al. 2019). Dies ist insbesondere für Unternehmen relevant, da die Transparenz von KI-Systeme ein wichtiger Aspekt in der geplanten KI-Verordnung (engl. AI Act) der EU darstellt und die Erklärbarkeit von der hochrangigen Expertengruppe (AI HLEG) der Europäischen Kommission explizit als eine von drei zentralen Elementen der Transparenz aufgeführt wird (High-Level Expert Group on Artificial Intelligence 2020). Im Gegensatz zur → **interpretierbaren KI**, müssen dafür zusätzliche Erklärungen erzeugt werden, um die Logik der Entscheidungsfindung und die Ausgaben eines Modells nachvollziehen zu können (Antoniadi et al. 2021; Gunning et al. 2019; Gunning und Aha 2019; IBM 2023; Meske et al. 2022; Schwalbe und Finzel 2023). In Abbildung 1 wird diese Unterscheidung sowie die Beziehung zu den verwandten Begriffen transparentes KI-System und → **nachvollziehbare KI** dargestellt. Dabei können KI-Modelle durch verschiedene XAI-Ansätze erklärt werden, beispielsweise durch das Dokumentieren des Entscheidungsprozesses in Form generierter Texte, visueller Darstellungen oder das Hervorheben relevanter

Merkmale zur Entscheidungsfindung (Eldrandaly et al. 2023). Dadurch lässt sich das KI-Verhalten für unterschiedliche Situationen abschätzen und somit ein Eindruck gewinnen, was die Stärken und Schwächen des Modells sind (Gunning und Aha 2019). Dieser Ansatz ermöglicht Nutzenden einen einfacheren Zugang zu den verwendeten KI-/ML-Methoden und bietet den Mehrwert einer Effizienz- und Produktivitätssteigerung in Digitalisierungs- und Arbeitsprozessen, sowie eine Stärkung des Vertrauens und der Akzeptanz in die Systeme (Gunning et al. 2019; Gunning und Aha 2019; IBM 2023). Beispiele solcher XAI-Methoden sind: SHAP, LIME, LRP, Counterfactuals.



Abbildung 1: Abgrenzung zu den Begriffen transparentes KI-System, nachvollziehbare KI und interpretierbare KI. Quelle: eigene Darstellung

Quellen

- Antoniadi AM, Du Y, Guendouz Y, Wei L, Mazo C, Becker BA, Mooney C (2021) Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review. *Applied Sciences* 11:5088. doi:10.3390/app11115088
- Eldrandaly KA, Abdel-Basset M, Ibrahim M, Abdel-Aziz NM (2023) Explainable and secure artificial intelligence: taxonomy, cases of study, learned lessons, challenges and future directions. *Enterprise Information Systems* 17 (9):2098537. doi:10.1080/17517575.2022.2098537
- Gunning D, Aha DW (2019) DARPA's Explainable Artificial Intelligence (XAI) Program. *AIMag* 40 (2):44–58. doi:10.1609/aimag.v40i2.2850
- Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang G-Z (2019) XAI-Explainable artificial intelligence. *Sci Robot* 4 (37). doi:10.1126/scirobotics.aay7120
- High-Level Expert Group on Artificial Intelligence (2020) The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>. Zugegriffen: 22 September 2023

IBM (2023) What is explainable AI? | IBM. <https://www.ibm.com/topics/explainable-ai>. Zugegriffen: 22. September 2023

Meske C, Bunde E, Schneider J, Gersch M (2022) Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management* 39 (1):53–63. doi:10.1080/10580530.2020.1849465

Schwalbe G, Finzel B (2023) A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Min Knowl Disc*:1–59. doi:10.1007/s10618-022-00867-8

Autor:innen

- Maria Schweren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)
- Mandy Wölke (Human-Computer Interaction und Innovationswissenschaften, Institut für Angewandte Informatik (InfAI), KMI)
- Robin Weitemeyer, M.Sc. (Informatik, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Expertensystem

Definition

Als Expertensysteme werden solche Computerprogramme bezeichnet, die Menschen bei der Problemlösung oder der Entscheidungsfindung in der Funktion eines Experten oder einer Expertin unterstützen oder die die Entscheidungsfindung von menschlichen Expertinnen oder Experten nachahmen. Dazu verwenden sie eine von Expertinnen oder Experten gepflegte Sammlung von Fachwissen, auf deren Basis sie Schlussfolgerungen ziehen und Entscheidungen ableiten.

Das erste Expertensystem wurde 1965 in Form des Programms DENDRAL (Feigenbaum et al. 1971) entwickelt. DENDRAL diente zur Identifikation von molekularen Strukturen in Massenspektrometern. Dazu verwendete DENDRAL chemisches Expert:innenwissen. Ein anderes frühes Expertensystem zur Diagnose und Therapie von bakteriellen Infektionskrankheiten war MYCIN (Shortliffe 1976).

Schon früh fand auch eine Diskussion arbeitswissenschaftlicher Fragestellungen im Zusammenhang mit Expertensystemen statt, etwa in (Zink und Hauer 1991). Eine umfassende Einführung und Übersicht zu Expertensystemen gibt Jackson (1998). Auch heute werden Expertensysteme noch für die unterschiedlichsten Anwendungen eingesetzt. Eine Sichtung von Anwendungsfällen aus knapp 30 Jahren gibt etwa Wagner (2017). Ein Beispiel aus der jüngsten Vergangenheit ist die Entscheidungsunterstützung im Zusammenhang mit COVID-19 (Chrimes 2023).

Quellen

- Chrimes D (2023) Using Decision Trees as an Expert System for Clinical Decision Support for COVID-19. *Interact J Med Res* 12 (1):e42540. doi:10.2196/42540
- Feigenbaum EA, Buchanan BG, Lederberg J (1971) On generality and problem solving. A case study using the DENDRAL program. In: Meltzer B, Michie D (eds) *Machine Intelligence 6*, Edinburgh University Press, Edinburgh, Scotland, pp 165–190
- Jackson P (1998) *Introduction to Expert Systems*, 3rd edn. Addison-Wesley Longman Publishing Co., Inc., USA. doi:10.5555/521024
- Shortliffe EH (1976) *Computer-based medical consultations: MYCIN*. Elsevier. doi:10.1016/B978-0-444-00179-5.X5001-X
- Wagner WP (2017) Trends in expert system development: A longitudinal content analysis of over thirty years of expert system case studies. *Expert Systems with Applications* 76:85–96. doi:10.1016/j.eswa.2017.01.028
- Zink KJ, Hauer R (1991) Arbeitswissenschaftlich relevante Aspekte bei der Entwicklung und Einführung von Expertensystemen, Springer Berlin Heidelberg, Berlin, Heidelberg, 145–177.

Autoren

- Dr. Stefan Schiffer (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)
- Prof. Dr. Alexander Ferrein (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)

Human In Command (Control)

Definition

Bei der Einführung neuer Techniken und Arbeitsprozesse ist oft die Rede von einem Trade-off zwischen menschlicher Kontrolle und dem Level der Automatisierung, das erreicht werden kann (Shneiderman 2022). Häufig bestehen Bedenken bei der Nutzung und Entwicklung von Technik, inwieweit die [Autonomie](#) von Maschinen Menschen nicht nur bei der Arbeit unterstützt, sondern ersetzt (Fischer 2018). Bedenken bei der Einführung Künstlicher Intelligenz (KI) im Arbeitsprozess können bis zur Angst vor dem Wegfall von Arbeitsstellen reichen und für negative Reaktionen vor allem bei Nutzenden sorgen. Um zu überprüfen, ob eine KI bestimmte Handlungen unerwartet ausführt, die nicht mehr für Nutzenden nachvollziehbar ist, sollte deshalb der Mensch als kritische Kontrollinstanz fungieren (Conn 2022).

Bei der Einführung von KI-gestützten Systemen in der Organisation, sollte daher beachtet werden, Nutzenden Kontrolle über den Prozess, vereinzelte Prozessschritte und über den Kontrollbereich der Maschinen zu gewähren. In der generellen Anwendung sollte die KI der Kontrolle durch den Menschen unterliegen, da bestimmte Entscheidungsfindungen ethisch-moralischen Richtlinien unterliegen und rechtliche oder gesellschaftliche Auswirkungen nach sich ziehen (Conn 2022, Cavalcante et al. 2023). Seit 2021 empfiehlt auch der Europäische Wirtschafts- und Sozialausschuss mit Nachdruck, dies in einem KI-Gesetz zu regeln (Muller 2021).

Besonders in „Bereichen, wie Justiz, Strafverfolgung, Sozialdienste, Gesundheitsversorgung, Wohnungswesen, Finanzdienstleistungen, Arbeitsbeziehungen und Bildung“ (Muller 2021) sollten weiterhin endgültige Entscheidungen von Menschen getroffen werden, um die damit verbundene Rechenschaftspflicht gegenüber Mitmenschen und der Gesellschaft zu erfüllen.

Quellen

- Cavalcante Siebert L, Lupetti ML, Aizenberg E, Beckers N, Zgonnikov A, Veluwenkamp H, Abbink D, Giaccardi E, Houben GJ, Jonker CM, van den Hoven J, Forster D, Lagendijk RL (2023) Meaningful human control: actionable properties for AI system development. *AI Ethics* 3 (1):241–255. doi:10.1007/s43681-022-00167-3.
- Conn A (2022) What Does “Human Control” Over Autonomous Systems Mean? The challenge of human control over autonomous weapons. *IEEE Spektrum*. <https://spectrum.ieee.org/autonomous-weapons-control#toggle-gdpr>. Zugegriffen: 19. Januar 2024
- Fischer G (2018) Exploring design trade-offs for quality of life in human centered design. *Interactions* 25 (1):26–33. doi:10.1145/3170706.
- Muller C (2021) Europäischer Wirtschafts- und Sozialausschuss/Fachgruppe Binnenmarkt, Produktion, Verbrauch (22.12.2021): Stellungnahme des Europäischen Wirtschafts- und Sozialausschusses zum „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 — 2021/106(COD)). (2021/C 517/09. Amtsblatt der Europäischen Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206>. Zugegriffen: 19. Januar 2024
- Shneiderman B (2022) Ensuring Human Control over AI-Infused Systems. *NAE Perspectives* offer practitioners, scholars, and policy leaders a platform to comment on developments and issues relating to engineering. Hg. v. National Academy of Engineering (NAE) (Perspectives). <https://www.nationalacademies.org/news/2022/04/ensuring-human-control-over-ai-infused-systems>. Zugegriffen: 19. Januar 2024

Autorin

- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)

Human In / On / Out of the Loop

Definition

Human-in-the-loop, kurz HITL, bezeichnet konzeptionelle Interaktionen zwischen Menschen und einer Künstlichen Intelligenz (KI) und zielt darauf ab, die KI zu verbessern (Rueckert und Riedl 2022). Dabei wird der Einfluss eines a-priori-Wissens genutzt, über welches Entwicklerinnen und Entwickler sowie Nutzende verfügen, um ein möglichst genaues Vorhersagemodell der KI mit minimalen Kosten zu trainieren (Wu et al. 2022).

➤ **Machine learning-Verfahren**, die zum Trainieren einer KI genutzt werden, unterliegen bestimmten Einschränkungen, die abhängig von den Trainingsdaten sind. Das bezieht sich vor allem auf die Menge der Daten und deren Vollständigkeit (Rueckert und Riedl 2022). Auch ist es möglich, dass sich Modelle sehr starr verhalten und eine schlechte Adaptionfähigkeit besitzen, wenn sich die Datenlage über einen bestimmten Zeitraum verändert (Rueckert und Riedl 2022, Wu et al. 2022). Standard-Kernel-Maschinen kommen beispielweise nicht mit bestimmten Extrapolationsproblemen zurecht, die jedoch für menschliche Nutzenden als trivial angesehen werden (Holzinger 2016).

Um also eine bessere Anpassungsfähigkeit einer Maschine zu erreichen, ist es also notwendig, menschliche Expertise in der Weiterentwicklung einer KI einzubeziehen, um Fehleinschätzungen zu korrigieren und Feedback zu geben, um mit Problemen wie spärlichen Daten umzugehen. All dies führt zur Entstehung eines neuen Arbeitsprozesses, mit dem Menschen in einer zentralen Agenten-Rolle gesehen werden, die beim Treffen wichtiger Entscheidungen und der gleichzeitigen Verbesserung des KI-Modells mit neuen Trainingsdaten mitwirken (Schmid et al. 2023).

Ein weiterer Begriff für das Platzieren eines menschlichen Agenten zur Überwachung ist der Human-on-the-loop. Dies beinhaltet eine Überwachung des Systems und ein Eingreifen und Korrigieren des Prozesses, sollten fragwürdige Entscheidungen getroffen werden (Li et al. 2020).

Human-out-of-the-loop bezeichnet das Prinzip eines Automationsgrades, bei dem KI komplett autonom arbeitet und dem menschlichen Agenten höchstens das Deaktivieren

zufällt (Wessel 2023). Alle hier genannten Begriffe beschreiben ebenfalls entsprechende Konzepte für Automatisierungsgrade von Künstlicher Intelligenz „von Human Control (maximal Vorschläge des Computers), Human-in-the-Loop (Mensch muss Vorschläge/Entscheidungen der KI vor Ausführung bewilligen), Human-on-the-Loop (KI arbeitet normalerweise autonom, aber Mensch kann korrigierend eingreifen), bis Human-out-of-the-Loop (KI arbeitet autonom, kann sie höchstens deaktivieren).“ (Wessel 2023; Poretschkin et al. 2021).

Quellen

- Holzinger A (2016) Machine learning for health informatics. State-of-the-art and future challenges. Springer Cham. doi:10.1007/978-3-319-50478-0
- Li N, Adepu S, Kang E, Garlan D (2020) Explanations for Human-on-the-loop: A Probabilistic Model Checking Approach. Proceedings of the IEEE/ACM 15th International Symposium on Software Engineering for Adaptive and Self-Managing Systems:181–187. doi:10.1145/3387939.3391592
- Poretschkin M, Schmitz A, Akila M, Adilova L, Becker D, Cremers, AB. et al. (2021) Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz (KI-Prüfkatalog). Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Sankt Augustin
- Rueckert M, Riedl M (2022) Human-in-the-Loop: Wie Mensch und KI Aufgaben besser lösen. Digitale Welt 6 (4):36–39. doi:10.1007/s42354-022-0542-x
- Schmid T, Hildesheim W, Holoyad T (2023) Künstliche Intelligenz managen und verstehen. Der Praxis-Wegweiser für Entscheidungsträger, Entwickler und Regulierer. 1. Auflage. Beuth, Berlin, Wien, Zürich
- Wessel D (2023) Mensch-KI-System. In: Heine M, Dhungel AK, Schrills T, Wessel D (Hrsg.) Künstliche Intelligenz in öffentlichen Verwaltungen. Grundlagen, Chancen, Herausforderungen und Einsatzszenarien. Springer Gabler, Wiesbaden, S. 83–101
- Wu X, Xiao L, Sun Y, Zhang J, Ma T, He L (2022) A survey of human-in-the-loop for machine learning. Future Generation Computer Systems 135:364–381. doi:10.1016/j.future.2022.05.014

Autorin

- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)

Humanisierung/Menschzentrierung

Definition

Die Humanisierung von Arbeit ist durch die menschliche Würde begründet, die sich auf die Wahrung der [Autonomie](#) und der Unversehrtheit des individuellen Menschen bezieht (Negt 2011). Das der Humanisierung zugrundeliegende [Menschenbild](#) respektiert den Einzelnen in seinem [selbstbestimmten Urteilen](#), Entscheiden und Handeln. Zugleich mutet es dem Menschen Verantwortung für Wahrhaftigkeit und Solidarität zu (Braun 2023).

Humanisierung der Arbeit ist ein Oberbegriff für die Verbesserung der menschlichen Arbeitsbedingungen. Das in den 1960er Jahren etablierte Humanisierungskonzept sollte den Menschen von schwerer, einseitig belastender Arbeit tayloristischer Prägung befreien; es gilt als Antwort auf unzureichende Automatisierungsansätze (Schmidt 1982). Arbeitsschutz- und Mitbestimmungsgesetze sollten die Umsetzung der Humanisierung in der Arbeitswelt forcieren. Das Programm Humanisierung der Arbeitswelt (1974-89) fasste über 1600 Projekte zusammen, die von gewerkschaftlicher, staatlicher oder wissenschaftlicher Seite initiiert wurden (Müller 2019). Eine makro-ergonomische Perspektive ergänzt mittlerweile die arbeitspolitische Dimension der Humanisierung (Mütze-Niewöhner und Nitsch 2020).

Arbeitsbezogene Humanisierungsziele münden in einen menschenzentrierten Gestaltungsansatz (Schlick et al 2018). Dieser soll menschliches Arbeiten bereichern und

erleichtern. Durch einen partizipativen Ansatz wird das Wissen der Mitarbeiter in den Gestaltungsprozess einbezogen, um zu menschengerechten Arbeitsbedingungen, -plätzen oder -mitteln zu gelangen. Ein menschenzentrierter Gestaltungsansatz bezieht sowohl die Bedürfnisperspektive (zum Beispiel Kundenanforderungen an Produkte, Beschäftigtenanforderungen an [Assistenzsysteme](#)) als auch die Fähigkeitsperspektive (zum Beispiel die qualifikatorischen Voraussetzungen, Optimierung) ein.

