

ÖAW

ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN

PROJEKTBERICHT

WIEN, NOVEMBER/2024
ITA-2024-07
WWW.OEAW.AC.AT/ITA

CAIL – CRITICAL AI LITERACY

KRITISCHE TECHNIKKOMPETENZ FÜR KONSTRUKTIVEN UMGANG
MIT KI-BASIERTER TECHNOLOGIE IN BETRIEBEN

GEFÖRDERT DURCH
Digifonds



CAIL – CRITICAL AI LITERACY

KRITISCHE TECHNIKKOMPETENZ FÜR KONSTRUKTIVEN
UMGANG MIT KI-BASIERTER TECHNOLOGIE IN BETRIEBEN

Institut für Technikfolgen-Abschätzung
der Österreichischen Akademie der Wissenschaften
Projektleitung: Stefan Strauß
Autor*innen: Stefan Strauß

Gefördert im Rahmen des Digitalisierungsfonds
der Arbeiterkammer Wien, Projekt-Nr. 5-349
(Unter Mitarbeit von Titus Udrea)

Wien, November 2024

Danksagung

Herzlicher Dank ergeht an alle Mitglieder:innen des Beirats für Arbeit und Technik (BAT) der Gewerkschaft der Privatangestellten (GPA), allen Interviewpartner:innen und Workshop-Teilnehmer:innen sowie an die Mitarbeiter:innen des Büros für digitale Agenden der Arbeiterkammer Wien.

IMPRESSUM

Medieninhaber:

Österreichische Akademie der Wissenschaften
Juristische Person öffentlichen Rechts (BGBl 569/1921 idF BGBl I 31/2018)
Dr. Ignaz Seipel-Platz 2, A-1010 Wien

Herausgeber:

Institut für Technikfolgen-Abschätzung (ITA)
Bäckerstraße 13, A-1010 Wien
www.oeaw.ac.at/ita

Die ITA-Projektberichte erscheinen unregelmäßig und dienen der Veröffentlichung der Forschungsergebnisse des Instituts für Technikfolgen-Abschätzung.

Die Berichte erscheinen in geringer Auflage im Druck und werden über das Internetportal „epub.oeaw“ der Öffentlichkeit zur Verfügung gestellt:

epub.oeaw.ac.at/ita/ita-projektberichte

ITA-Projektbericht Nr.: ITA-2024-07 (Wien, November/2024)

ISSN: 1819-1320

ISSN-online: 1818-6556

epub.oeaw.ac.at/ita/ita-projektberichte/ITA-2024-07.pdf

Dieser Bericht unterliegt der Creative Commons Attribution 4.0 International License:
creativecommons.org/licenses/by/4.0/

INHALT

	ZUSAMMENFASSUNG	5
	EXECUTIVE SUMMARY	9
1	EINLEITUNG	12
1.1	Gängige Narrative und Erwartungen an KI	14
1.2	Erwartete Effekte und Nutzungsverhalten	15
1.2.1	Nutzungsverhalten	16
1.2.2	Gründe gegen die Nutzung	17
2	GRUNDLAGEN UND FUNKTIONSWEISE	18
2.1	Zum Begriff KI	18
2.2	Wie verändert KI die Gesellschaft?	19
2.3	Funktionsweise von KI	20
2.3.1	Überblick zu Machine Learning Verfahren	22
2.3.2	Nutzen und Grenzen KI-generierter Ergebnisse	24
3	WISSENSARBEIT UND KI: EIN WIDERSPRUCH?	27
3.1	Übersicht zu ausgewählten Anwendungsbereichen	30
3.1.1	Medizin und Gesundheit	31
3.1.2	Journalismus und Medien	34
3.1.3	Schule und Bildungswesen	36
3.1.4	Bereichsübergreifender KI-Einsatz im Berufsalltag	37
4	BESONDERHEITEN KI-BASIRTER AUTOMATION	40
4.1	Höhere Dynamik und Volatilität im Systemverhalten	41
4.2	(Deep) Automation Bias: Ein zentrales Risiko des KI-Einsatzes	42
5	ZENTRALE HERAUSFORDERUNGEN UND PROBLEMFELDER	46
5.1	Welchen Mehrwert kann KI bringen?	46
5.2	Kluft zwischen KI-Nutzung als Expertensystem und „Nebenbei“-Technologie	49
5.3	Akzeptanz und Akzeptabilität	51
5.4	„Verdeckte KI“: Intransparenz und Datenanalysen im Hintergrund	54
5.5	Neuer Qualifikationsbedarf aber uneindeutige Tätigkeitsbereiche	57
6	HANDLUNGSFÄHIGKEIT MIT CRITICAL AI LITERACY STÄRKEN	61
6.1	CAIL-Framework	64
6.2	Handlungsempfehlungen für Betriebe	67
7	ZUSAMMENFASSUNG UND AUSBLICK	71
8	LITERATUR	75