Die menschenzentrierte Gestaltung grenzt sich von einem technikzentrierten Gestaltungsansatz ab. Letzterer ist darauf gerichtet, menschliche Leistungen im Rahmen der Arbeitsrationalisierung durch technische Systeme (zum Beispiel Robotik, [Künstliche Intelligenz](#)) zu ersetzen (Ulich 2011). Die digitale Transformation beschreibt eine aktuelle Stufe eines technikdeterministischen Rationalisierungskonzepts (Raehlmann 2019).

Angesichts fortschrittlicher Rationalisierungsmaßnahmen gewinnt die Humanisierung von digital unterstützter Arbeit – gleichsam als komplementäres Element – an Bedeutung. Unabdingbare menschliche Leistungsbeiträge in sozio-technischen Arbeitssystemen betreffen unter anderem eine Entscheidungskompetenz bei Ungewissheit (Hacker 2022), eine resonante Kommunikationsfähigkeit und eine damit verbundene Anpassungsfähigkeit an veränderte Umweltbedingungen.

Quellen

Braun M (2023) Wissenstransfer durch plattformgestützte Vernetzung – Impulse des Projekts CoCo für die Arbeitsforschung. In: Borowski E, Cernavin O, Hees F, Joerißen T (Hrsg) Erfolgreicher Transfer in der Arbeitsgestaltung. Waxmann, Münster, S 83-102

Hacker W (2022) Arbeitsgestaltung bei Digitalisierung. Merkmale menschenzentrierter Gestaltung informationsverarbeitender Erwerbsarbeit. Zeitschrift für Arbeitswissenschaft 76 (1):90-98. doi:10.1007/s41449-022-00302-0

Müller S (2019) Das Forschungs- und Aktionsprogramm „Humanisierung des Arbeitslebens“ (1974-1989). In: Kleinöder N, Müller S, Uhl K (Hrsg.) Humanisierung der Arbeit. Aufbrüche und Konflikte in der rationalisierten Arbeitswelt des 20. Jahrhunderts. Transcript, Bielefeld, S 59-88

Mütze-Niewöhner S., Nitsch V. (2020) Arbeitswelt 4.0. In: Frenz W. (Hrsg.) Handbuch Industrie 4.0: Recht, Technik, Gesellschaft. Springer, Berlin, Heidelberg, S 1187-1217. doi:10.1007/978-3-662-58474-3_61

Negt O (2011) Arbeit und menschliche Würde. Aus Politik und Zeitgeschichte 15:3-5

Raehlmann I (2019) Arbeit in der Digitalwirtschaft. Springer VS, Wiesbaden. doi:10.1007/978-3-658-26939-5

Schlick C, Bruder R, Luczak H (2018) Arbeitswissenschaft. 4. Auflage. Springer Vieweg, Wiesbaden

Schmidt G (1982) Humanisierung der Arbeit. In: Littek W, Rammert W, Wachtler G (Hrsg.)

Einführung in die Arbeits- und Industriesoziologie. Campus, Frankfurt, S 163-186

Ulich E (2011) Arbeitspsychologie. 7. Auflage. Schäffer-Poeschel, Stuttgart

Autor:innen

- Dr. Martin Braun (Arbeitswissenschaft, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Projekt CoCo – Connect & Collect)
- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)
- Mandy Wölke (Human-Computer Interaction und Innovationswissenschaften, Institut für Angewandte Informatik (InfAI), KMI)

Industrie 4.0

Definition

Industrie 4.0 ist die Bezeichnung für eines der zehn Zukunftsprojekte der Bundesregierung im Rahmen ihrer Hightech-Strategie für Deutschland (BMBF 2014). Die Entstehung des Begriffes basiert auf der Einschätzung, dass Wirtschaft und Gesellschaft – nach drei industriellen Revolutionen, geprägt durch Dampfmaschine, Fließband und Computer – zu Beginn der 2010er Jahre am Anfang einer vierten industriellen Revolution stünden. Diese werde durch das Internet der Dinge (engl. Internet of Things, IoT), das die Trennung von virtueller und realer Welt weitgehend aufhebt, ausgelöst. Dabei werden physische Systeme (things) drahtlos miteinander verbunden. Zusätzlich erhalten diese die Fähigkeit, über das Internet zu kommunizieren. Eng verbunden mit dem Begriff Internet der Dinge ist der Begriff des cyber-physischen Systems (CPS) (Kagermann et al. 2015). Hier wird zusätzlich zur reinen Kommunikation der physischen Systeme untereinander eine Analyse der physischen Systeme und eine daraus hervorgehende Steuerung und Optimierung ermöglicht (Lesch et al. 2023). Kernelement der Industrie 4.0 ist das Zukunftsbild der Smart Factory, in der eine Vielzahl von cyber-physischen Systemen zusammenwirken (Lucke 2013). Dieses Zukunftsbild bezieht sich auf alle Bestandteile eines Arbeitssystems. Dadurch ergeben sich vielfältige Gestaltungspotentiale zur Unterstützung diverser Aufgaben in der Arbeit. Das wird in Anlehnung an die Begrifflichkeit von Industrie 4.0 als Arbeitswelt 4.0 bezeichnet (ifaa 2018).

Insgesamt handelt es sich bei dem Konzept Industrie 4.0 um eine Vision, die neben der intelligenten Vernetzung von CPS vor allem Automatisierung, intelligente Datennutzung, Individualisierung, Flexibilität (in Produktionssystemen) und die [Mensch-Maschine-Interaktion](#) in den Fokus stellt (BMW 2023). Dabei ist das übergeordnete Ziel, die Wettbewerbsfähigkeit des Produktionsstandorts Deutschland zu stärken. Auch wenn die Realisierung des Gesamtszenarios noch nicht in greifbarer Nähe ist, lässt sich in den letzten Jahren ein zunehmender Implementierungsgrad von Industrie 4.0 in der betrieblichen Wirklichkeit beobachten.

Quellen

- Bundesministerium für Bildung und Forschung (BMBF) (2014) Die neue Hightech-Strategie; Innovationen für Deutschland. https://www.bmbf.de/bmbf/shreddocs/downloads/upload_filestore/pub_hts/hts_broschure_web.pdf?__blob=publicationFile&v=1. Zugegriffen: 28. September 2023
- Bundesministerium für Wirtschaft und Klimaschutz (BMWK) (2023) Was ist Industrie 4.0?. <https://www.plattform-i40.de/IP/Navigation/DE/Industrie40/WasIndustrie40/was-ist-industrie-40.html>. Zugegriffen 28. September 2023
- Institut für angewandte Arbeitswissenschaft e. V. (ifaa) (2018) Digitalisierung und Industrie 4.0 – Teil 1. https://www.arbeitswissenschaft.net/fileadmin/Downloads/Angebote_und_Produkte/Broschueren/ifaa_2016_Digitalisierung_I40.pdf. Zugegriffen: 28. September 2023
- Kagermann H, Lukas W-D, Wahlster W Industrie 4.0 (2015) Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. VDI Nachrichten 2015
- Lesch V, Züfle M, Bauer A, Iffländer L, Krupitzer C, Kounev S (2023) A literature review of IoT and CPS—What they are, and what they are not. *Journal of Systems and Software* 200:111631. doi:10.1016/j.jss.2023.111631
- Lucke D (2013) Smart Factory. In: Westkämper E, Spath D, Constantinescu C, Lentjes J (Hrsg.) *Digitale Produktion*. Springer Berlin Heidelberg, Berlin, Heidelberg, S 251–269

Autorin

- Jennifer Link, M.Sc. (Maschinenbau, ifaa – Institut für angewandte Arbeitswissenschaft e. V., WIRKSam)

Industrie 5.0

Definition

Industrie 5.0 steht für ein Konzept, welches die vorhandenen Ansätze der [Industrie 4.0](#) in ein erweitertes Wirtschaftsmodell überführt, in dem Wachstum nicht mehr nur wirtschafts- oder technologiegestützt betrachtet wird. Wachstum stellt sich hier als transformatorischer Ansatz dar, der den menschlichen Fortschritt und das Wohlbefinden aller beteiligten Personen fokussiert. Industrie 5.0 ist somit als eine Erweiterung des Fokus auf die Erhöhung der Produktivität der technologischen Systeme durch technische Innovationen, wie sie in der Industrie 4.0 forciert wurde, um ein stark menschzentriertes und nachhaltig orientiertes Paradigma zu verstehen.

Arbeitnehmerinnen und Arbeitnehmer werden in diesem mensch-zentrierten Ansatz noch stärker befähigt, ihre Aufgaben durch Nutzung digitaler Geräte und Unterstützung von [KI-Systemen](#) zu bewältigen, ohne dass ein Gefühl des Machtverlustes gegenüber der Technologie entsteht (European Commission 2021a). Dabei wird explizit darauf hingewiesen, den Menschen ins Zentrum der digitalen Transformation zu rücken und somit in allen Aspekten des Gestaltungsprozesses digitaler Systeme zu inkludieren.

Innovation und Forschung werden vor allem genutzt, um einen Fokus auf [Nachhaltigkeit](#), [Menschzentrierung](#) und Resilienz (Widerstandsfähigkeit) im gesamten Prozess zu legen. Diese drei Säulen sichern die ganzheitliche Transformation hin zur ökologisch nachhaltigen Nutzung von Technologien, die Erweiterung der Aufgabenbereiche und

die Verantwortung der Unternehmen gegenüber ihren Mitarbeitenden, und die Widerstandsfähigkeit von Unternehmen im Falle von Krisen (zum Beispiel Coronakrise) sowie die Ausschöpfung von Wertschöpfungsketten (European Commission 2021b).

Quellen

European Commission, Directorate-General for Research and Innovation, Breque M, De Nul L, Petridis A. (2021a) Industry 5.0 – Towards a sustainable, human-centric and resilient European industry, Publications Office of the European Union. doi:10.2777/308407

European Commission, Directorate-General for Research and Innovation, Renda A, Schwaag Serger S, Tataj D et al. (2021b) Industry 5.0, a transformative vision for Europe – Governing systemic transformations towards a sustainable industry, Publications Office of the European Union. doi:10.2777/17322

Autorinnen

- Maria Schveren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)
- Mandy Wölke (Human-Computer Interaction und Innovationswissenschaften, Institut für Angewandte Informatik (InfAI), KMI)

Intelligenz

Definition

Bei dem Begriff der Intelligenz (lat. intelligo: inter = zwischen, lego = wählen, lesen) handelt es sich nicht um ein klar umrissenes Konzept. Je nach disziplinärem Hintergrund wird er sehr unterschiedlich gedeutet (Legg und Hutter 2007), gleichwohl bezieht sich der Referenzrahmen in der Regel auf die Definition menschlicher Intelligenz. In der Psychologie umfasst Intelligenz einen Wissensbestand für nicht direkt beobachtbare Phänomene des Menschen. So bezeichnet es hier „die erworbenen kognitiven Fähigkeiten und Wissensbestände einer Person, die ihr zu einem gegebenen Zeitpunkt zur Verfügung stehen“ (Maier 2023).

Diese Fähigkeiten sind kulturell geprägt und beziehen sich auf die Gesamtheit kognitiver Kompetenzen zur Lösung bestimmter Probleme. Diese können logischer, sprachlicher, mathematischer Art sein. Darüber hinaus können sinnorientierte Probleme ebenfalls mittels dieser Fähigkeiten gelöst werden. In der Regel werden im Rahmen der menschlichen Intelligenz sieben Aspekte der Intelligenz beobachtet und gemessen: „Raumvorstellung, Sprachverständnis, Wortflüssigkeit, Rechenfertigkeit, Induktion, Wahrnehmungsgeschwindigkeit und mechanisches Gedächtnis“ (ebd.). Diese Aspekte werden in hohem Maße in sozio-technische Arbeitsumgebungen übersetzt. Darüber hinaus gibt es noch andere Formen der Intelligenz wie beispielsweise die Fähigkeit, Beziehungsstrukturen wahrzunehmen und zu gestalten oder die Fähigkeit, intuitives Wissen anzuwenden. Auch diese Beschreibungen werden mehr und mehr in [Mensch-Maschine-Interaktionen](#) umgesetzt und/oder (kritisch diskutiert) (Kopp et al. 2023).

Vor allem beim Einsatz von robotischen Systemen in der Krankenpflege oder in anderen sozialen Kontexten werden diese Formen der so genannten Künstlichen Intelligenz zunehmend mehr fokussiert. Dadurch entstehen wieder neue Abgrenzungen zu menschlichen Formen der Intelligenz.

Quellen

- Kopp T, Schäfer A, Kinkel S, Krings BJ, Cuhls K (2023) Editorial introduction to the Special Issue: Interactions between humans and social robots put to the test. *Technological Forecasting and Social Change*, 192, Art.-Nr.: 122556. doi:10.1016/j.techfore.2023.122556
- Legg Sh, Hutter M (2007) A Collection of Definitions of Intelligence. *Frontiers in Artificial Intelligence and Applications* 157:14-24
- Maier G (2023) Intelligenz. <https://wirtschaftslexikon.gabler.de/definition/intelligenz-37696>. Zugegriffen: 03. Juli 2023

Autorin

- Dr. Bettina-Johanna Krings (Technik- und Arbeitssoziologie sowie Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Interpretierbare KI

Definition

Interpretierbare KI bezeichnet KI-Modelle, deren Logik der Entscheidungsfindung in einer so strukturierten Form vorliegen, dass sie direkt vom Menschen verstanden werden können (Antoniadi et al. 2021; Meske et al. 2022; Mohseni et al. 2021; Schwalbe und Finzel 2023). Voraussetzung dafür ist die geringe Komplexität eines Modells und das notwendige technische Wissen, damit die Abbildung der Eingabedaten auf die KI-Ausgabe und die entsprechenden kausalen Zusammenhänge nachvollzogen werden können (Doran et al. 2017; Mohseni et al. 2021; Schwalbe und Finzel 2023). Im Gegensatz zur [erklärbaren KI](#) (XAI, explainable artificial intelligence) werden dafür keine weiteren Informationen aus zusätzlich erzeugten Erklärungen benötigt (Meske et al. 2022; Schwalbe und Finzel 2023). In Abbildung 2 wird diese Unterscheidung sowie die Beziehung zu den verwandten Begriffen [transparentes KI-System](#) und [nachvollziehbare KI](#) dargestellt.



Abbildung 2: Abgrenzung zu den Begriffen transparentes KI-System, nachvollziehbare KI und erklärbare KI. Quelle: eigene Darstellung

Neben den allgemeinen Nutzen einer nachvollziehbaren KI, wie gesteigertes Vertrauen und Akzeptanz der Endnutzenden (Gunning und Aha 2019), bieten interpretierbare KI-Modelle Unternehmen weitere Vorteile bei der Umsetzung transparenter KI-Systeme. Die Nachvollziehbarkeit der KI hängt hier nicht vom erfolgreichen Einsatz einer zusätzlichen XAI-Methode ab. Der Arbeitsaufwand für die Integration und Evaluation solcher

Methoden entfällt somit. Weiterhin bieten interpretierbare KI-Modelle einen direkten Einblick in ihre Logik der Entscheidungsfindung, während durch XAI erzeugte Erklärungen diese meist nur approximieren (Vale et al. 2022). Dies ist insbesondere relevant, wenn Transparenz aus rechtlichen Gründen gefordert wird (Vale et al. 2022).

Beispiele für interpretierbare KI-Modelle sind: Entscheidungsbäume, lineare Regressionen.

Quellen

- Antoniadi AM, Du Y, Guendouz Y, Wei L, Mazo C, Becker BA, Mooney C (2021) Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review. *Applied Sciences* 11:5088. doi:10.3390/app11115088
- Doran D, Schulz SC, Besold TR (2017) What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *CEUR Workshop Proceedings*
- Meske C, Bunde E, Schneider J, Gersch M (2022) Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management* 39:53–63. doi:10.1080/10580530.2020.1849465
- Mohseni S, Zarei N, Ragan ED (2021) A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Trans. Interact. Intell. Syst.* 11:1–45. doi:10.1145/3387166
- Schwalbe G, Finzel B (2023) A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Min Knowl Disc*:1–59. doi:10.1007/s10618-022-00867-8
- Vale D, El-Sharif A, Ali M (2022) Explainable artificial intelligence (XAI) post-hoc explainability methods: risks and limitations in non-discrimination law. *AI Ethics* 2:815–826. doi:10.1007/s43681-022-00142-y

Autor

- Robin Weitemeyer, M.Sc. (Informatik, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

KI-Akzeptanz

Definition

Im arbeitswissenschaftlichen Kontext bezieht sich KI-Akzeptanz auf die Bereitschaft und die Einstellungen von Arbeitnehmer:innen und Arbeitgeber:innen gegenüber KI-basierenden Technologien am Arbeitsplatz (Jung und von Garrel 2021). Generell ist Akzeptanz ein wichtiger sozialer und auch persönlicher Prozess, der es ermöglicht, Vielfalt und Verschiedenheit anzuerkennen und Neuerungen als Bereicherung statt als Hindernis zu sehen.