ABBILDUNGSVERZEICHNIS

Abbildung 1: KI als vereinfachtes Prozess-Modell	21
Abbildung 2: Beispiel einer KI-Bildererkennung mit Deep Learning	21
Abbildung 3: Einflussfaktoren von Deep Automation Bias	44
Abbildung 4: CAIL-Framework	64
Abbildung 5: Entscheidungshilfe bei der KI-Einführung	67

TABELLENVERZEICHNIS

Tabelle 1: Gründe gegen KI-Nutzung	17
------------------------------------	----

ZUSAMMENFASSUNG

Künstliche Intelligenz (KI) ist seit einigen Jahren zu einer starken Triebfeder der Digitalisierung geworden. Der anhaltende Hype wurde zuletzt noch weiter verstärkt durch generative KI-Anwendungen und intensive mediale Berichterstattung über mittlerweile auch in der breiten Öffentlichkeit bekannte generative KI-Tools. Dieser Hype verstellt jedoch etwas den Blick darauf, dass nicht alles neu ist, was nun KI genannt wird und es sehr unterschiedliche Anwendungsfelder gibt. Debatten über die Auswirkungen KI-basierter Technologien auf Wirtschaft und Arbeitswelt sind seit längerem Gegenstand einer fortwährenden Entwicklung. Eine wachsende Zahl an Unternehmen setzt diese Technologien bereits in unterschiedlichen Bereichen ein. Das betrifft immer mehr Berufssparten, die bislang kaum von diesem technologischen Wandel berührt waren. Viele Arbeitnehmer:innen sind daher in ihrem betrieblichen Alltag mittelfristig mit deutlichen Veränderungen konfrontiert.

Das besondere Novum KI-basierter Technologien sind nicht nur ihre mittlerweile breiten Anwendungsvarianten, sondern vor allem die neuen Automatisierungsformen, die ihr Einsatz mit sich bringt. Der damit verbundene Wandel betrifft zusehends die Wissensarbeit, jedoch ist noch weitgehend offen, wie sich diese Veränderungen manifestieren. Das im Rahmen des Digitalisierungsfonds der Arbeiterkammer Wien geförderte Forschungsprojekt CAIL¹ hat daher untersucht, wie sich Arbeitspraktiken der Wissensarbeit durch den Einsatz KI-basierter Technologien und ihren neuen Automatisierungsformen verändern und welche Herausforderungen damit verbunden sind. Darauf aufbauend wurden Ansätze zur Etablierung einer Critical AI Literacy (CAIL) – also einer kritischen Kompetenz im Umgang mit KI-basierten Technologien – ausgearbeitet. Der vorliegende Bericht fasst wesentliche im Rahmen des Forschungsprojekts gewonnene Erkenntnisse zusammen und geht dabei insbesondere auf folgende Fragestellungen ein:

- Welche Erwartungen haben Unternehmen an den KI-Einsatz?
- Was sind wesentliche Besonderheiten KI-basierter Automatisierung?
- Wie wirkt sich KI-Einsatz auf Arbeitspraktiken der Wissensarbeit aus?
- Welche Herausforderungen und Problemfelder sind damit verbunden und welche Zusammenhänge zum Risiko von (Deep) Automation Bias bestehen?
- Welche Einflussfaktoren sind für den konstruktiven KI-Einsatz in Unternehmen relevant?
- Welche Grundkompetenzen umfasst Critical AI Literacy und mit welchen Ansätzen ist das entsprechende Wissen vermittelbar?