KI-Akzeptanz, als Teil der Technikakzeptanz, untersucht, wie Mitarbeitenden und Organisationen die Integration von KI-Technologien in ihren Arbeitsalltag wahrnehmen und annehmen. Die Herstellung oder Förderung der KI-Akzeptanz bei Mitarbeitenden im Rahmen von KI-Einführungen im Arbeitskontext, nimmt einen besonderen Stellenwert ein, da diese eine zentrale Voraussetzung für die erfolgreiche Interaktion zwischen KI-basierten Arbeitssystemen und den Mitarbeitenden herstellt (Schüth et al. 2021). KI-Akzeptanz ist dynamischer Natur und verläuft von der Einstellungsakzeptanz über die Nutzungsakzeptanz hin zur Handlungsakzeptanz (Kollmann 1998) hinsichtlich spezifischer Funktionen und Anwendungen von Technologien, die auf Verfahren der Künstlichen Intelligenz beruhen. Genauer gesagt versteht man unter diesen drei Phasen:

- **Einstellungsakzeptanz:** Dies bezieht sich auf die allgemeine Haltung oder Einstellung eines Mitarbeitenden gegenüber KI-Technologien. In dieser Phase prüfen die Individuen ihre Vorurteile, Bedenken und Erwartungen im Zusammenhang mit KI.
- **Nutzungsakzeptanz:** Nach der Einstellungsakzeptanz geht es darum, ob die Menschen bereit sind, [KI-Systeme](#) tatsächlich zu verwenden und in ihren Arbeitsablauf zu integrieren. Dies hängt von ihrer Wahrnehmung der Nützlichkeit und Benutzerfreundlichkeit der KI-Technologie ab.
- **Handlungsakzeptanz:** Diese Phase bezieht sich darauf, ob die Mitarbeitenden bereit sind, KI in ihren täglichen Aufgaben und Arbeitsabläufen aktiv einzusetzen. Dies erfordert oft eine Veränderung in ihrem Verhalten und auch in der Interaktion mit der KI-Technologie.

Des Weiteren gestaltet sich KI-Akzeptanz dabei in der Triade von Akzeptanzobjekt, -subjekt & -kontext aus (Schäfer und Keppler 2013). Genauer gesagt versteht man unter:

- **Akzeptanzobjekt:** Die Technologie selbst inklusive der Eigenschaften und Funktionen. Die Nutzenden müssen die Technologie verstehen (zum Beispiel in Bezug auf [Nachvollziehbarkeit](#) & [Transparenz](#)) (Mockenhaupt 2021).
- **Akzeptanzsubjekt:** Hierbei geht es um die Individuen, die die KI-Technologie verwenden. Ihre Einstellungen, Erfahrungen und auch Fähigkeiten spielen eine Rolle bei der KI-Akzeptanz.
- **Akzeptanzkontext:** Dies bezieht sich auf den Kontext oder auch die Umgebung, in der die KI-Technologie eingesetzt wird. Dies kann die Unternehmenskultur, Branche oder auch die gesetzlichen Bestimmungen umfassen.

Die Triade von Akzeptanzobjekt, -subjekt & -kontext verdeutlicht also, dass die erfolgreiche Einführung von KI am Arbeitsplatz nicht nur von der Technologie selbst abhängt, sondern auch von den Einstellungen der Mitarbeitenden und der Arbeitsumgebung, in der sie arbeiten.

Quellen

- Jung M, von Garrel J (2021) Mitarbeiterfreundliche Implementierung von KI-Systemen in Bezug auf Akzeptanz und Vertrauen. Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis (TATuP) 30 (3):37-43. doi:10.14512/tatup.30.3.37
- Kollmann T (1998) Akzeptanz innovativer Nutzungsgüter und Nutzungssysteme. Springer, Berlin
- Schäfer M, Keppler D (2013) Modelle der technikorientierten Akzeptanzforschung (Diskussionspapier 34/2013). Zentrum Technik und Gesellschaft, Technische Universität Berlin.
- Schüth NJ, Peifer Y, Weber, MA (2021) Entwicklungspotenziale der Künstlichen Intelligenz für die Mensch-Roboter-Kollaboration. In Arbeit interdisziplinär analysieren-bewerten-gestalten (Beitrag B 5). GfA-Press.

Autor:innen

- Prof. Dr. Joerg von Garrel (Sozioökonomie, Hochschule Darmstadt, KompAKI)
- Maria Jung (Arbeitsfähigkeit von Mitarbeitenden im KI-Arbeitskontext, Fachbereich Gesellschaftswissenschaften, Hochschule Darmstadt, KompAKI)
- Samantha Werens (Arbeitsfähigkeit von Mitarbeitenden im KI-Arbeitskontext, Fachbereich Gesellschaftswissenschaften, Hochschule Darmstadt, KompAKI)

KI-getriebenes Geschäftsmodell

Definition

Ein Geschäftsmodell beschreibt, wie ein Unternehmen mit seinen Aktivitäten einen Mehrwert schafft, wie es diesen seinen Kundinnen und Kunden bereitstellt, und wie es Einnahmen monetärer und nicht-monetärer Art es als Gegenleistung erhält. Geschäftsmodell-Komponenten sind nach dem weitverbreiteten Verständnis von Osterwalder und Pigneur (2010): Wertangebot, Schlüsselaktivitäten, Schlüsselressourcen, Schlüsselpartner, Kundenbeziehungen, Vertriebskanäle, Kundengruppen, Kosten, und Erlöse.

Geschäftsmodelle werden als KI-getrieben bezeichnet, wenn KI-Technologien elementarer Bestandteil mindestens einer ihrer Geschäftsmodell-Komponenten sind (Vetter et al. 2022; Hahn et al. 2020). Das bedeutet, dass die entsprechende Geschäftsmodell-Komponente ohne den Einsatz von KI-Technologien in der Form nicht oder nicht ähnlich profitabel realisierbar wäre.

So könnte [Künstliche Intelligenz \(KI\)](#) zum Beispiel eingesetzt werden, um Schlüsselaktivitäten im Unternehmen möglich zu machen, die von Menschen ohne KI-Unterstützung gar nicht oder nur mit sehr hohem Aufwand durchführbar wären, wie etwa durch die Verarbeitung sehr großer Datenmengen. Heutzutage wird KI bei KI-getriebenen Geschäftsmodellen jedoch am häufigsten in der Komponente Wertangebot eingesetzt. Hierbei werden KI-Technologien zur Erstellung neuer Dienstleistungen oder Produkte verwendet und kommen somit zur Erschaffung fundamental neuer Märkte mit neuen Kundengruppen zum Einsatz. Exemplarisch ermöglichen [KI-Systeme](#) somit das Angebot bahnbrechender Übersetzungsdienste wie DeepL und erlauben es Streamingdiensten wie Netflix und Spotify, ihren Nutzenden maßgeschneiderte Empfehlungen anzubieten, wodurch bei diesen neuartige Werte geschaffen werden.

Quellen

- Hahn C, Traunecker T, Niever M, Basedow GN (2020) Exploring AI-driven business models: Conceptualization and expectations in the machinery industry. Proceedings of the 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2020), S 567-570
- Hartmann PM, Zaki M, Feldmann N, and Neely A (2016) Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. International Journal of Operations & Production Management 36 (10):1382-1406
- Osterwalder A, Pigneur Y (2010) Business model generation: A handbook for visionaries, game changers, and challengers. John Wiley & Sons, Hoboken, NJ.
- Vetter OA, Hoffmann FS, Pumplun L, Buxmann P (2022) What constitutes a machine-learning-driven business model? A taxonomy of B2B start-ups with machine learning at their core. Proceedings of the 30th European Conference on Information Systems (ECIS 2022), S 1-18

Autor:innen

- Oliver A. Vetter, M.Sc. (KI-Geschäftsmodellentwicklung, Lehrstuhl Wirtschaftsinformatik | Software & Digital Business, Technische Universität Darmstadt, KompAKI)
- Maren F. Mehler, M.Sc. (KI-Geschäftsmodellentwicklung, Lehrstuhl Wirtschaftsinformatik | Software & Digital Business, Technische Universität Darmstadt, KompAKI)
- Prof. Dr. Peter Buxmann (Wirtschaftsinformatik, Lehrstuhl Wirtschaftsinformatik, Technische Universität Darmstadt, KompAKI)

KI-Methoden

Definition

KI-Methoden sind diejenigen Verfahren und Algorithmen, die bei der Umsetzung und Erstellung von [Künstlicher Intelligenz](#) zum Einsatz kommen (Nilsson 1998; Russell und Norvig 2020). Der Begriff Künstliche Intelligenz (KI) wurde durch John McCarthy und Kollegen in Rahmen eines Forschungsantrags aus dem Jahr 1955 (McCarthy et al. 1955) geprägt (Nachdruck in McCarthy et al. (2006)). Es gibt verschiedene Versuche, den Begriff der KI genauer zu fassen. So hat etwa die Europäische Union eine Expertengruppe gebildet, die eine Definition von KI erarbeitet hat (High-Level Expert Group on Artificial Intelligence 2019). Trotzdem gibt es keine eindeutige oder allgemeingültige Definition. Im Folgenden wird deshalb davon ausgegangen, dass KI die Menge von Techniken und Verfahren meint, die dazu nötig sind und die es erlauben, intelligentes Verhalten von und in technischen Artefakten zu erzeugen. Dies stützt sich weitgehend auf Nilsson (1998).

Häufig wird in der KI zwischen symbolischen und sub-symbolischen Methoden unterschieden (Shapiro 2003). Bei symbolischen Systemen findet eine explizite Repräsentation der Information statt, häufig in einer formalen Sprache. Diese Symbole erlauben üblicherweise eine Interpretation durch den Menschen. Sub-symbolische Systeme dagegen sind von außen meist nicht ohne weiteres interpretierbar. Es gibt jedenfalls keine direkt ablesbare semantische Bedeutung der internen Elemente (Shapiro 2003). Heute finden vermehrt Versuche statt, Systeme mit einer Kombination aus symbolischen und sub-symbolischen Anteilen zu entwickeln (Sarker et al. 2021).

Zur Problemlösung finden dabei ganz unterschiedliche Methoden Anwendung (Schmid et al. 2021). Dazu gehören unter anderem Suche, automatische Planung, mathematische Optimierung, Verfahren aus der Statistik, künstliche neuronale Netze aber auch Methoden der formalen Logik. Einige Methoden werden dabei aus anderen Disziplinen und Anwendungsfeldern wie der Mathematik adaptiert, andere sind eher spezifische Entwicklungen (Shapiro 2003). Die in der KI verwendeten Methoden decken ein sehr breites Spektrum ab und berühren häufig nicht nur einzelne, sondern mehrere Teilgebiete der Künstlichen Intelligenz. Die Teilgebiete umfassen unter anderem Problemlösung durch Suche, Wissensrepräsentation und Schlussfolgern, Schlussfolgerungen unter Unsicherheit, [maschinelles Lernen](#), maschinelle Bild- und Sprachverarbeitung, Mensch-Maschine Interaktion und Robotik (Shapiro 2003). Eine umfassende Darstellung der Verfahren gibt das Standardwerk von Russell und Norvig (2020).

Eine ähnliche Übersicht der KI-Methoden wird auch in dem von Shapiro aufgebrachten Konzept der KI-Vollständigkeit (AI-completeness) reflektiert (Shapiro 2003). Darin wird ein Teilgebiet der KI als KI-vollständig bezeichnet, wenn „[...] in the sense that solving the problem of the area is equivalent to solving the entire AI problem – producing a generally intelligent computer program“ (Shapiro 2003). Die angesprochenen Gebiete umfassen (1) Natural Language, (2) Problem Solving and Search, (3) Knowledge Representation and Reasoning, (4) Learning, (5) Vision, (6) Robotics, (7) Integrated Systems, and (8) Autonomous Agents. Eine weitere Übersicht, welche Bereiche und Methoden in der KI eine Rolle spielen, gibt beispielhaft auch Kersting et al. (2019). Eine Landkarte der Verfahren und Gebiete von KI-Anwendungen findet sich unter anderem auch in Humm (2020).

Quellen

High-Level Expert Group on Artificial Intelligence (2019) A Definition of AI: Main Capabilities and Disciplines

Humm BG (2020) Applied Artificial Intelligence: An Engineering Approach. Independently published

Kersting K, Peters J, Rothkopf C (2019) Was ist eine Professur für Künstliche Intelligenz? DOI 10.48550/ARXIV.1903.09516, in German

McCarthy J, Minsky ML, Rochester N, Shannon CE (1955) A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE. <http://jmc.stanford.edu/articles/dartmouth.html>

McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI Magazine 27 (4):12. doi:10.1609/aimag.v27i4.1904

Nilsson NJ (1998) Artificial Intelligence: A New Synthesis. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. doi:10.5555/2974990

Russell S, Norvig P (2020) Artificial Intelligence: A Modern Approach, 4th edn. Pearson

Sarker MK, Zhou L, Eberhart A, Hitzler P (2021) Neuro-symbolic Artificial Intelligence. AI Communications 34:197–209. doi:10.3233/AIC-210084, 3

Schmid T, Hildesheim W, Holoyad T, Schumacher K (2021) The AI Methods, Capabilities and Criticality Grid. KI – Künstliche Intelligenz 35 (3):425–440. doi:10.1007/s13218-021-00736-4

Shapiro SC (2003) Artificial intelligence (AI). In: Encyclopedia of Computer Science, pp 89–93

Autoren

- Dr. Stefan Schiffer (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)
- Prof. Dr. Alexander Ferrein (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)
- Joschua Schulte-Tigges, M.Eng. (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, AKzentE4.0)

KI-System

Definition

Ein KI-System – oft auch nur „Intelligentes System“ (Negnevitsky 2005) – ist ein System, das [Künstliche Intelligenz \(KI\)](#) anwendet. Ein KI-System besteht aus mehreren Komponenten, typischerweise mindestens aus einer Einheit für die Eingabe, die Ausgabe und einer KI-basierten Methode, welche heutzutage meistens auf [maschinellern Lernen](#) beruht. KI-Systeme können in verschiedenen Anwendungsbereichen eingesetzt werden, wie beispielsweise im Gesundheitswesen, in der Finanzindustrie, im Verkehrswesen, in der Produktion oder in der Sicherheit. Ein KI-System verarbeitet Eingabedaten, welche von Sensoren in Maschinen oder der Umwelt, Kameras, ERP- oder CRM-Systemen, bis hin zu Textdaten aus dem Internet kommen. Je nach Anwendungsbereich unterscheidet sich die Ausgabe eines KI-Systems. Beispiele hierfür sind die Beurteilung der Eingabedaten (Qualität, Fehler, usw.), Vorhersagen und Ziehen von Schlussfolgerungen, oder das direkte Ausführen von Aktionen. Jüngere Entwicklungslinien der Forschung (Kamar 2016) beschäftigen sich auch mit der Einbindung von Menschen in KI-Systeme ([Human-in-the-Loop](#)). Einige Beispiele für KI-Systeme sind:

- **Chatbots:** KI-Systeme, die in der Lage sind, menschenähnliche Gespräche zu führen und Anfragen zu beantworten, ohne dass ein menschlicher Operator involviert ist

- **Bild- und Spracherkennung:** KI-Systeme, die Bilder oder Sprache erkennen und analysieren, um bestimmte Muster oder Merkmale zu identifizieren
- **Autonome Fahrzeuge:** KI-Systeme, die in der Lage sind, ein Fahrzeug zu steuern und Entscheidungen auf der Grundlage von Sensor- und Kartendaten zu treffen
- **Personalisierte Empfehlungssysteme:** KI-Systeme, die Benutzerverhalten analysieren und personalisierte Empfehlungen für Produkte, Dienstleistungen oder Inhalte abgeben können
- **Predictive Analytics:** KI-Systeme, die Datenanalysen durchführen und Vorhersagen über zukünftige Ereignisse oder Trends treffen können

Quellen

Kamar E (2016) Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI):4070-4073.

Negnevitsky M (2005) Artificial intelligence: a guide to intelligent systems. Pearson education.

Autor

- Prof. Dr. Ralf Krestel (Informatik, Leibniz-Informationszentrum Wirtschaft ZBW, Projekt CoCo: Connect & Collect)

Komplexitätskompetenz

Definition

Mit Komplexitätskompetenz ist jenes Wissen benannt, die es Menschen ermöglicht, in nicht-transparenten und sich immer wieder verändernden Umgebungen selbstbewusst und souverän, dem eigenen Willen folgend zu agieren. Dabei bedeutet Kompetenz die Verknüpfung von Wissen und Handeln-Können. Angesichts der technischen Potenziale von sich selbst ständig verändernden Delegationssystemen kann ein gescheiterter Erwerb von Komplexitätskompetenz eine erhebliche Beeinträchtigung im Berufsleben und eine Ausgrenzung samt Zugangsbarriere darstellen (Schröter 2019). Das Vordringen von [Delegationstechnik](#) in den Arbeitsalltag verlangt die Erweiterung der beruflichen Bildung um Komplexitätskompetenz. Diese geht in ihrer Wirkungstiefe über das bisherige Verständnis von Medienkompetenz hinaus (Cernavin 2018).

Quellen

- Cernavin O, Schröter W, Stowasser S (2018) Prävention 4.0 – Neue Perspektiven für Führung, Organisation, Sicherheit und Gesundheit im Betrieb. Springer Fachmedien, Wiesbaden
- Schröter W (2019) Der mitbestimmte Algorithmus. Arbeitsweltliche Kriterien zur sozialen Gestaltung von Algorithmen und algorithmischen Entscheidungssystemen. Talheimer, Mössingen

Autor

- Welf Schröter (Arbeitsgestaltung, Forum Soziale Technikgestaltung, KARL)
- Dr.-Ing. Dipl.-Psych. Rico Ganßauge (Arbeitsorganisation, Fachgebiet Arbeitswissenschaft und Arbeitspsychologie der BTU Cottbus-Senftenberg, PAL)

Künstliche Intelligenz (KI)

Definition

Künstliche Intelligenz (KI) beschreibt die Fähigkeit von Maschinen, menschliche Fähigkeiten zu imitieren, die zur Lösung komplexer Probleme benötigt werden und dadurch für den Menschen intelligent wirken (McCarthy 2007, Russell S J 2010, Europäisches Parlament 2022, IBM Deutschland GmbH 2023, Humm et al. 2022). Die Maschine kann mit Hilfe von Datenerfassung ihre Umgebung wahrnehmen, indem sie die Daten verarbeitet und interpretiert. Darauf aufbauend zieht sie Schlussfolgerungen basierend auf ihrem verfügbaren Wissen und trifft entsprechend geeignete Maßnahmen, um teils autonom ein spezifisches Ziel zu erreichen (Winston 1992, Russell S J 2010, Europäische Kommission 2019, Europäisches Parlament 2022, Humm et al. 2022).

Zwei Domänen der KI bilden das maschinelle Lernen und die wissensbasierte KI, wobei sich diese beiden Bereiche überlappen und auch hybride Verfahren ermöglichen (Europäische Kommission 2019, Humm et al. 2022):

- [↗ Maschinelles Lernen \(ML\)](#) ermöglicht es Programmen, Vorhersagen zu treffen, indem sie Modelle auf der Grundlage vergangener Beobachtungen verwenden. Dafür nutzt ML numerische Verfahren (bspw. künstliche neuronale Netzwerke, Entscheidungs bäume), welche versuchen, Beobachtungen und Muster in Daten zu erkennen und somit implizit Wissen über ein Anwendungsfeld aufzubauen;

- Wissensbasierte KI, erlaubt es Programmen, Schlussfolgerungen auf Basis einer Wissensbasis durchzuführen. Eine Wissensbasis bildet dabei Wissen aus einem Anwendungsfeld ab und wird explizit vom Menschen modelliert (bspw. Wissensgraphen).

Die beiden Domänen sind von den menschlichen Fähigkeiten des Kommunizierens, Wahrnehmens, Lernens, Wissens, Handelns und Denkens inspiriert (Humm et al. 2022). Jede dieser Domänen umfasst eine Vielzahl von [↗ KI-Methoden](#) und [↗ KI-Systemen](#).

In der [↗ Arbeitswissenschaft](#) bezieht sich KI auf den Einsatz von maschinellen Systemen, um menschliche Arbeitsaufgaben zu unterstützen oder zu automatisieren (IFAA 2022). Solche Systeme können zu einer allgemeinen Skepsis gegenüber KI-Anwendungen bei Arbeitnehmenden führen (Harlacher et al. 2023). Ein breiter Einsatz von KI wird zwar zu einer Veränderung der Arbeit führen, Arbeitnehmende dabei jedoch vorrangig unterstützen und ihre Tätigkeiten vereinfachen (Osborne 2023).

Quellen

Europäische Kommission, Generaldirektion Kommunikationsnetze, Inhalte und Technologien (2019) Ethik-leitlinien für eine vertrauenswürdige KI.

Publications Office, <https://data.europa.eu/doi/10.2759/22710>

Europäisches Parlament (2022) Was ist künstliche Intelligenz und wie wird sie genutzt. <https://www.europarl.europa.eu/news/de/headlines/society/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt>. Zugegriffen: 19. Juni 2023

Harlacher M, Feggeler N, Peifer Y, Ottersböck N (2023) Produzierendes Gewerbe auf internationalem Niveau: Ergebnisse der Online-Befragung zum Thema „Künstliche Intelligenz in produzierenden Unternehmen“. Zeitschrift für wirtschaftlichen Fabrikbetrieb 118 (3):173-177. doi:10.1515/zwf-2023-1012

Humm BG, Buxmann P, Schmidt JC (2022) Grundlagen und Anwendungen von KI. In: Gethmann CF, Buxmann P, Distelrath J, Humm BG, Lingner S, Nitsch V, Schmidt JC, Spiecker genannt Döhmann I (Hrsg.) Künstliche Intelligenz in der Forschung. Ethics of Science and Technology Assessment, vol 48. Springer, Berlin, Heidelberg, S 13-42. doi:10.1007/978-3-662-63449-3_2

IBM Deutschland GmbH (2023) Was ist Künstliche Intelligenz (KI)? <https://www.ibm.com/de-de/topics/artificial-intelligence>. Zugegriffen: 19. Juni 2023

Institut für angewandte Arbeitswissenschaft (IFAA) e. V. (2022). Künstliche Intelligenz – Grundlagen, Anwendungen und Praxisbeispiele https://www.arbeitswissenschaft.net/fileadmin/Downloads/Angebote_und_Produkte/Broschueren/ifaa_Praxisbroschuere_KI_2022.pdf. Zugegriffen: 19. Juni 2023

McCarthy J (2007) What is artificial intelligence?. Stanford University, Stanford

Osborne M (2023) Generative AI and the future of work: a reappraisal. Brown Journal of World Affairs, Brown University, Providence, Rhode Island. <https://ora.ox.ac.uk/objects/uuid:f52030f5-23eb-4481-a7f1-8006685edbae>

Russell S J (2010) Artificial intelligence: a modern approach. Pearson Education, Inc., 3. Auflage, Upper Saddle River, New Jersey

Winston PH (1992) Artificial intelligence. Addison-Wesley Longman Publishing Co., Inc., 3. Auflage, Reading, MA

Autoren

- Alexander Zender, M.Sc. (Wissensbasierte KI, Institut für angewandte Informatik (aida) der Hochschule Darmstadt, KompAKI)
- Heiko Webert, M.Sc. (KI und Machine Learning in der Produktion, Zentrum für Robotik und Automation (CRA), Hochschule Darmstadt, KompAKI)
- Robin Weitemeyer, M.Sc. (Informatik, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Lernförderliche Arbeitsgestaltung

Definition

Unter dem Begriff lernförderliche [Arbeitsgestaltung](#) versteht man die Schaffung von Arbeitsbedingungen, die im Arbeitsablauf Lernprozesse begünstigen oder lerngünstige Voraussetzungen schaffen (Bigalk 2006). Hacker (2015) stellt vier Dimensionen lernförderlicher Arbeitsgestaltung heraus: Vollständigkeit der Aufgabe (planende, steuernde, ausführende und kontrollierende Elemente), Handlungsspielraum (zeitlich und inhaltlich), Anforderungsvielfalt (verschiedene Teiltätigkeiten) sowie Transparenz (soziale Unterstützung und Rückmeldungen über die Arbeit).

Hacker (2022) erweitert dieses Konzept der lernförderlichen Arbeitsgestaltung, indem er die Auswirkungen der Digitalisierung und des Einsatzes Künstlicher Intelligenz (KI) berücksichtigt. Besonders betont er, dass die Automatisierung anspruchsvoller kognitiver Tätigkeiten durch KI das Lernen nicht erübrigen sollte. Darüber hinaus hebt er hervor, dass ausführende Aufgaben und gestaltende (innovative) Tätigkeiten miteinander kombiniert werden sollten, um eine ganzheitliche und informatisierte Arbeit zu schaffen. Schließlich unterstreicht er die Bedeutung nicht-digitaler Lerngegenstände (zum Beispiel nicht verbalisiertes Wissen der Mitarbeitenden) und Lernprozesse (zum Beispiel Vorbildlernen), da diese die Grundlage für die Ausübung von Tätigkeiten und Entscheidungsfindung bereitstellen. Die Bedeutsamkeit einer lernförderlichen Arbeitsgestaltung zeigt sich insbesondere in den positiven Auswirkungen auf die Mitarbeitenden. Dazu zählen die Kompetenzentwicklung, gesteigerte Motivation,

erhöhte Lernbereitschaft, Reduktion von Fehlern und Persönlichkeitsförderung. Zudem zeigen sich langfristig positive Effekte auf die Organisation, da sie in der Folge zu einer höheren Produktivität, besseren Arbeitsergebnissen, der Erschließung von Innovationspotenzialen sowie zu einer verringerten Mitarbeitendenfluktuation führt (Bigalk 2006).

Quellen

- Bigalk D (2006) Lernförderlichkeit von Arbeitsplätzen – Spiegelbild der Organisation? Eine vergleichende Analyse von Unternehmen mit hoch und gering lernförderlichen Arbeitsplätzen. Kassel University Press, Kassel
- Hacker, W (2015) Lern-, gesundheits- und leistungsförderliche Arbeitsgestaltung in kleinen und mittleren Unternehmen – Warum und wie? Technische Universität Dresden, Dresden
- Hacker W (2022) Arbeitsgestaltung bei Digitalisierung. Zeitschrift für Arbeitswissenschaft 76 (1):90–98. doi:10.1007/s41449-022-00302-0

Autorin

- Paula Danhausen, M.Sc. (Psychologie, FIR an der RWTH Aachen, Bereich Smart Mobility, AKzentE4.0)

Maschinelles Lernen

Definition

Maschinelles Lernen (ML) ist ein Teilgebiet der [Künstlichen Intelligenz \(KI\)](#), das IT-Systemen auf Basis von (Beispiel-) Daten und Erfahrungswerten ermöglicht, aus diesen zu lernen, ohne dafür explizit programmiert worden zu sein. Mit Hilfe von großen Datenmengen werden Lernalgorithmen darauf trainiert, Muster sowie Zusammenhänge zu erkennen, Regeln abzuleiten und somit Wissen aus Erfahrung zu generieren (Döbel et al. 2018). Dies erfolgt meist mittels künstlicher neuronaler Netze ([Deep Learning](#)) (Terstegen und Link 2022). Weitere mögliche Algorithmen sind beispielsweise Entscheidungsbäume, lineare oder logistische Regressionen und Support-Vektor Maschinen. Maschinelles Lernen kann dabei in drei Methoden unterteilt werden, dem supervised learning, unsupervised learning und reinforcement learning. Beim supervised learning werden bekannte und gekennzeichnete Trainingsdaten, wie beispielsweise Bilder, Texte oder Sensordaten, sowie eine definierte Zielvariable verwendet, um Muster und Zusammenhänge zu erlernen. Dahingegen lernt der Algorithmus beim unsupervised learning selbständig, aus nicht gekennzeichneten Eingabedaten Muster und Cluster zu erkennen. Bei der dritten Methode – reinforcement learning – werden keine Beispieldaten benötigt, da die Algorithmen Belohnungen von Handlungen mit positivem Ergebnis erhalten und dadurch lernen, welche Handlungen die richtigen waren. Diese erlernten Muster und gewonnenen Erkenntnisse werden dann auf unbekannte Daten angewendet, um Vorhersagemodelle zu erstellen, Entscheidungen zu treffen und Empfehlungen zu generieren (Mich 2020; Seifert et al. 2018). Dadurch können wiederholende und zeitaufwändige Aufgaben automatisiert und somit Arbeitsprozesse und deren Effizienz und Genauigkeit optimiert werden. Mitarbeitende sollen dabei keinesfalls ersetzt, sondern deren Arbeitsumgebung und -ergebnisse optimiert werden, indem maschinelles Lernen als Hilfsmittel eingesetzt wird, um mit Hilfe der daraus resultierenden Modelle Fachkräfte unter Berücksichtigung ihrer Fähigkeiten und Bedürfnisse zu unterstützen.

Die Erzeugung dieser Anwendungen des Maschinellen Lernens und dessen Modelle sind mit großem Aufwand sowie hohen Kosten verbunden und lediglich auf eine geringe Aufgabenbreite trainiert und anwendbar. Dies hängt unter anderem damit zusammen, dass eine Vielzahl von Trainingsdaten gesammelt, und von Fachpersonal analysiert sowie teilweise gelabelt werden muss. Neben dem hohen Trainingsaufwand der Modelle zählen die Robustheit, Sicherheit, [Transparenz](#) sowie Adaptivität dieser und die hohe Ressourcenintensität sowie das benötigte geschulte Fachpersonal zu weiteren Restriktionen von Maschinellern Lernen (Döbel et al. 2018).

Quellen

- Döbel I, Leis M, Molina Vogelsang M, Neustroev D, Petzka H, Riemer A, Rüping S, Voß A, Wegele M, Welz J (2018) Maschinelles Lernen: Eine Analyse zu Kompetenzen Forschung und Anwendung. <https://publica.fraunhofer.de/handle/publica/299081>
- Mich L (2020) Artificial Intelligence and Machine Learning. In: Xiang Z, Fuchs M, Gretzel U, Höpken W (Hrsg.) Handbook of e-Tourism. Springer International Publishing, Cham, S 1–21
- Seifert I, Bürger M, Wangler L, Christmann-Budian S, Rohde M, Gabriel P, Zinke G (2018). Potenziale der Künstlichen Intelligenz im produzierenden Gewerbe in Deutschland: Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi) im Rahmen der Begleitforschung zum Technologieprogramm PAiCE — Platforms | Additive Manufacturing | Imaging | Communication | Engineering. Berlin: iit – Institut für Innovation und Technik in der VDI/VDE Innovation und Technik GmbH
- Terstegen S, Link J (2022) Künstliche Intelligenz: Grundlagen, Anwendungen und Praxisbeispiele

Autorin

- Elena Kick, M.Sc. (Wirtschaftswissenschaften, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Menschenbild

Definition

Das Konzept des Menschenbild wird schon sehr früh in der philosophischen Anthropologie verhandelt und ist durch die Diagnose über die Besonderheit des Menschen im Kosmos gekennzeichnet. Mit Autoren wie Max Scheler, Helmut Plessner und Arnold Gehlen seien beispielhaft drei Autoren genannt, die sich intensiv mit dem Konzept des Menschenbilds beschäftigt und zu deren inhaltlichen Ausgestaltung wesentlich beigetragen haben. So resultiert beispielsweise für Scheler die Sonderstellung des Menschen im Charakter als Mängelwesen im Hinblick auf seine fehlende Kapazität, sich in seine biologische Umwelt einzubinden. Vor seinen Annahmen (Gutmann 2013) hat sich der Mensch historisch als ein handelndes Wesen bestimmt. „Dabei steht im Zentrum die Vermutung, dass Technik als Medium der Transformation von Natur zu menschlichen Zwecken diejenige Eigenschaft sei, die das biologisch defiziente Wesen ‚in der Natur‘ halte“ (Gutmann 2013, S. 96). Technologien als erweiterter Handlungsspielraum des Menschen werden hier zum Kontrapunkt zum „mangelhaften“ Menschen.

Im Laufe der letzten Jahrzehnte gibt es eine Vielzahl an Beschreibungen im Hinblick auf das Menschenbild. In den letzten Jahren entwickelte sich hier ein starker Schwerpunkt in Mensch-Maschine-Modellen in den Wissenschaften. Dies ist auch und vor allem auf das sich veränderte Mensch-Maschine-Verhältnis zurückzuführen. Während beispielsweise thermodynamische Maschinen in Arbeitsbereichen „die Funktion der Arbeitserleichterung und Leistungssteigerung hatten“ (Weber 2020, S. 319), werden digitale Technologien vielseitig genutzt und verschmelzen mit den Handlungsräumen des Menschen in ungekanntem Maße. Diese Veränderungen mobilisieren einerseits traditionelle Vorstellungen des Menschenbildes, andererseits werden neue Menschenbilder generiert, wie beispielsweise das Bild des Cyborgs anschaulich zeigt (Weber 2020).

Insgesamt gilt, wenn der Mensch als Teil der Welt verstanden wird, so ist das Menschenbild auch Teil des Weltbildes. Menschenbilder und demnach auch Weltbilder prägen umfassende Perspektiven und Lehren wie beispielsweise das humanistische Menschen- und Weltbild. Dieses Beispiel weist darüber hinaus darauf, dass Menschenbilder einen normativen Charakter haben können und im Hinblick auf soziale Sinnstiftung auch sollen. So beherbergen die übergeordneten Menschenbilder religiöse/ ethische, soziale und politische Dimensionen gleichermaßen und können einen sinnstiftenden und gestaltenden Charakter auf der sozialen sowie auf der institutionellen Ebene bilden.

Quellen

- Gruevska J, Lessing HU (2020) Philosophische Anthropologie. In: Heßler M, Liggieri K (Hrsg.) Technikanthropologie – Handbuch für Wissenschaft und Studium. Nomos, Baden-Baden, S 33-44
- Gutmann M (2013) Philosophische Anthropologie. In: Grunwald A (Hrsg.) Handbuch Technikethik. Metzler, Stuttgart, S 94-98
- Weber J (2020) MenschMaschine. In: Heßler M, Liggieri K (Hrsg.) Technikanthropologie – Handbuch für Wissenschaft und Studium. Nomos, Baden-Baden, S 318-322

Autorin

- Dr. Bettina-Johanna Krings (Technik- und Arbeitssoziologie sowie Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Menschengerechte Arbeitsgestaltung

Definition

Das Prinzip der menschengerechten [Arbeitsgestaltung](#) legt den Fokus auf gesundheitsförderliche Aspekte der Arbeitstätigkeit, welche über die reine Vermeidung von Risiken und Gefährdungen hinausgehen, und die körperlichen und psychischen Bedingungen des Menschen berücksichtigen (Weber et al. 2022). Luczak und Volpert (1987) führen in ihrer Kriterienliste die Beurteilungskriterien von Rohmert (1983) und Hacker und Richter (1984) zusammen und ergänzen sie. Folglich definieren sie fünf Bewertungskriterien menschengerechter Arbeitsgestaltung: Ausführbarkeit, Schädigungslosigkeit, Beeinträchtigungsfreiheit, Persönlichkeitsförderlichkeit sowie Sozialverträglichkeit.