In Abschnitt 1 wird kurz auf wesentliche Erwartungen, Narrative und aktuelle Nutzungszahlen von KI in Österreich eingegangen. Die Bedeutung und Funktionsweise von KI wird in Abschnitt 2 erläutert. In Abschnitt 3 wird die Frage behandelt, wie sich Wissensarbeit verändert und anhand ausgewählter Anwendungsfelder gezeigt, welche Folgen die Technologie auf Arbeitspraktiken bereits hat. Abschnitt 4 geht auf wesentliche Besonderheiten KI-basierter Automatisierung

Abseits des Hypes ist KI in vielen Bereichen bereits im Einsatz

KI bedeutet vor allem: neue Formen der Automatisierung

Übersicht

¹ Projekt-Nr. 5-349

ein und zeigt auf, warum (Deep) Automation Bias ein Meta-Risiko des KI-Einsatzes darstellt. In Abschnitt 5 werden zentrale Herausforderungen und Problemfelder identifiziert und diskutiert. Abschnitt 6 geht auf den wesentlichen Aspekt der Handlungsfähigkeit ein, erläutert die Bedeutung von CAIL und zeigt anhand des entwickelten CAIL-Frameworks, wie kritische KI-Kompetenz in Betrieben vermittelt werden kann. Zur Orientierung bei der Planung und Gestaltung des KI-Einsatzes wurden zudem Handlungsempfehlungen für Unternehmen ausgearbeitet (Abschnitt 6.2). In Abschnitt 7 werden wesentliche Ergebnisse kurz zusammenfassend dargestellt.

Die Ergebnisse verdeutlichen eine zentrale Voraussetzung für die konstruktive Nutzung KI-basierter Technologien: Wissen über die Funktionsweisen und Grenzen in konkreten Anwendungskontexten. KI-basierte Automatisierung ist komplexer, dynamischer und volatiler als klassische Automatisierungsformen (Abschnitt 4). Dadurch entstehen neue Herausforderungen und Bedarf nach Strategien zu deren Bewältigung. Entgegen gängiger Narrative führt KI-Einsatz nicht automatisch zu Produktivitäts- und Effizienzgewinnen. Ein Mehrwert ist nur dann realisierbar, wenn die Technologie in Einklang mit Arbeitspraktiken und Arbeitsorganisation in betriebliche Abläufe integrierbar ist. Das setzt auch Klarheit über Anwendungszweck und Ziele des KI-Einsatzes voraus – es braucht daher neben Transparenz der Funktionsweise von KI-Systemen auch eine Nutzungs-Transparenz.

Ein kritischer Faktor des KI-Einsatzes betrifft die Anwendungs- und Prozessqualität sowie Maßnahmen zu ihrer Gewährleistung. KI kann zwar rasch Ergebnisse produzieren, jedoch ohne zusätzliche Maßnahmen keine Korrektheit oder Zuverlässigkeit bieten. Die Interpretation und Überprüfung bleibt eine wesentliche menschliche Aufgabe, die künftig noch wichtiger wird. Gleiches gilt für Tätigkeiten des Datenmanagements (Datenaufbereitung, Datenpflege, Bereinigung) sowie der Qualitätssicherung und Fehlerkorrektur. Es entstehen also auch neue Aufgaben, die auch mehr Arbeitsaufwand bedeuten – vor allem für das Qualitätsmanagement von KI-Systemen. Eine Prozess-Sicht auf die KI-Einbettung in Arbeitsabläufe ist daher wichtig, um diese Tätigkeiten so zu gestalten, dass die möglichen Entlastungen in einem Arbeitsbereich nicht zur Überlastung in anderen Arbeitsbereichen führen.

Hierbei macht die Art des Nutzungsgrades einen wichtigen Unterschied, die entlang zweier Pole beschreibbar ist (Abschnitt 5): Zum einen die KI-Nutzung als bewusst eingesetztes Werkzeug in einem Expertensystem, zum anderen als Art „En-passant“-Technologie ohne klare Tätigkeitsbezüge und geringen Qualitätsanforderungen. In ersterem besteht tendenziell ein relativ hohes Maß an Fachexpertise und Problembewusstsein (z. B. in der Medizin). Das begünstigt die Integration der Technologie in Einklang mit bestehenden Arbeitspraktiken zur Unterstützung konkreter Tätigkeiten mit praktischem Nutzen. Dem steht ein unkonkreter, wenig Zweck- und zielorientierter Einsatz gegenüber, wo oftmals wage bleibt, welche Tätigkeiten mit KI unterstützt werden sollen und ob das realisierbar ist. D. h. je breiter und unspezifischer die KI-Nutzung ausfällt, desto geringer auch der potenzielle Mehrwert und desto wahrscheinlicher werden praktische Probleme. Zudem können bei mangelhafter Einbeziehung von Arbeitnehmer:innen und deren Praktiken viele Anwendungsdefizite erst nach Technologie-Einführung sichtbar werden. Erfolgt der KI-Einsatz zu vorschnell, können neue technisch-organisatorische Abhängigkeiten entstehen, die sich eher belastend auf die

*KI-basierte
Automatisierung
ist komplexer,
dynamischer und
volatiler*

*KI bringt mehr
Arbeitsaufwand für
Datenpflege und
Qualitätsmanagement*

*Kluft zwischen KI
als Expertensystem
und beiläufiger
Nutzung*

Effizienz der Arbeitsabläufe und damit auch negativ auf die Akzeptanz und Akzeptabilität im Unternehmen auswirken.

Damit einher geht ein wachsendes und unterschätztes Problemfeld: Die schleichende Einführung von KI-Funktionen über Standardsoftware wie z. B. weit verbreitete Betriebssysteme, die sukzessive KI-Funktionen in verschiedene Office-Anwendungen integrieren (siehe Abschnitt 5.4). Das kann zu mehr Intransparenz und mangelnder Kontextualisierung führen. Etwa, wenn die Funktionsweisen und Datenflüsse von verdeckten, im Hintergrund laufenden KI-Anwendungen nicht mehr nachvollziehbar und kontrollierbar sind. Die Folge können verstärkte Datenschutz- und Sicherheitsprobleme sein. Aufgrund der starken Marktmacht einzelner großer Softwareanbieter besteht i. d. R. ein hohes Abhängigkeitsverhältnis zu diesen. Die meisten Betriebe sind auf diese Systeme angewiesen und haben oftmals wenig Möglichkeiten zur Mitgestaltung in Einklang mit den betrieblichen Bedingungen. Das kann sich negativ auf die notwendige Kontextualisierung auswirken, also die bewusste Einbindung von KI-Anwendungen für bestimmte Nutzungskontexte entsprechend den Anforderungen der betroffenen Arbeitsabläufe. Zudem kann die Abhängigkeit zentraler Geschäftsprozesse zu KI-Systemen als Bestandteil der unternehmensinternen kritischen Infrastruktur die Vulnerabilität des Unternehmens erhöhen. Bei der betrieblichen Entscheidung zur KI-Einführung spielen daher auch Open-Source Varianten für KI-Anwendungen eine wichtige Rolle, um einige der mit Standardanwendungen einher gehenden Probleme zu vermeiden.

Der durch KI-Einsatz erzielbare Mehrwert (Abschnitt 5.1) und die Auswirkungen auf die Arbeitsorganisation hängen wesentlich von der Art der Nutzung der Anwendungen und ihrer Integration in konkrete Tätigkeiten ab. Abseits hoher Erwartungen und hoher Automationsgrade ist das konkrete Ausmaß der Automatisierung bestimmter Tätigkeiten im betrieblichen Arbeitskontext mitgestaltbar. Ein grundsätzlich hoher KI-Automationsgrad impliziert noch keinen praktischen Nutzen. D. h. es gilt zu unterscheiden zwischen dem, was eine KI-Anwendung für bestimmte Arbeitsschritte leisten kann, und inwieweit das auch in bestehende Arbeitsabläufe der betrieblichen Arbeitspraxis sinnvoll integrierbar ist. Entgegen gängiger Annahmen eignet sich KI weniger für komplexe Aufgaben, sondern vielmehr für Routinetätigkeiten. Hierbei sind die Form der Wissensarbeit und konkret betroffene Tätigkeiten wesentlich. Wie in Abschnitt 3 ausgearbeitet, umfasst Wissensarbeit i. d. R. eine Kombination aus „manuellen“ bzw. repetitiven Routinetätigkeiten und komplexen Aufgaben. Letztere sind nicht ohne weiteres mit KI automatisierbar, weil es keine formalisierbaren Bearbeitungswege gibt. Zudem sind in der Arbeitspraxis beide Tätigkeitsformen nicht völlig isoliert voneinander unterscheidbar. Ein konstruktiver Einsatz von KI in der Wissensarbeit setzt voraus, dass die Einbettung der Technologie zur Entlastung bei monotonen Routineaufgaben führt, sodass Freiräume für komplexere, vielseitigere Aufgaben entstehen. KI-Anwendungen können – sofern die Ergebnisse zuverlässig sind – Wissensarbeit und Wissensproduktion in vielerlei Hinsicht ergänzen. Jedoch nicht durch Generierung neuen Wissens, sondern nur durch Kombination und Aufbereitung bestehender Informationen. Das Verstehen, Interpretieren und Kontextualisieren KI-generierter Information, d. h. ihre Einbettung in konkrete Anwendungskontexte wird daher Teil des Aufgabenspektrums von Wissensarbeit. D. h., Wissensarbeit benötigt zusätzlich zur „klassischen“ Problemlösungskompetenz nun auch die Kompetenz zu erkennen, ob und wie KI-Einsatz eine Tätigkeit sinnvoll unterstützen kann oder nicht. Das ist mit einem wachsenden