In Zusammenhang mit der menschengerechten Arbeitsgestaltung wurden diverse Gesetze und Richtlinien erlassen (Arbeitsschutzgesetz). Betrachtet man die menschengerechte Arbeitsgestaltung im Kontext der Digitalisierung und der zunehmenden Bedeutsamkeit von Künstlicher Intelligenz (KI), sollten insbesondere die Erhaltung und Erweiterung der Kompetenzen berücksichtigt werden. Zudem sollte die Förderung des gesundheitlichen Wohlbefindens und der Arbeitsleistung betrachtet werden (Hacker 2022). So gilt es zum Beispiel bei der Einführung von intelligenten Maschinen in der Produktion, die Auswirkungen auf den Menschen im Vorhinein zu berücksichtigen. So kann zum Beispiel eine veränderte Arbeitsteilung zwischen Mensch und Maschine mit monotonen Restaufgaben für den Menschen einhergehen, wodurch negative Auswirkungen zum Beispiel durch einseitige Belastungen, Demotivation und Dequalifizierung zu erwarten sind. Solche Auswirkungen sind im Sinne einer menschengerechten Arbeitsgestaltung zu vermeiden.

Siehe auch [Humanisierung/Menschzentrierung](#)

Quellen

- Arbeitsschutzgesetz (ArbSchG). <https://www.gesetze-im-internet.de/arbSchg/>. Zugegriffen: 07. September 2023
- Hacker W, Richter P (1984) Psychische Fehlbeanspruchung. Psychische Ermüdung, Monotonie, Sättigung, Stress. VEB Deutscher Verlag der Wissenschaften, Berlin
- Hacker W (2022) Arbeitsgestaltung bei Digitalisierung. Zeitschrift für Arbeitswissenschaft (76):90-98. doi:10.1007/s41449-022-00302-0
- Luczak H, Volpert W (1987) Arbeitswissenschaft. Springer, Berlin, Heidelberg
- Rohmert W, (1983) Formen menschlicher Arbeit. In: Rohmert W, Rutenfranz J (Hrsg.) Praktische Arbeitsphysiologie. Georg Thieme Verlag, Stuttgart, New York, S 5-29
- Weber C, Tegmeier P, Sommer S, Tisch A & Wischniewski S (2022) I. Kriterien einer menschengerechten Gestaltung von Arbeit in der digitalisierten Arbeitswelt. Sicherheit und Gesundheit in der digitalisierten Arbeitswelt, S 11–46. doi:10.5771/9783748927372-11

Autorin

- Annika Franken, M.Sc. (Psychologie, FIR an der RWTH Aachen, Bereich Smart Mobility, AKzentE4.0)

Mensch-Maschine-Interaktion

auch: HCI (Human-Computer Interaction) oder HRC (Human-Robot Collaboration)

Definition

In der Arbeitswissenschaft, im Bereich des [maschinellen Lernens](#) und der Robotik bezieht sich die Mensch-Maschine-Interaktion (MMI), auch Mensch-Maschine-Kommunikation, auf die Analyse, Gestaltung und Optimierung von Kommunikations- und Interaktionsprozessen zwischen Menschen und technischen Systemen, insbesondere Computersystemen und Maschinen. Die Maschine kann jegliches Gerät sein, das über eine Schnittstelle mit dem Menschen kommunizieren kann. Im Fachbereich der Robotik werden bereits aufeinander abgestimmte Handlungen im selben Arbeitsraum als Formen der [Mensch-Roboter-Interaktion](#) angesehen, auch wenn es nicht zu einer Kommunikation über eine gemeinsame Schnittstelle kommt (Kopp et al. 2021; Aaltonen et al. 2018; Malik und Bilberg 2019; Onnasch 2016). Dieses interdisziplinäre Forschungsgebiet umfasst Aspekte der Informatik, Psychologie, Ergonomie, Design und Sozialwissenschaften. Ziel ist es, die Effizienz, Sicherheit und Zufriedenheit der Nutzenden zu verbessern und die Zusammenarbeit zwischen Mensch und Maschine zu fördern (Sanders und McCormick 1993; Dix 2004) sowie die Schnittstelle möglichst bedienungsfreundlich und intuitiv zu gestalten.

Im Kontext der [Arbeitswissenschaft](#) fokussiert die Mensch-Maschine-Interaktion die Gestaltung von Arbeitsplätzen, Arbeitsabläufen und Arbeitsumgebungen, in denen Menschen und Maschinen effektiv und sicher zusammenarbeiten können. Dabei spielen ergonomische Aspekte eine wichtige Rolle, wie beispielsweise die Anpassung von Maschinen und Arbeitsplätzen an die körperlichen und kognitiven Fähigkeiten sowie der Wahrnehmung des Menschen, die Unterstützung bei Entscheidungsprozessen und die daraus resultierende Minimierung von Fehlern, Unfällen sowie die Wahrung der körperlichen und psychischen Gesundheit (Wickens et al. 1998).

Im Bereich des [maschinellen Lernens](#) liegt ein Schwerpunkt der Mensch-Maschine-Interaktion auf der Entwicklung von Algorithmen und Systemen, die in der Lage sind, menschliches Verhalten und menschliche Entscheidungen zu erkennen, zu unterstützen, vorherzusagen und darauf zu reagieren. Hierzu gehören unter anderem die automatische Anpassung von Benutzeroberflächen, die Erkennung von Benutzerintentionen, die Entwicklung von Sprach- und Gestenerkennungssystemen sowie die Gestaltung von kollaborativen Robotern (Cobots) und Entscheidungsunterstützungssysteme für komplexe Situationen, die in enger Zusammenarbeit mit Menschen arbeiten können.

Die Mensch-Maschine-Interaktion in der Robotik umfasst die Kommunikation und Zusammenarbeit zwischen Menschen und Robotern auf physischer, sprachlicher und kognitiver Ebene. Die Gestaltung von Benutzerschnittstellen ist dabei entscheidend, wobei Aspekte wie Benutzerfreundlichkeit, Ästhetik und Sicherheit eine zentrale Rolle spielen, um eine positive und effiziente Interaktion zu gewährleisten.

In allen drei Kontexten ist die Mensch-Maschine-Interaktion von zentraler Bedeutung, um die Potenziale neuer Technologien auszuschöpfen, die Arbeitsbedingungen zu verbessern und gleichzeitig eine hohe Arbeitsleistung, -qualität und -motivation zu gewährleisten (Bauernhansl 2014).

Quellen

- Aaltonen I, Salmi T, Marstio I (2018) Refining levels of collaboration to support the design and evaluation of human-robot interaction in the manufacturing industry. *Procedia CIRP* 72: 93–98, Jan. 2018. doi:10.1016/j.procir.2018.03.214
- Bauernhansl T, ten Hompel M, Vogel-Heuser B (2014) *Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung, Technologien, Migration*. Springer Fachmedien, Wiesbaden. doi:10.1007/978-3-658-04682-8
- Dix A (2004) *Human-computer interaction*, 3rd ed. Pearson/Prentice-Hall, Harlow, England, New York
- Kopp T, Baumgartner M, Kinkel S (2021) Success factors for introducing industrial human-robot interaction in practice: an empirically driven framework. *Int J Adv Manuf Technol* 112 (3):685–704. doi:10.1007/s00170-020-06398-0
- Malik AA, Bilberg A (2019) Collaborative robots in assembly: A practical approach for tasks distribution. *Procedia CIRP* 81:665–670. doi:10.1016/j.procir.2019.03.173
- Onnasch L (2016) *Mensch-Roboter-Interaktion – Eine Taxonomie für alle Anwendungsfälle*. *baua: Fokus* 1:1–12. doi:10.21934/baua:fokus20160630
- Sanders MS, McCormick EJ (1993) *Human factors in engineering and design*, 7th ed. McGraw-Hill, New York
- Wickens CD, Gordon SE, Liu Y (1998) *An introduction to human factors engineering*. Longman, New York

Autoren

- Philip Schäfer, M.Sc. (Informatik, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Mohammad Ehsan Matour, M. Sc. (Elektro- und Automatisierungstechnik, Professur Fertigungsautomatisierung der Hochschule Mittweida, PAL)



Mensch-Roboter Interaktion

Definition

Laut Fong et al. (2003) beschreibt die Mensch-Roboter Interaktion Untersuchungen von Menschen, Robotern und die Art und Weise, wie sich diese gegenseitig beeinflussen. Dieses Studienfeld beschäftigt sich mit dem Verständnis, dem Design und der Evaluation von robotischen Systemen, die von oder im Zusammenhang mit Menschen genutzt werden (Gervasi et al. 2020). Sein Ursprung liegt im Feld der [Mensch-Maschine Interaktion](#) (Salvine et al. 2011).

Die Interaktionsform kann unterschiedlich ausgeprägt sein: Ko-Existenz bezieht sich auf eine bestimmte Episode der Interaktion und ist damit räumlich und zeitlich begrenzt. Bei einer Kooperation oder Kollaboration geht es um Zusammenarbeit. Es besteht entweder eine klare Aufgabenteilung oder eine gemeinsame Zielverfolgung (Onnasch et al. 2016).

Ein mögliches Einsatzgebiet ist hierbei der Arbeitskontext, hier ist die Triade Mensch-Technologie-Organisation von Bedeutung. Für die erfolgreiche Einführung von Robotern in Unternehmen ist es wichtig, die Mitarbeitenden im Sinne eines [menschenzentrierten Ansatzes](#) einzubinden, [Partizipation](#) zu ermöglichen und das System an ihren Bedürfnissen auszurichten.

Für die Mitarbeitenden ist eine erfolgreiche Einführung des Roboters wichtig. Dies schafft einerseits Vertrauen bei den Mitarbeitenden und ermöglicht andererseits eine geeignete Konfiguration des Roboters, beides sind zentrale Erfolgsfaktoren der Mensch-Roboter-Interaktion. Als Erfolgsfaktoren zählen Vertrauen und die passende Konfiguration des Roboters (Kopp et al. 2021). Als Vorteil für Organisationen kann die Integration von Robotern Effizienz erhöhen (Gombolay et al. 2015). Aus technologischer Sicht profitieren Roboter von Entwicklungen in verwandten Bereichen, wie zum Beispiel der [Künstlichen Intelligenz](#). Dies ermöglicht eine bessere Unterstützung der Mitarbeitenden (Chui et al. 2022).

Quellen

- Chui M, Roberts R, Yee L (2022) Generative AI is here: How tools like ChatGPT could change your business. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/generative-ai-is-here-how-tools-like-chatgpt-could-change-your-business#/>
- Fong T, Thorpe C, Baur C (2003) Collaboration, Dialogue, Human-Robot Interaction. In: Jarvis RA, Zelinsky A (Hrsg) Robotics Research. Springer Berlin Heidelberg, Berlin, Heidelberg, S 255–266
- Gervasi R, Mastrogiacomo L, Franceschini F (2020) A conceptual framework to evaluate human-robot collaboration. Int J Adv Manuf Technol 108:841–865
- Gombolay MC, Gutierrez RA, Clarke SG, Sturla GF, Shah JA (2015) Decision-making authority, team efficiency and human worker satisfaction in mixed human-robot teams. Autonomous robots 39:293–312
- Kopp T, Baumgartner M, Kinkel S (2021) Success factors for introducing industrial human-robot interaction in practice: an empirically driven framework. Int J Adv Manuf Technol 112:685–704
- Onnasch L, Maier X, Jürgensohn T (2016) Mensch-Roboter-Interaktion - Eine Taxonomie für alle Anwendungsfälle
- Salvine P, Nicolescu M, Ishiguro H (2011) Benefits of HumanRobot Interaction [TC Spotlight]. IEEE Robotics & Automation Magazine 18:98–99

Autorin

- Lea Heitlinger, M.Sc. (Psychologie, Fachgebiet für Marketing & Personalmanagement der TU Darmstadt, KompAKI)

Menschenzentrierte KI

Definition

Eine menschenzentrierte KI (human-centered artificial intelligence, HCAI) ist eine KI-Anwendung, die nach den Prinzipien des menschenzentrierten Designs (HCD, ISO 9241-210:2019) entwickelt wurde. Der Fokus liegt neben technischen, algorithmischen und ökonomischen Aspekten vor allem auf einer menschenzentrierten Perspektive (Shneiderman 2022; Zinke-Wehlmann et al. 2022; Auernhammer 2020).

Zu den zentralen Prinzipien des HCD gehören das Verständnis von Nutzenden, Aufgaben und Umgebungen (Zinke-Wehlmann et al. 2022), die Einbeziehung der Nutzenden in Design und Entwicklung (Auernhammer 2020) und das Vorhandensein von multidisziplinären Fähigkeiten und Perspektiven im Designteam (International Organization for Standardization 2019). Im Gegensatz zum traditionellen nutzerzentrierten Design (user-centered design, UCD) werden Menschen ganzheitlicher und umfassender betrachtet, nicht nur als Nutzende oder als „Faktor Mensch“ (Auernhammer 2020; Hyysalo und Johnson 2015).

HCD hat zum Ziel, menschlichen Bedürfnissen gerecht zu werden, zum Beispiel indem eine KI die betroffenen Personen am Arbeitsplatz unterstützt, anstatt sie komplett zu ersetzen (International Organization for Standardization 2019). Insbesondere bei KI-Anwendungen sind häufig die Bedürfnisse großer Personengruppen und verschiedenster Stakeholder wie Aufsichtsbehörden, Nutzende, Entwickelnde sowie

von KI-basierten Entscheidungen Betroffene zu berücksichtigen (Langer et al. 2021). Daher wird mitunter gar von einem menscheitszentrierten Ansatz gesprochen (Dignum und Dignum 2020).

HCAI erfordert die Betrachtung der KI-Anwendung als Endprodukt und aller Prozesse entlang des kompletten Lebenszyklus (Shneiderman 2022; Hyysalo und Johnson 2015). Durch Design-Entscheidungen werden Menschen in die Lage versetzt, die KI-Entwicklung und deren Einsatz zu beeinflussen und Einsichten über die Funktionsweise der KI zu bekommen (Chromik und Butz 2021). Neben Transparenz und dem Zusammenspiel aus Mensch, Technik, und Organisation spielen Vertrauenswürdigkeit, Ethik, soziale Werte und [Nachhaltigkeit](#) eine zentrale Rolle bei der Entwicklung und Nutzung menschenzentrierter KI (Shneiderman 2022; Zinke-Wehlmann et al. 2022; Auernhammer 2020; Willhelm et al. 2021).

Häufig wird das Konzept der HCAI daher mit [erklärbarer KI \(explainable AI, XAI\)](#), ethischer KI (ethical AI) und – im industriellen Kontext – mit [Industrie 5.0](#) in Verbindung gebracht (Vainio-Pekka et al. 2023; Ehsan et al. 2021). Die konzeptionellen Zusammenhänge sind allerdings noch nicht genau spezifiziert. Mitunter wird auch eine Integration dieser Konzepte gefordert (Garibay et al. 2023).

Quellen

- Auernhammer J (2020) Human-centered AI: The role of Human-centered Design Research in the development of AI. In: Boess S, Cheung M, Cain R (Hrsg.) Synergy – DRS International Conference 2020, 11-14 August. Human-centered AI: The role of Human-centered Design Research in the development of AI
- Chromik M, Butz A (2021) Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces. In: Ardito C, Lanzilotti R, Malizia A, Petrie H, Piccinno A, Desolda G, Inkpen K (Hrsg.) Human-Computer Interaction – INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30 – September 3, 2021, Proceedings, Part II. Springer. S. 619–640
- Dignum F, Dignum V (2020) How to center AI on humans. In: Alessandro S, Luciano S, Paul L (Hrsg.) NeHuAI 2020. First International Workshop on New Foundations for Human-Centered AI: Proceedings of the First International Workshop on New Foundations for Human-Centered AI (NeHuAI)co-located with 24th European Conference on Artificial Intelligence (ECAI 2020). CEUR workshop proceedings. S. 59-62
- Ehsan U, Wintersberger P, Liao QV, Mara M, Streit M, Wachter S, Riener A, Riedl MO (2021) Operationalizing Human-Centered Perspectives in Explainable AI. In: Kitamura Y, Quigley A, Isbister K, Igarashi T (Hrsg.) CHI EA '21. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. Yokohama Japan. CHI Extended Abstracts. S. 1–6. doi: 10.1145/3411763
- Garibay OO, Winslow B, Andolina S, Antona M, Bodenschatz A, Coursaris C, Falco G, Fiore SM, Garibay I, Grieman K, Havens JC, Jirotko M, Kacorri H, Karwowski W, Kider J, Konstan J, Koon S, Lopez-Gonzalez M, Maifeld-Carucci L, McGregor S, Salvendy G, Shneiderman B, Stephanidis C, Strobel C, Holter CT, Xu W (2023) Six Human-Centered Artificial Intelligence Grand Challenges. International Journal of Human-Computer Interaction 39(3):391-437. doi: 10.1080/10447318.2022.2153320
- Hyysalo S, Johnson M (2015) The user as relational entity. Options that deeper insight into user representations opens for human-centered design. Information Technology & People 28(1): 72–89. doi: 10.1108/ITP-01-2014-0011
- International Organization for Standardization (2019) Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. ISO-9241-210-2019.pdf (iteh.ai)
- Langer M, Oster D, Speith T, Hermanns H, Kästner L, Schmidt E, Sasing Am Baum K (2021) What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. Artificial Intelligence 296: 103473. doi: 10.1016/j.artint.2021.103473
- Shneiderman B (2022). Human-Centered AI. Oxford: Oxford University Press.
- Vainio-Pekka H, Agbese MOO, Jantunen M, Vakkuri V, Mikkonen T, Rousi R, Abrahamsson P (2023) The Role of Explainable AI in the Research Field of AI Ethics. ACM Trans. Interact. Intell. Syst. 13(4)(26): 26-39. doi: 10.1145/3599974
- Wilhelm B, Oliver R, Thomas R, Matthias P (Hrsg.) (2021) Menschzentrierte KI-Anwendungen in der industriellen Produktion. Fraunhofer IAO, Stuttgart
- Zinke-Wehlmann C, Friedrich J, Kirschenbaum A, Wölke M, Brückner A (2022) Conceptualizing Sustainable Artificial Intelligence Development. In: Camarinha-Matos LM, Angel O, Xavier B, Osório AL (Hrsg.) Collaborative Networks in Digitalization and Society 5.0. Springer. S. 545–554

Autor:innen

- Dr. Sandra Schumann (Wirtschafts- und Medizininformatik / Physik, Institut für Angewandte Informatik (InfAI), KMI)
- Dr. Tobias Kopp (Mensch-Technik-Interaktion, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Mitbestimmung/ Partizipation

Definition

Unter Mitbestimmung in der Arbeitswelt wird verstanden, dass Beschäftigte – oder ihre betrieblichen Interessensvertretungen – über die Ausgestaltung von Arbeitsprozessen mitentscheiden. Dies umfasst auch und insbesondere Fragen der Technikeinführung und -nutzung. Von einem weiten Verständnis von Partizipation unterscheidet sich Mitbestimmung durch ihre rechtliche Verankerung, etwa in Form des Betriebsverfassungsgesetzes in der Privatwirtschaft oder des Personalvertretungsgesetzes im Öffentlichen Dienst, und damit durch ihren verbindlichen Charakter. Dieser besteht aber weitgehend in Form von Einspruchsrechten, etwa in Fragen des Beschäftigten-datenschutzes, der beruflichen Weiterbildung oder des Gesundheitsschutzes (Klebe und Wenckebach 2020). Als Reaktion auf die besonderen Gestaltungsbedarfe und -Herausforderungen im Bereich [Künstliche Intelligenz \(KI\)](#) wurden im Betriebsräte-modernisierungsgesetz von 2021 die Hinzuziehung von Sachverständigen durch den Betriebsrat erleichtert und die Notwendigkeit einer robusten Mitbestimmung im Hinblick auf die Einführung von KI im Betrieb unterstrichen (BetrVG § 80 Absatz 3).