*Schleichende
KI-Einführung ist
ein unterschätztes
Problem*

*KI kann Wissensarbeit
ergänzen, wenn es bei
Routineaufgaben
entlastet und
Handlungsfähigkeit
stärkt*

*KI kann kein neues
Wissen generieren, aber
bei seiner Entwicklung
stimulieren*

*Kritische
KI-Kompetenz wird
zum Bestandteil von
Wissensarbeit*

Bedarf an Wissen verknüpft, um im Sinne einer Critical AI Literacy zu einem breiteren Grundverständnis über Funktionsweise, Eignung und Relevanz KI-basierter Technologien in bestimmten Anwendungskontexten zu gelangen. Das hierzu erarbeitete CAIL-Framework (Abschnitt 6) bietet ein niedrighschwelliges Instrument zur Vermittlung von Basiswissen, das bei der Beurteilung des praktischen Nutzens und der Grenzen von KI-Anwendungen und KI-generierten Ergebnissen unterstützen kann. Die daran anknüpfenden Handlungsempfehlungen können als Unterstützungshilfe für Unternehmen und Betriebsräte dienen, um gemeinsam mit der Belegschaft zu einer konstruktiven, sozialverträglichen Gestaltung und Nutzung KI-basierter Technologien zu gelangen.

Die Transformation der Wissensarbeit durch KI ist in vielerlei Hinsicht noch am Anfang und mit vielen offenen Fragen verbunden. Das derzeitige Entwicklungsstadium lässt sich als Teil eines größeren, institutionellen und gesellschaftlichen Lernprozesses darüber begreifen, ob und wie KI-basierte Technologien konstruktiv und praktikabel einsetzbar sind. Dabei muss zunächst auch in Unternehmen erst geklärt werden, was Technologieeinsatz leisten kann, ob das in der Arbeitspraxis zielführend ist, und welche Folgen sich daraus für die jeweiligen Arbeitsbereiche ergeben. In vielen Unternehmen hat gerade erst eine Art Lern- und Experimentierphase mit KI begonnen, die idealerweise dazu beiträgt, über die Organisation von Arbeit, die Bedeutung sozialverträglicher Arbeitsbedingungen und Mitarbeiterzufriedenheit neu zu reflektieren. Wenn es gelingt, in den nächsten Jahren zu einem tieferen Grundverständnis von KI-basierten Technologien und der Etablierung einer kritischen Kompetenz im Umgang damit zu gelangen, ist ein wesentliches Lernziel für die konstruktive Nutzung dieser Technologien erreicht.

*Gesellschaftlicher
Lernprozess über die
konstruktive
KI-Nutzung*

EXECUTIVE SUMMARY

Artificial intelligence (AI) has become a strong driver behind digitalization in recent years. Prominent and intensively promoted generative AI-tools have been reinforcing the ongoing hype also in the general public. The hype obscures the fact that not everything now called AI is new and that a broad range of different applications already exist. There are continuing debates on the impact of AI-based technologies on labour and the economy and a growing number of companies already apply these technologies in various areas. This technological change affects more and more professions that were previously not involved. On the longer run, many employees have to encounter significant changes in their daily work life.

The specific novelty of AI-based technologies is not the wide range of applications, but basically the new forms of automation that their use entails. The associated change increasingly affects knowledge work, though, yet widely uncertain to what extent. The CAIL research project, funded by the Digitalization Fund of the Viennese Chamber of Labour, thus investigated how practices of knowledge work alter through the use of AI-based technologies and the related new forms of automation and what challenges result from this development. Based on this, approaches for establishing Critical AI Literacy (CAIL) – i.e. critical competence in dealing with AI-based technologies – were developed. This report summarizes the key findings of the research project and particularly addresses the following questions:

- What do companies expect from the use of AI?
- What are peculiarities of AI-based automation?
- How does the use of AI affect practices of knowledge work?
- What are corresponding main challenges and problems and how do they relate to the risk of (deep) automation bias?
- What are main influencing factors for the constructive use of AI in companies?
- What are basic skills of a Critical AI Literacy and what approaches can be used to impart the corresponding knowledge?

Following this introduction, section 1 briefly discusses key expectations, narratives and current usage figures of AI in Austria. The significance and functionality of AI is explained in section 2. Section 3 deals with the question how knowledge work is changing and – based on selected fields of application – outlines what impact on working practices is already observable. Section 4 explores crucial peculiarities of AI-based automation and explains why (deep) automation bias represents a meta-risk of AI usage. Section 5 identifies and discusses key challenges and problem areas, section 6 addresses the essential aspect of agency, explains the importance of CAIL and outlines – based on the elaborated CAIL framework – how critical AI literacy can be communicated within companies. The subsequent recommendations can support decision-makers in the planning and deployment of AI-based technologies (section 6.2). Section 7 briefly summarizes the key results.

Despite the hype AI is already used in many domains

AI implies new forms of automation

Overview

The results highlight a central prerequisite for the constructive use of AI-based technologies: Knowledge about the modes of operation and limitations in specific application contexts of AI. AI-based automation is more complex, more dynamic and more volatile than traditional forms of automation. This creates new challenges and demand for coping strategies. Contrary to common narratives, the use of AI does not automatically lead to more productivity and efficiency gains. Added value can only emerge if the technology is integrated into operational processes in line with working practices and the organization of work. This also requires clarity about the purpose and objectives of the use of AI. Hence, not only transparency about how AI systems work is required but also transparency of utilization.

A critical factor of AI usage relates to the process quality and measures to ensure it. AI can produce results quickly, but it cannot provide accuracy or reliability without additional measures. Interpretation and verification thus remain crucial human tasks that becomes even more important. The same applies to data management activities (data preparation, data maintenance, cleansing) as well as quality assurance, failure detection and correction. Thus, also new tasks arise, which implies increasing work load – especially for quality management of AI systems. A process view on the integration of AI in workflows is therefore crucial in order to achieve a proper implementation that effectively relieves workload and does not lead to overload in other working domains.

The type of usage makes an important difference here, which can be described along two poles (section 5): On the one hand, AI deployed as a specific tool involved in an expert system; and on the other as a kind of „en passant“ technology without clear task-related purpose and low quality requirements. In expert environments, there tends to be a relatively high level of technical expertise and problem awareness (e.g. in medicine). This facilitates the integration of the technology in line with existing work practices to support concrete activities to achieve practical benefits. This contrasts with an unspecific, less purpose- and goal-oriented use, where it often remains unclear which activity is to be supported with AI and whether this is feasible. In other words, the broader and less specific the use of AI, the lower the potential added value and the more likely practical problems will arise. Moreover, if employees and their practices are not sufficiently considered, many practical deficits may become apparent only afterwards, i.e., when the technology is already in use. Premature AI usage can aggravate additional technical and organizational dependencies and hamper efficiency, working conditions and thus also lower acceptance and acceptability within the company.

This relates to an increasing and yet underestimated problem area: the creeping integration of AI functions via standard software such as widespread operating systems, that gradually integrate AI functions into various applications (see section 5.4). This can increase a lack of transparency and contextualization. For example, if the functions and data flows of hidden AI applications running in the background are no longer traceable and controllable. This can aggravate data protection and security issues. Moreover, only few large software providers dominate global markets. Most companies are thus dependent on their systems and often lack in options adjust the applications according to their operational needs and working conditions. This can have a negative impact on the necessary contextualization, i.e., entails an insufficient integration of AI applications that is not tailored to specific usage contexts in accordance with the requirements of the corresponding work practices. Furthermore, the dependency of central business

AI-based automation is more complex, dynamic and volatile

AI entails more workload for data maintenance & quality management

Gap between AI as expert system and „en passant“ usage

creeping AI integration is an underestimated problem

processes on AI systems as part of the internal critical infrastructure can increase the company's vulnerability. Therefore, considering open source variants for AI applications also play an important role in the operational decision to introduce AI to avoid some of the outlined problems with standard software.