Gerade im Bereich der Einführung von KI kann eine frühzeitige Einbindung der Beschäftigten und ihrer Interessensvertretungen dazu führen, das Beschäftigtenwissen als Innovationsressource zu erschließen. Dies stellt eine größere Legitimität des Innovationsprozesses und damit dessen Erfolg sicher und vermeidet spätere Fehlentwicklungen (zu den epistemischen Vorzügen von demokratischer Mitbestimmung im Betrieb siehe Gerlsbeck und Herzog 2020). Für die gemeinschaftliche Organisation von Prozessen der Technikentwicklung bieten sich Formen erweiterter Mitbestimmung

an, die beispielsweise in Betriebsvereinbarungen auf einzelbetrieblicher Ebene geregelt werden und bis hin zu Vetorechten der betrieblichen Interessensvertretungen reichen können, um eine Vertrauensgrundlage im Sinne einer konstruktiven Zusammenarbeit zu legen (Schröter 2019).

Quellen

- Bundesministerium der Justiz und für Betriebsverfassungsgesetz (BetrVG). BetrVG - nichtamtliches Inhaltsverzeichnis (gesetze-im-internet.de) Zugriffen: 16. Januar 2024
- Bundesministerium der Justiz und für Bundespersonalvertretungsgesetz (BPersVG). BPersVG - Bundespersonalvertretungsgesetz (gesetze-im-internet.de) Zugriffen: 16. Januar 2024
- Gerlsbeck F, Herzog L (2020) The epistemic potentials of workplace democracy. *Review of Social Economy* 78 (3), 307–330. doi:10.1080/00346764.2019.1596299
- Klebe T, Wenckebach J (2020) Künstliche Intelligenz – Handlungsfeld für betriebliche Mitbestimmung und Arbeitsrechtsregulierung. In: Verena B, Stephan K (Hrsg.) *Arbeit in der Data Society. Zukunftsfähige Unternehmensführung in Forschung und Praxis*. Springer, Wiesbaden. S. 331–350.
- Schröter W (2019) *Der mitbestimmte Algorithmus: Gestaltungskompetenz für den Wandel der Arbeit: Beiträge zum 65. Geburtstag eines Netzwerkes*. Talheimer Verlag, Mössingen-Talheim.

Autor

- Philipp Frey, M.A. (Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Moderierter Spezifikationsdialog

Definition

Das Format Moderierter Spezifikationsdialog stellt einen Anwendungsbaustein für sozialpartnerschaftliche Prozesse der Anpassung algorithmischer Steuerungs- und Entscheidungssysteme an die Anforderungen von Arbeits-, Geschäfts- und Verwaltungsvorgängen dar. Diese Dialogform bildet Ort und Zeitpunkt der Anwendung von ethischen, rechtlichen und sozialen Gestaltungskriterien zur Vorbereitung der Einführung von sich selbst verändernden Delegationssystemen (Schröter 2022). Das Format wird vor der Beschaffung und vor der Implementierung von [↗ Delegations-technik](#) als Aushandlungsvorgang zwischen dem Arbeitgeber, dem IT-Team und der Beschäftigtenvertretung realisiert. Er wird extern moderiert. Kennzeichnend für den Moderierten Spezifikationsdialog ist die Gleichgewichtung von arbeitsweltlichem Erfahrungswissen und dem IT-Fach- und Sachwissen. Die Sozialpartner der Landeshauptstadt Stuttgart – der öffentliche Arbeitgeber und der Gesamtpersonalrat – haben in ihrer Rahmendienstvereinbarung 2022 dieses Format rechtlich formalisiert und verankert. Dafür erhielten die Stuttgarter Partner den Personalrätepreis 2022 in Silber. Das Format Moderierter Spezifikationsdialog wurde in die Deutsche Normungsroadmap Künstliche Intelligenz Ausgabe 2 (Wahlster 2022) aufgenommen. Das Format wurde vom Forum Soziale Technikgestaltung entwickelt.

Siehe auch [↗ Der mitbestimmte Algorithmus](#)

Quellen

Schröter W (2022) Wie gestalten wir die sogenannte „Künstliche Intelligenz“?

<http://www.blog-zukunft-der-arbeit.de/wie-gestalten-wir-die-sogenannte-kuenstliche-intelligenz-einladung-zur-mitwirkung/> Zugegriffen: 5. Oktober 2023

Wahlster W, Winterhalter C (Hrsg.) (2022) Deutsche Normungsroadmap.

Künstliche Intelligenz. Ausgabe 2. DIN e. V., Berlin

Autor

- Welf Schröter (Arbeitsgestaltung, Forum Soziale Technikgestaltung, KARL)

Nachhaltigkeit (in der KI)

Definition

Mit dem Fokus auf eine nachhaltige Künstliche Intelligenz (KI) wird eine Neuausrichtung des gesamten Lebenszyklus von KI-Produkten (sprich: Ideenfindung, Training, Anpassung, Umsetzung und Verwaltung) im Sinne des Konzeptes einer nachhaltigen Entwicklung gefördert. Das Konzept der Nachhaltigkeit fußt hierbei auf der Berücksichtigung dreier Dimensionen: Soziales/ Gesellschaft, Ökologie und Ökonomie. Es geht darum, bereits in der Entwicklung von KI Nachhaltigkeitskriterien bei der Gestaltung zu inkludieren (Van Wynsberghe 2021). Im ökologischen Sinne zielt nachhaltige KI unter anderem auf eine Reduktion von Kohlenstoffemissionen durch die Verwendung von zertifizierter ressourcenschonender Hardware ab. Ein weiterer Aspekt der Energieeffizienz lässt sich durch die Verwendung sparsamer [KI-Methoden](#), die mit weniger Rechenleistung als beispielsweise Neuronale Netze auskommen, erreichen. Auf sozialer Ebene bedeutet nachhaltige KI einerseits, die [Nachvollziehbarkeit](#) der Entscheidungsfindung der KI zu gewährleisten. Andererseits muss die Verantwortungsübernahme bei der KI sichergestellt sein. Außerdem ist darauf zu achten, Diskriminierung zu unterbinden oder Fairness zu gewährleisten. Dies betrifft die KI einerseits in der Anwendung, wenn es um die Verwendung personalisierter Daten geht. Andererseits auch in der Entwicklung (beispielsweise beim Labeln von Daten), da hier darauf geachtet werden muss, dass Arbeitskräfte nicht benachteiligt oder sogar

ausgebeutet werden. Auf Seiten der ökonomischen Nachhaltigkeit ist darauf zu achten, dass die Marktvielfalt sowie die Ausschöpfung von Innovationspotenzialen gegeben sind. Außerdem müssen eventuelle Auswirkungen auf den Arbeitsmarkt berücksichtigt werden, dies betrifft mitunter die Sicherstellung guter Arbeitsbedingungen sowie die Entgegenwirkung von möglichem Stellenabbau (Rohde et al. 2021).

Quellen

Rohde F, Wagner J, Reinhard P, Petschow U, Mayer A, Voss M, Mollen A (2021)

Nachhaltigkeitskriterien für Künstliche Intelligenz. Entwicklung eines Kriterien- und Indikatorensets für die Nachhaltigkeitsbewertung von KI-Systemen entlang des Lebenszyklus,

Schriftenreihe des IÖW 220/2021. Institut für ökologische Wirtschaftsforschung GmbH, Berlin

Van Wynsberghe A (2021) Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 1: 213-218. doi: 10.1007/s43681-021-00043-6

Autorin

- Mandy Wölke (Human-Computer Interaction und Innovationswissenschaften, Institut für Angewandte Informatik (InfAI), KMI)

Nachvollziehbare KI (understandable, comprehensible, intelligible artificial intelligence)

Definition

Ist es für Menschen möglich, die zugrundeliegende Logik der Entscheidungsfindung einer Künstlichen Intelligenz (KI) zu verstehen, so spricht man von nachvollziehbarer KI (Doran et al. 2017; Mohseni et al. 2020). Neben den Eingabedaten und KI-Ausgabe werden dafür weitere Informationen benötigt (Doran et al. 2017): entweder die Funktionen und Prozesse eines direkt interpretierbaren KI-Modells oder zusätzlich erzeugte Erklärungen, falls das Modell eine zu hohe Komplexität aufweist (Meske et al. 2022; Mohseni et al. 2021). → [Interpretierbare](#) oder → [erklärbare KI](#) sind somit eine Vorbedingung für nachvollziehbare KI (Schwalbe und Finzel 2023). Diese Beziehung sowie das Verhältnis zum verwandten Begriff → [transparentes KI-System](#) werden in Abbildung 3 verdeutlicht. Zum Verstehen gehört allerdings auch die Fähigkeit eines Menschen, den Kontext und die Zusammenhänge eines Problems zu erkennen (Schwalbe und Finzel 2023). Es hängt daher vom einzelnen Menschen und dessen Vorwissen ab, ob er die zusätzlichen Informationen verstehen und die KI somit nachvollziehen kann (Doran et al. 2017).

Eine nachvollziehbare KI befähigt den Endnutzenden, die KI-Entscheidungen auf ihre Sinnhaftigkeit und Konsistenz zu prüfen und im Zweifel anzufechten (Antoniadi et al. 2021; Gunning und Aha 2019). Dies kann zu einer Stärkung des Vertrauens und der → [Akzeptanz](#) in die KI und zu einer Effizienz- und Produktionssteigerung in

KI-gestützten Digitalisierungs- und Arbeitsprozessen führen (Gunning et al. 2019; Gunning und Aha 2019; IBM 2023). Nachvollziehbare KI kann jedoch nicht nur die Implementierung von KI im Organisationskontext positiv beeinflussen, sondern auch die rechtskonforme Umsetzung transparenter KI-Systeme ermöglichen (Gerlings et al. 2021; Goodman und Flaxman 2017).



Abbildung 3: Abgrenzung zu den Begriffen transparentes KI-System, interpretierbare KI und erklärbare KI, Quelle: eigene Darstellung

Quellen

- Antoniadi A, Du Y, Guendouz Y, Wei L, Mazo C, Becker BA, Mooney C (2021) Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review. *Applied Sciences* 11(11), 5088. doi:10.3390/app11115088
- Doran D, Schulz S, Besold TR (2017) What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *ArXiv*. doi:10.48550/arXiv.1710.00794
- Gerlings J, Shollo A, Constantiou I (2021) Reviewing the Need for Explainable Artificial Intelligence (xAI). *Proceedings of the 54th Hawaii International Conference on System Sciences*:1284–1293. doi: 10.24251/HICSS.2021.156
- Goodman B, Flaxman S (2017) European union regulations on algorithmic decision making and a „right to explanation“. *AI Magazine* 38(3):50–57. doi:10.1609/aimag.v38i3.2741
- Gunning D, Aha DW (2019) DARPA's Explainable Artificial Intelligence Program. *AIMag* 40(2):44–58. doi:10.1609/aimag.v40i2.2850
- Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang GZ (2019) XAI-Explainable artificial intelligence. *Sci Robot* 4(37). doi: 10.1126/scirobotics.aay7120
- IBM (2023). What is explainable AI? | IBM Zugegriffen: 22. September 2023

- Meske C, Bunde E, Schneider J, Gersch M (2022) Explainable artificial intelligence: Objectives, stakeholders, and future research opportunities. *Information Systems Management* 39(1):53–63. doi:10.1080/10580530.2020.1849465
- Mohseni S, Zarei N, Ragan ED (2020) A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Trans. Interact. Intell. Syst.* 1:1–45. doi: 10.1145/3387166
- Schwalbe G, Finzel B (2023) A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Min Knowl Disc* 37:1–59. doi: 10.1007/s10618-022-00867-8

Autor

- Robin Weitemeyer, M.Sc. (Informatik, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

Psychische Belastung und Beanspruchung

Definition

Ergonomische Grundlagen psychischer Belastung sind Inhalt der DIN EN ISO 10075-1 (2017), welche auf dem Belastungs- und Beanspruchungsmodell aufbaut (Rohmert 1984). Psychische Belastung kann dabei als die Gesamtheit aller feststellbaren Einflüsse beschrieben werden, die auf den arbeitenden Menschen geistig, sprich kognitiv und emotional, einwirken. Die psychische Beanspruchung hingegen ist die unmittelbare Auswirkung der psychischen Belastung innerhalb des Individuums in Abhängigkeit der aktuellen Verfassung (Neuner 2021; Kirchner 1986). Psychische Belastungen beschreiben Faktoren, die sich psychisch auf die arbeitende Person auswirken, sie entstehen unter anderem durch die Arbeitsaufgabe, die [Arbeitsgestaltung](#) und -organisation aber auch durch soziale Gegebenheiten, beispielsweise soziale Kontakte am Arbeitsplatz (Neuner 2021). Die Auswirkungen der Belastung können grundsätzlich sowohl positiv als auch negativ sein. Allerdings können insbesondere bei lang andauernder hoher Beanspruchung negative Auswirkungen wie arbeitsbedingter Stress, Burnout und Depressionen auftreten (DIN EN ISO 10075-1 2017).

Durch die rasant fortschreitende Automatisierung und Digitalisierung sind die aktuellen Arbeitsbedingungen einem dynamischen Wandel unterworfen – wobei der Wandel der Arbeitswelt eine Zunahme psychischer Belastungen am Arbeitsplatz mit sich

bringt. Zum einen nehmen psychische Belastungen bei der Arbeit zu, zum Beispiel durch Arbeitsverdichtung, beschleunigte Kommunikation, Entgrenzung und ständige Erreichbarkeit. Andererseits verändern sich die eigentlichen Arbeitsaufgaben, von ehemals eher körperlichen Tätigkeiten hin zu zunehmend kognitiven und informationellen Arbeitsanforderungen (Müller-Thur et al. 2018; Süß et al. 2022). Dieser Wandel zur digitalen Arbeit ist zunehmend auch im industriellen Umfeld zu beobachten. Denn auch dieses ist zunehmend durch intelligente und vernetzte technische Systeme geprägt, die hochautomatisierte physische Prozesse und digitale Informationen integrieren (Cascio und Montealegre 2016).

Um psychisch beeinträchtigende Auswirkungen zu vermeiden und förderliche Auswirkungen zu unterstützen, ist es daher notwendig, das Arbeitssystem auf die arbeitende Person abzustimmen. Der DIN EN ISO 10075-2 (2023) lassen sich einige Gestaltungsmaßnahmen entnehmen, die eine psychisch gesunde Gestaltung der Arbeitsumgebung fördern. Dazu gehören die Beeinflussung der Belastungsintensität und der Belastungsdauer in Bezug auf die primären Arbeitsaufgaben, aber auch Maßnahmen auf organisatorischer und sozialer Ebene.

Quellen

Cascio WF, Montealegre R (2016) How technology is changing work and organizations. Annual review of organizational psychology and organizational behavior 3:349–375. doi:10.1146/annurev-orgpsych-041015-062352

DIN EN ISO 10075-1 (2017) Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung – Teil 2: Gestaltungsgrundsätze. doi: 10.31030/3435527

DIN EN ISO 10075-2 (2023) Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung – Teil 2: Gestaltungsgrundsätze. doi: 10.31030/3435527

Kirchner JH (1986) Belastungen und Beanspruchungen — Einige begriffliche Klärungen zum Belastungs-Beanspruchungs-Konzept. In: Hackstein R, Heeg FJ, von Below F (Hrsg.) Arbeitsorganisation und Neue Technologien: Impulse für eine weitere Integration der traditionellen arbeitswissenschaftlichen Entwicklungsbereiche. Springer, Berlin, Heidelberg, S. 553–569

Müller-Thur K, Angerer P, Körner U, Dragano N (2018) Arbeit mit digitalen Technologien, psychosoziale Belastungen und potenzielle gesundheitliche Konsequenzen. Zeitschrift für medizinische Prävention (ASU) 53:388–391.

Nachreiner F, Schütte M (2018) Ergonomische Grundlagen bezüglich psychischer Arbeitsbelastung – Teil 1: Allgemeine Aspekte und Konzepte und Begriffe. Zeitschrift für Arbeits- und Organisationspsychologie A&O 62(3): 166. doi: 10.1026/0932-4089/a000276

Neuner R (2021) Psychische Gesundheit bei der Arbeit. Gefährdungsbeurteilung und gesunde Organisationsentwicklung. 4. Aufl. Springer, Wiesbaden

Rohmert W (1984) Das Belastungs-Beanspruchungs-Konzept. Zeitschrift für Arbeitswissenschaft 38:193–200

Süß S, Dragano N, Klingenberg I (2022) Belastungen durch die digitale Arbeit. In: Roth S, Corsten H (Hrsg.) Handbuch Digitalisierung. Franz Vahlen, München, S. 273–290

Autor:innen

- Vera Rick, M.A. (Psychologie, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Prof. Dr.-Ing. Dr. rer. medic. Alexander Mertens (Arbeitswissenschaft, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)
- Univ.-Prof. Dr.-Ing. Verena Nitsch (Arbeitswissenschaft, Institut für Arbeitswissenschaft der RWTH Aachen University, AKzentE4.0)

Reliabilität (reliability)

Definition

Reliabilität beschreibt die Verlässlichkeit von einheitlichen Prozessen oder Algorithmen, bezogen auf verschiedene Aspekte der Überprüfbarkeit des gesamten Arbeitsprozesses. Viele Faktoren beeinflussen die Aussagekraft von zum Beispiel [machine learning-Algorithmen](#), die für das Trainieren eines [KI-Systems](#) verwendet werden. Reproduzierbarkeit, Genauigkeit (accuracy), Limitation, [Transparenz](#) in der Entwicklung, ethische Fairness, sind einige wichtige Faktoren, die beschreiben können, wie verlässlich und vor allem wie vergleichbar ein Prozess abläuft (Balagurunathan et al. 2021).

Auch kann sich Reliabilität auf die Ausfallsicherheit eines Prozesses oder eines Systems beziehen. Aus einer qualitativen Sicht kann die Reliabilität auch als Funktionsfähigkeit eines Prozesses definiert werden und aus quantitativer Sicht gibt sie die Wahrscheinlichkeit für mögliche Betriebsunterbrechung innerhalb eines festen Zeitrahmens an (Birolini 2017).

Dabei ist der Schutz vor sowohl äußeren Einflüssen als auch vor internen Ausfällen eines Systems zu gewährleisten. Ausfälle könnten dafür sorgen, dass der Prozess deutlich verlangsamt wird, keine korrekten Ergebnisse liefert oder komplett zum Erliegen kommt. Auch die Akzeptanz zum Beispiel eines KI-Systems hängt stark von seiner Verlässlichkeit hinsichtlich Ausfallsicherheit und Performance ab (Balagurunathan et al. 2021; Ryan 2020).

Quellen

- Balagurunathan Y, Mitchell R, El Naqa I (2021) Requirements and reliability of AI in the medical context. *Physica medica (PM)* 83:72–78. doi: 10.1016/j.ejmp.2021.02.024
- Birolini A (2017) *Reliability engineering. Theory and practice*. 8. Aufl. Springer, Berlin, Heidelberg
- Ryan M (2020) In AI We Trust: Ethics, Artificial Intelligence, and Reliability. *Sci Eng Ethics* 26(5):2749–2767. doi: 10.1007/s11948-020-00228-

Autorin

- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)

Selbstbestimmung

Definition

Selbstbestimmung im Kontext der [Arbeitsforschung](#) und [-gestaltung](#) ist ein Aspekt der [Humanisierung](#) und Subjektivierung von Arbeit, der sowohl auf Individuen als auch auf Gruppen – etwa teilautonome Arbeitsgruppen – bezogen wird. Sie stellt zudem eine normative Prämisse dar: Der Mensch wird als grundsätzlich zur Selbstbestimmung fähig definiert (Ethikrat 2023). Empirisch fasst Selbstbestimmung die Freiheitsgrade, über die Individuen oder Gruppen innerhalb einer [Arbeitsorganisation](#) verfügen. Sie betrifft insbesondere die Entscheidung über und Gestaltung von Zielen und Instrumenten der eigenen Arbeit in quantitativer wie qualitativer Hinsicht (Schlick et al. 2018, S. 690).

Die Selbstbestimmungstheorie (Deci und Ryan 2008) unterscheidet unter anderem extrinsische Motivation, die durch äußere Faktoren (etwa leistungsabhängige Entgeltanteile) induziert wird, von intrinsischer Motivation, die durch innere Faktoren des Menschen (wie inhaltliches Interesse an einer Aufgabe, Selbstverwirklichung) entsteht. Soziologisch und philosophisch lässt sich Selbstbestimmung als [Autonomie](#) von Gruppen oder Individuen als etwas fassen, das nicht etwa von Individuen erlangt, sondern von einer Arbeitsorganisation oder einem Staat gewährt wird. Beide Lesarten verweisen darauf, dass Selbstbestimmung immer sowohl von individuellen als auch sozialen – auf die Arbeitswelt bezogen: organisatorischen – Faktoren abhängt.

Im Kontext von Künstlicher Intelligenz (KI) umreißt der deutsche Ethikrat Selbstbestimmung als ein zu schützendes demokratisches Gut. Er verweist auf das Recht zur informationellen Selbstbestimmung, die Notwendigkeit, über den Einsatz von KI selbstbestimmt entscheiden zu können, sowie die Notwendigkeit, sich digitalen Plattformen selbstbestimmt entziehen zu können. Auch diese vom Ethikrat angesprochenen Aspekte treffen auf soziale Bedingungen, die mit Blick darauf, diese Selbstbestimmung zumindest in Teilen zu ermöglichen, mittels Arbeitsorganisation gestaltbar sind.

Quellen

- Deci EL, Ryan RM (2008): Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Canadian Psychology* 49 (3):182–185 doi:10.1037/a0012801
- Deutscher Ethikrat (2023) Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. STELLUNGNAHME. [stellungnahme-mensch-und-maschine.pdf](#) (ethikrat.org) Zugegriffen: 09. Januar 2024
- Schlick C, Bruder R, Luczak H (2018) *Arbeitswissenschaft*. 4. Aufl. Springer, Wiesbaden

Autor

- Dr. Christian Bernhard-Skala (Erziehungswissenschaft, Deutsches Institut für Erwachsenenbildung Leibniz-Zentrum für Lebenslanges Lernen e. V., Projekt CoCo: Connect & Collect)

Sprachmodelle (wie etwa ChatGPT)

Definition

Ein Sprachmodell basiert auf statistischen Modellen, die auf der Analyse von großen Mengen an Textdaten beruhen, und ist in der Lage, automatisch menschenähnlichen Text zu erzeugen oder Texte zu analysieren, um automatisch Antworten zu generieren oder Schlussfolgerungen zu ziehen (Li et al. 2021). Dabei lernen Sprachmodelle, aufeinander folgenden Wörtern Wahrscheinlichkeiten zuzuordnen, die auf statistischen Mustern in der Verteilung von Wörtern in der Sprache beruhen (Li 2022). Sprachmodelle wurden in den letzten Jahren dank der Fortschritte in der [Künstlichen Intelligenz](#) und insbesondere im Bereich des [maschinellen Lernens](#) erheblich leistungsfähiger. Heute sind sie in der Lage, natürliche Sprache in vielen Anwendungsbereichen zu verstehen und zu generieren, einschließlich Chatbots, Übersetzung und Zusammenfassung von Texten, automatischer Textkorrektur, Beantworten von Fragen und vielem mehr. Ein Sprachmodell besteht aus mehreren Komponenten, darunter eine Eingabe, die typischerweise eine Abfolge von Wörtern oder Buchstaben darstellt, und eine Ausgabe, die eine Vorhersage für das nächste Wort oder die nächste Abfolge von Wörtern darstellt. Der Kern eines Sprachmodells ist der Algorithmus, der auf Basis von Trainingsdaten und der Verarbeitung von Sprachstrukturen wie Grammatik, Syntax und Semantik arbeitet, um Vorhersagen zu erstellen und Texte zu generieren oder zu analysieren.

Quellen

- Li J, Tang T, Zhao WX, Wei Z, Yuan NJ, Wen JR (2021) Pretrained Language Model for Text Generation: A Survey. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI). 4492-4499. doi: 10.24963/ijcai.2021/612
- Li H (2022) Language models: past, present, and future. Communications of the ACM, 65(7), 56-63. doi: 10.1145/3490443
- Trott S, Jones C, Chang T, Michaelov J, Bergen B (2023) Do Large Language Models Know What Humans Know? Cognitive Science 47: e13309. doi:10.1111/cogs.13309

Autorin

- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)

Sustainable Development Goals (SDGs)

Definition

Die Sustainable Development Goals (SDGs) (United Nations 2016) stellen eine Sammlung von 17 Zielen für nachhaltige Entwicklung dar, auf die sich die internationale Staatengemeinschaft im Rahmen der Vereinten Nationen verständigt hat (Abbildung 4). Sie sind Ausdruck eines umfassenden Nachhaltigkeitsverständnisses und umfassen beispielsweise die Umstellung auf bezahlbare und saubere Energie und Maßnahmen zum Klimaschutz gleichermaßen wie Armuts- und Hungerbekämpfung und gelten für den Globalen Süden wie Industrieländer. Das Nachhaltigkeitsverständnis, das die Formulierung der SDG orientiert hat, lässt sich auf den Einsatz von [KI-Systemen](#) operationalisieren. Dabei rücken sowohl der Anwendungszweck der Künstlichen Intelligenz (KI) als auch Fragen der Energie- und Ressourcennutzung und soziale Implikationen, etwa mögliche Qualifizierungs- und Beschäftigungseffekte, in den Blick.

ZIELE FÜR NACHHALTIGE ENTWICKLUNG



Abbildung 4: Die 17 SDGs mit ihren einzelnen eigenen Logos, Quelle: https://de.wikipedia.org/wiki/Ziele_f%C3%BCr_nachhaltige_Entwicklung#/media/Datei:Sustainable_Development_Goals_de.svg

Das Thema [Nachhaltigkeit](#) und die Nachhaltigkeitsziele stehen in engem Zusammenhang mit der [menschengerechten Arbeitsgestaltung](#) und damit auch mit der [Arbeitswissenschaft](#) als wissenschaftliche Disziplin. Denn eine menschengerechte Arbeitsgestaltung ist eine wichtige Voraussetzung für die Erreichung der Nachhaltigkeitsziele, beispielweise in der Schaffung einer gesunden und schadungsfreien Arbeitsumgebung, die sowohl wirtschaftlich als auch menschengerecht ist (Barth et al. 2016; Sträter 2022)

Weitere Informationen zu den SDGs:

➔ https://sdg-portal.de/de/ueber-das-projekt/17-ziele?gclid=EAlaIqobChMlwY6tvs6lgAMV1YZoCROUOAuPEAAAYiAAEgl_KPD_BwE

Weitere Informationen zu Nachhaltigkeitskriterien für Künstliche Intelligenz:

➔ https://www.ioew.de/fileadmin/user_upload/BILDER_und_Downloaddateien/Publikationen/2021/IOEW_SR_220_Nachhaltigkeitskriterien_fuer_Kuenstliche_Intelligenz.pdf

Quellen

- Barth T, Jochum G, Littig B (Hrsg.) (2016) Nachhaltige Arbeit: Soziologische Beiträge zur Neubestimmung der gesellschaftlichen Naturverhältnisse. Campus Frankfurt / New York
- Sträter, O (2022) Die Rolle und Bedeutung der Arbeitswissenschaft für die beginnende Arbeit an der Nachhaltigkeit. Eine erste Bestandsaufnahme wichtiger Themen. GfA Frühjahrskongress 2022.
- United Nations (2016) The 17 Goals. THE 17 GOALS | Sustainable Development (un.org) Zugriffen:12 July 2023

Autor

- Philipp Frey, M.A. (Technikfolgenabschätzung, Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am Karlsruher Institut für Technologie (KIT), KARL)

Transparente KI/Transparentes KI-System

Definition

Eine zentrale Anforderung an ein transparentes KI-System in der Arbeitswelt ist die Nachvollziehbarkeit der Entscheidungsprozesse des verwendeten KI-Modells (Antoniadi et al. 2021; Mohseni et al. 2021; Schwalbe und Finzel 2023). Abbildung 5 stellt diesen Zusammenhang eines transparenten KI-Systems und der [nachvollziehbaren KI](#) sowie den verwandten Begriffen [interpretierbarer KI](#) und [erklärbarer KI](#) dar. Viele KI-Modelle stellen aufgrund ihrer hohen Komplexität jedoch für den Menschen eine Black Box dar, da die Funktionen und Logik, nach der die KI-Ausgaben aus den Eingabedaten erzeugt werden, nicht vom Menschen verstanden werden können (Antoniadi et al. 2021; Doran et al. 2017). Dies kann bei Endnutzenden und betroffenen Mitarbeitenden zu einer höheren Skepsis gegenüber [KI-Systemen](#) führen, sowie in der gesamten Gesellschaft gegenüber der Technologie KI als Ganzes (Eschenbach 2021). Um im Falle einer Black Box die Nachvollziehbarkeit trotzdem zu gewährleisten, müssen weitere Maßnahmen ergriffen werden, indem man die Komplexität des Modells reduziert oder zusätzliche Erklärungen zu dem KI-Verhalten erzeugt. Transparenz stellt jedoch ein ganzheitliches Merkmal an ein KI-basiertes System dar (Antoniadi et al. 2021). Es müssen Informationen zu den Trainingsdaten, dem Trainingsprozess, der Verarbeitung der Eingabedaten, den Prozessen der Entscheidungsfindung des KI-Modells und der weiteren Nutzung der KI-Ausgaben dokumentiert und bei Bedarf bereitgestellt werden (Antoniadi et al. 2021). Die Transparenz eines KI-Systems stellt somit eine Voraussetzung dar, um den in der Datenschutzgrundverordnung (DSGVO) festgelegten Anforderungen nachzukommen (Goodman und Flaxman 2017). Des Weiteren führt die hochrangige Expertengruppe für Künstliche Intelligenz (AI HLEG) der Europäischen Kommission in ihrer Bewertungsliste für vertrauenswürdige KI ([ALTAI](#)) drei zentrale Elemente der Transparenz an (High-Level Expert Group on Artificial Intelligence 2020):

1. die Rückverfolgbarkeit der Datenverarbeitung und Entscheidungsfindung des KI-Systems,
2. die Erklärbarkeit der KI-gestützten Entscheidungsprozesse für direkt und indirekt betroffene Personen,
3. Kommunikation der Fähigkeiten und Limitierungen des KI-Modells an die Nutzenden und in Kenntnis setzen im Falle einer Interaktion mit der KI.

Diese Liste dient Unternehmen als wichtige Orientierungshilfe, welche rechtlichen Anforderungen an die Transparenz von KI-Systemen in Zukunft im Rahmen der geplanten KI-Verordnung (AI Act) der Europäischen Union (EU) gestellt werden (Europäische Kommission 2021). Diese Verordnung hat das Ziel, den Einsatz bestimmter KI-Anwendungen zu regulieren, die Einhaltung der Rechte von EU-Bürgern sicherzustellen und somit das Vertrauen in diese Technologie zu stärken. Das Offenlegen der für Transparenz relevanten Informationen kann für Unternehmen zwar zu potenziellen Interessenkonflikten mit der Wahrung ihrer Betriebsgeheimnissen führen, jedoch auch dazu beitragen, die Skepsis in der Gesellschaft gegenüber KI zu verringern und somit die Akzeptanz von KI-gestützten Produkten zu erhöhen (Fine Licht und Fine Licht 2020).



Abbildung 5: Abgrenzung zu den Begriffen nachvollziehbare KI, interpretierbare KI und erklärbare KI, Quelle: eigene Darstellung

Quellen

- Antoniadi A, Du Y, Guendouz Y, Wei L, Mazo C, Becker BA, Mooney C (2021) Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review. *Applied Sciences* 11(11), 5088. doi:10.3390/app11115088
- Doran D, Schulz S, Besold TR (2017) What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *ArXiv*. doi:10.48550/arXiv.1710.00794
- Eschenbach WJ von (2021) Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology* 34:1607–1622. doi: 10.1007/s13347-021-00477-0
- Europäische Kommission (2021) 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. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF. Zugegriffen: 29. September 2023
- Fine Licht K de, Fine Licht J de (2020) Artificial intelligence, transparency, and public decision-making. *AI & Soc* 35:917–926. doi: 10.1007/s00146-020-00960-w
- Goodman B, Flaxman S (2017) European Union Regulations on Algorithmic Decision Making and a “Right to Explanation”. *AIMag* 38:50–57. doi: 10.1609/aimag.v38i3.2741

- High-Level Expert Group on Artificial Intelligence (2020) The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>. Zugegriffen: 22. September 2023
- Mohseni S, Zarei N, Ragan ED (2021) A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Transactions on interactive intelligent systems*. 11:1–45. doi:10.1145/3387166
- Schwalbe G, Finzel B (2023) A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Min Knowl Disc* 37:1–59. doi: 10.1007/s10618-022-00867-8

Autor

- Robin Weitemeyer, M.Sc. (Informatik, Institut für Lernen und Innovation in Netzwerken (ILIN) der Hochschule Karlsruhe (HKA), KARL)

UX Design

Definition

Interagieren Menschen mit Elementen ihrer Umwelt, so findet stets ein interner, kontextabhängiger Prozess des Erlebens statt. Dieser lässt sich auch für den Umgang von Nutzenden spezieller Systeme, zum Beispiel Mitarbeitenden in Unternehmen, die mit technischen Systemen unter anderem [KI-Systemen](#) interagieren, übersetzen. Bei den Nutzenden wird, sobald eine Interaktion mit dem System stattfindet, eine Reaktion hervorgerufen. Die Gesamtheit dieser Reaktion im Rahmen der Interaktion wird als User Experience, kurz UX betitelt (Harston und Pyla 2018). Das UX-Design zielt darauf ab, eine User Experience zu schaffen, die als produktiv, erfüllend, oder als erfreulich wahrgenommen werden soll.

Im Gegensatz zur Usability (Gebrauchstauglichkeit oder Benutzerfreundlichkeit), welche das Ausmaß, in dem ein Produkt, System oder Dienst effektiv, effizient und zufriedenstellend zur Zielerreichung verwendet werden kann, bezeichnet die User Experience das Nutzungserlebnis und die Nutzerfreundlichkeit bei der Verwendung.

Faktoren für gutes UX-Design werden als usability, usefulness, emotional impact und meaningfulness beschrieben und umfassen die UX unter Einbezug aller Sinne (Harston und Pyla 2018). Dabei ist zu beachten, dass gutes UX-Design darauf abzielt, eine Erfahrung bei Nutzung einer Software zu schaffen, die als positiv von Nutzenden wahrgenommen wird. Sie sollte mit dem Ansatz des menschenzentrierten Arbeitens

vereinbar sein und durch zum Beispiel Feedback der Nutzenden weiterentwickelt werden können. Für eine gut funktionierende UX spricht, dass die Nutzenden gern und freiwillig mit der Software oder Applikation interagieren. Eine gut entwickelte UX kann somit bei Anwendenden für mehr Motivation, Effizienz und Gesamtzufriedenheit sorgen, ausgelöst durch eine positive interne Reaktion auf das angewandte System (Unger und Chandler 2012). Im Unternehmenskontext kann dies auch zu positiven Auswirkungen auf allgemeine Effizienz wirken, was finanzielle Auswirkungen wie zum Beispiel Gewinnsteigerungen nach sich ziehen kann.

Quellen

- Hartson R, Pyla P (2018) The UX book. Agile UX design for a quality user experience. 2. Aufl. Morgan Kaufmann, Cambridge
- Unger R, Chandler C (2012) A Project Guide to UX Design: For User Experience Designers in the Field or in the Making. 2.Aufl. New Riders, Berkeley

Autorin

- Maria Schwaren (Digital Humanities, Institut für Angewandte Informatik (InfAI), KMI)

Wissensbasiertes System

Definition

Der Wesenskern wissensbasierter Systeme ist, dass sie eine explizite Repräsentation von Wissen besitzen, und dieses Wissen maßgeblich für das Verhalten des Systems ist. Dies wurde etwa im Jahr 1982 als sogenannte KR Hypothese (engl. knowledge representation hypothesis) von Brian Smith formuliert (Smith 1982).

Demnach beinhaltet ein (intelligentes) wissensbasiertes System eine Komponente, die das Wissen des Systems repräsentiert und Komponenten, in denen dieses Wissen das Verhalten des Systems maßgeblich beeinflusst. Wissensbasierte Systeme stellen ein Teilgebiet der [Künstlichen Intelligenz](#) dar. Eine frühe, aber gute Darstellung von wissensbasierten Systemen findet sich in (Davis 1986). Eine aktuelle Übersicht über Methoden und Systeme geben (Beierle und Kern-Isberner 2019).

Einer der Vorzüge wissensbasierter Systeme ist, dass eine einfache Übertragung eines Systems auf einen neuen Anwendungsfall oder eine neue Domäne möglich ist. Dazu muss lediglich das Wissen im alten System durch das Wissen der neuen Domäne ausgetauscht werden. Die restlichen Systemkomponenten können unverändert bleiben. Außerdem kann durch die explizite Repräsentation des Wissens oft einfacher eine Form der Erklärbarkeit hergestellt werden (Meske et al. 2022). Die ist auch angesichts der Datenschutzgrundverordnung (Goodman und Flaxman 2017) und der KI-Regulierung der EU von zunehmender Bedeutung.

Quellen

- Beierle C, Kern-Isberner G (2019) Computational Intelligence. Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen, 6. Aufl. Springer Vieweg, Wiesbaden
- Davis R (1986) Knowledge-based systems. Science 231(4741):957–963. doi: 10.1126/science.231.4741.957
- Goodman B, Flaxman S (2017) European union regulations on algorithmic decision making and a „right to explanation“. AI Magazine 38(3):50–57. doi:10.1609/aimag.v38i3.2741
- Meske C, Bunde E, Schneider J, Gersch M (2022) Explainable artificial intelligence: Objectives, stakeholders, and future research opportunities. Information Systems Management 39(1):53–63. doi:10.1080/10580530.2020.1849465
- Smith BC (1982) Reflection and semantics in a procedural language. Ph.D. Thesis. M.I.T., Computer Science lab. MIT

Autoren

- Dr. Stefan Schiffer (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)
- Prof. Dr. Alexander Ferrein (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)

Wissensgraphen

Definition

Ein Wissensgraph (engl. Knowledge Graph) (Hogan et al. 2020) ist eine semantische Netzwerkstruktur, die verwendet wird, um Informationen und Beziehungen zwischen verschiedenen Datenpunkten zu organisieren und darzustellen. Der Wissensgraph wird häufig in der [Künstlichen Intelligenz \(KI\)](#) und der Datenanalyse eingesetzt, um große Informationsmengen effizient zu verwalten und die Suche nach Informationen zu erleichtern. Im Kern besteht ein Wissensgraph aus einer Sammlung von Knoten und Kanten, die miteinander verbunden sind. Die Knoten repräsentieren dabei entweder konkrete Entitäten wie Personen, Orte oder Objekte oder abstrakte Konzepte wie Ideen oder Kategorien. Die Kanten zeigen wiederum Beziehungen zwischen den Knoten auf, zum Beispiel, dass eine Person ein bestimmtes Objekt besitzt oder dass ein bestimmter Ort sich in einem bestimmten Land befindet. Diese Fakten können auch als sogenannte (Subjekt-, Prädikat-, Objekt-)Tripel abgespeichert werden, zum Beispiel (Berlin, ist_Hauptstadt_von, Deutschland). Eingesetzt werden Wissensgraphen in vielen Bereichen vom Personalwesen über das Modellieren von Zusammenhängen bei Banken und Versicherungen bis zum Digitalen Zwilling in der Entwicklung und Fertigung.

Quelle

Hogan A, Blomqvist E, Cochez M, d'Amato C, de Melo G, Gutierrez C, Gayo JEL, Kirrane S, Neumaier S, Polleres A, Navigli R, Ngomo AN, Rashid SM, Rula A, Schmelzeisen L, Sequeda J, Staab S, Zimmermann A (2020). Knowledge graphs. *ACM Computing Surveys*, 54(4), 1-37. doi: 10.1145/3447772

Autor

- Prof. Dr. Ralf Krestel (Informatik, Leibniz-Informationszentrum Wirtschaft ZBW, Projekt CoCo: Connect & Collect)

Wissensrepräsentation und Schlussfolgern

Definition

Wissensrepräsentation und Schlussfolgern ist ein Teilgebiet der [Künstlichen Intelligenz](#), das sich mit der formalen Repräsentation von Wissen und der automatischen Berechnung von neuem Wissen befasst. Ziel dabei ist, das Wissen in einer für Maschinen (und meist auch Menschen) interpretierbaren Form zu hinterlegen. Für die Repräsentation werden häufig (formale) Sprachen verwendet, oft auch formale Logiken. Grundlage ist hier, wie bei [wissensbasierten Systemen](#) die Wissensrepräsentationshypothese von Brian C. Smith (Smith 1982). Da das Wissen nicht zum Selbstzweck repräsentiert wird, gehört üblicherweise ein Mechanismus dazu, der es erlaubt, Schlussfolgerungen zu ziehen beziehungsweise neues Wissen aus dem vorhandenen Wissen abzuleiten. Das geschieht durch ein automatisiertes Verfahren zur Inferenz. Dies führt dann zum gemeinsamen Begriff Knowledge Representation and Reasoning (KRR). Eine gute Einführung in die Thematik gibt etwa (Levesque 1986). In deutscher Sprache ist das Buch von (Bibel et al. 1993) zu empfehlen. Weiterführende Details werden in (Brachman und Levesque 2004) umfassend behandelt. Auch Kapitel 3 in (Beierle und Kern-Isberner 2019) behandelt das Themengebiet und gibt Beispiele.

Quellen

- Beierle C, Kern-Isberner G (2019) Computational Intelligence. Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen, 6. Aufl. Springer Vieweg, Wiesbaden
- Bibel W, Hölldobler St, Schaub T (1993) Wissensrepräsentation und Inferenz - Eine grundlegende Einführung. Künstliche Intelligenz. Vieweg, Wiesbaden
- Brachman RJ, Levesque HJ (2004) Knowledge Representation and Reasoning. Elsevier
- Levesque HJ (1986) Knowledge representation and reasoning. Annual Review of Computer Science 1(1):255–287. doi:10.1146/annurev.cs.01.060186.001351
- Smith BC (1982) Reflection and semantics in a procedural language. Ph.D. Thesis. M.I.T., Computer Science lab. MIT

Autoren

- Dr. Stefan Schiffer (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)
- Prof. Dr. Alexander Ferrein (Informatik, Institut für Mobile Autonome Systeme und Kognitive Robotik (MASKOR), FH Aachen University of Applied Sciences, WIRKsam & AKzentE4.0)

Danksagung

Die Herausgebenden danken herzlich den nachfolgenden Personen für ihr Engagement, ihren kritischen Blick und vor allem die wertschätzende Rückmeldung, die Argumente geschärft und damit die Diskussion weitergebracht haben.

- Dr. Andrea Altepost
Kompetenzzentrum WIRKsam
Institut für Textiltechnik, RWTH Aachen University
- Marco Baumgartner
Kompetenzzentrum KARL
Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken (ILIN)
- Dr. Christian Bernhard-Skala
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen (DIE)
- Dr. Martin Braun
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
- Anja Brückner
Kompetenzzentrum KMI (Künstlich-Menschlich-Intelligent)
Institut für Angewandte Informatik (InfAI)
- Maximilian Duisberg
Kompetenzzentrum AKzentE4.0
Institut für Arbeitswissenschaft, RWTH Aachen University
- Robert Eckardt
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Hochschule Mittweida
- Mohammad Ehsan Matour
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Hochschule Mittweida
- Farah Elaroussi
Kompetenzzentrum WIRKsam
Institut für Textiltechnik, RWTH Aachen University
- Gilbert Engert
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Produktionsmanagement, Technologie und Werkzeugmaschinen (PTW)
- Annika Franken
Kompetenzzentrum AKzentE4.0
FIR an der RWTH Aachen
- Philipp Frey
Kompetenzzentrum KARL
Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am KIT
- Dr.-Ing. Rico Ganßauge
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Brandenburgische Technische Universität Cottbus-Senftenberg (BTU)
- Nadja Geisler
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Data and AI Systems
- Michael Gühne
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Technische Universität Dresden
- Dr. Adjan Hansen-Ampah
Kompetenzzentrum WIRKsam
Institut für Textiltechnik, RWTH Aachen University
- Dr. Markus Harlacher
Kompetenzzentrum WIRKsam
ifaa – Institut für angewandte Arbeitswissenschaft e. V.

- Lea Heitlinger
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Marketing und Personal (MuP)
- Lena Hintzen
Kompetenzzentrum AKzentE4.0
MA&T Sell & Partner GmbH
- Dr. Tim Jeske
Kompetenzzentrum WIRKsam
ifaa – Institut für angewandte Arbeitswissenschaft e.V.
- Philip Joisten
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Arbeitswissenschaften (IAD)
- Maria Jung
Kompetenzzentrum KompAKI
Hochschule Darmstadt, Fachbereich Gesellschaftswissenschaften
- Elena Kick
Kompetenzzentrum KARL
Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken (ILIN)
- Dr. Amit Kirschenbaum
Kompetenzzentrum KMI (Künstlich-Menschlich-Intelligent)
Institut für Angewandte Informatik (InfAI)
- Dr. Tobias Kopp
Kompetenzzentrum KARL
Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken (ILIN)
- Prof. Dr. Ralf Krestel
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Leibniz-Informationszentrum Wirtschaft (ZBW)
- Dr. Bettina-Johanna Krings
Kompetenzzentrum KARL
Institut für Technikfolgenabschätzung und Systemanalyse (ITAS) am KIT

- Enno Lang
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Produktionsmanagement,
Technologie und Werkzeugmaschinen (PTW)
- Dr. Athanasios Mazarakis
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Leibniz-Informationszentrum Wirtschaft (ZBW)
- Maren Mehler
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Lehrstuhl Wirtschaftsinformatik | Software & Digital Business
- Dr. Charlott Menke
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Fraunhofer-Zentrum für Internationales Management und Wissensökonomie IMW
- Prof. Alexander Mertens
Kompetenzzentrum AKzentE4.0
Institut für Arbeitswissenschaft, RWTH Aachen University
- Wolfgang Merx
Kompetenzzentrum WIRKsam
Institut für Textiltechnik, RWTH Aachen University
- Prof. Susanne Mütze-Niewöhner
Kompetenzzentrum AKzentE4.0
Institut für Arbeitswissenschaft, RWTH Aachen University
- Sina Niehues
Kompetenzzentrum WIRKsam
ifaa – Institut für angewandte Arbeitswissenschaft e.V.
- Leon Pfenning
Kompetenzzentrum KompAKI
Hochschule Darmstadt, Assisted Working and Automation (AWA)
- Christina Pietschmann
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Hochschule Mittweida

- Ulrike Pietrzyk
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Technische Universität Dresden
- Lucas Polanski-Schräder
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Arbeitswissenschaften (IAD)
- Gerda Maria Ramm
Kompetenzzentrum WIRKsam,
ifaa – Institut für angewandte Arbeitswissenschaft e. V.
- Vera Rick
Kompetenzzentrum AKzentE4.0
Institut für Arbeitswissenschaft, RWTH Aachen University
- Sophie Sandner
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Institut für Produktionsmanagement,
Technologie und Werkzeugmaschinen (PTW)
- Edgar Schero
Kompetenzzentrum KMI (Künstlich-Menschlich-Intelligent)
Institut für Angewandte Informatik (InfAI)
- Ina Schiedermaier
Kompetenzzentrum KARL
Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken (ILIN)
- Dr. Stefan Schiffer
Kompetenzzentren AKzentE 4.0 & WIRKsam
MASKOR Institut, FH Aachen University of Applied Sciences
- Carsten Schmidt
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
- Welf Schröter
Kompetenzzentrum KARL
Forum Soziale Technikgestaltung (FST)
- Maria Schweren
Kompetenzzentrum KMI (Künstlich-Menschlich-Intelligent)
Institut für Angewandte Informatik (InfAI)
- Florian Spitzbarth
Kompetenzzentrum PerspektiveArbeit Lausitz (PAL)
Hochschule Mittweida
- Dr. Anne-Sophie Tombeil
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
- Karl Trela
Metavorhaben und Begleitforschungsprojekt CoCo – Connect & Collect
Fraunhofer-Zentrum für Internationales Management und Wissensökonomie IMW
- Oliver Vetter
Kompetenzzentrum KompAKI
Technische Universität Darmstadt, Lehrstuhl Wirtschaftsinformatik | Software & Digital Business
- Aline Vugrincic
Kompetenzzentrum KARL
FZI Forschungszentrum Informatik
- Heiko Webert
Kompetenzzentrum KompAKI
Hochschule Darmstadt, Zentrum für Robotik und Automation (CRA)
- Samantha Werens
Kompetenzzentrum KompAKI
Hochschule Darmstadt, Fachbereich Gesellschaftswissenschaften
- Mandy Wölke
Kompetenzzentrum KMI (Künstlich-Menschlich-Intelligent)
Institut für Angewandte Informatik (InfAI)
- Alexander Zender
Kompetenzzentrum KompAKI
Hochschule Darmstadt, Institut für angewandte Informatik

Impressum

Herausgeber

Dennis Richter (Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken)

Dr. Christian Bernhard-Skala (Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen)

Prof. Dr. Steffen Kinkel (Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken)

Konzeption und Umsetzung

Dennis Richter (Kompetenzzentrum KARL)

Dr. Christian Bernhard-Skala (Projekt CoCo – Connect & Collect)

Prof. Dr. Steffen Kinkel (Kompetenzzentrum KARL)

Gestaltung/Layout

Medienarchitekten Bonn

Stand

März 2024

Förderhinweis

Die Projekte „KARL- Künstliche Intelligenz für Arbeit und Lernen in der Region Karlsruhe“ und „CoCo – Connect & Collect – KI-gestützte Cloud für die interdisziplinäre vernetzte Forschung und Innovation für die Zukunftsarbeit“ werden durch das Bundesministerium für Bildung und Forschung (BMBF) im Programm „Zukunft der Wertschöpfung – Forschung zu Produktion, Dienstleistung und Arbeit“ gefördert und vom Projektträger Karlsruhe (PTKA) betreut (Förderkennzeichen: KARL: 02L19C250; CoCo: 02L19C000 ff).