The extent, to which AI usage can create to added value depends largely on usage patterns how applications are integrated into specific activities. Apart from high expectations and high levels of automation, the specific extent of automation of certain activities in the operational work context can be shaped. A fundamentally high degree of AI automation does not yet imply any practical benefit. In other words, a distinction needs to be made between what an AI application can achieve for certain work steps and the extent to which this can be meaningfully integrated into existing workflows in operational work practice. In contrast to common assumptions, AI is less suitable for complex tasks than for simpler routine activities. The form of knowledge work and the specific activities involved are essential here. As elaborated in section 3, knowledge work usually involves a combination of „manual“ routine activities and complex tasks. The latter cannot be easily automated with AI as they have no formalizable procedures. Moreover, both types cannot be completely distinguished from each other in practice. A prerequisite for a constructive use of AI in knowledge work is that the technology eases from monotonous routine tasks and enables more resources for complex, varying tasks. Provided the results are reliable, AI applications can complement knowledge work and knowledge production in many ways. However, not by generating new knowledge, but only by combining and processing existing information. Understanding, interpreting and contextualizing AI-generated information, i.e. embedding it into specific application contexts, therefore becomes part of the range of tasks involved in knowledge work. In other words, in addition to „classic“ problem-solving skills, knowledge work now also requires the ability to recognize whether and how the use of AI can support an activity or not. This highlights the increasing demand to establish a critical AI Literacy, i.e. knowledge required to understand the functionality, practicability and relevance of AI-based technologies in specific application contexts. The CAIL framework (Section 6) offers a low-threshold instrument for imparting basic knowledge that can support assessing the practical benefits and limitations of AI applications and AI-generated results. The related recommendations aim to support companies and works councils in achieving a constructive, socially acceptable implementation and use of AI-based technologies.

In many respects, the transformation of knowledge work is still in its infancies with many open questions. This ongoing development involves a larger, institutional and social learning process about whether and how constructive and practical use of AI-based technologies is feasible on the long run. This implies clarifying usage purposes, objectives, whether they are achievable in practice and what the consequences this will have for the respective working areas. In many companies, a kind of learning and experimentation phase with AI has only just begun. This ideally contributes to a new reflection on the organization of work, the importance of socially acceptable working conditions and employee satisfaction. If it succeeds in the next years, to gain a deeper basic understanding of AI-based technologies and establish a critical competence in dealing with them a key learning objective for the constructive use of these technologies is achieved.

AI can complement knowledge work by easing routine tasks and fostering agency

AI cannot produce new knowledge but stimulate its emergence

Critical AI Literacy becomes part of knowledge work

1 EINLEITUNG

Künstliche Intelligenz (KI) ist in den letzten Jahren zur zentralen Triebfeder der Digitalisierung und des damit verbundenen gesellschaftlichen Wandels geworden. Der anhaltende Hype wurde zuletzt noch weiter verstärkt durch generative KI-Anwendungen und intensive mediale Berichterstattung über mittlerweile auch in der breiten Öffentlichkeit bekannte generative KI-Tools mit Chatbot-Funktionalität, die auf Basis großer Sprachmodelle (Large Language Models – LLMs) vereinfachte Interaktionsformen zwischen Menschen und Maschinen in natürlicher Sprache ermöglichen. Dieser Hype verstellt jedoch etwas den Blick darauf, dass nicht alles neu ist, was nun KI genannt wird (siehe dazu auch Abschnitt 2). Die Auswirkungen KI-basierter Technologien auf Wirtschaft und Arbeitswelt sind seit längerem Gegenstand einer breiten gesellschaftlichen Debatte in einer fortwährenden Entwicklung. Eine wachsende Zahl an Unternehmen setzt diese Technologien bereits in unterschiedlicher Form ein. Daher sind viele Arbeitnehmer:innen in ihrem betrieblichen Alltag mittelfristig mit deutlichen Veränderungen konfrontiert. Die betroffenen Bereiche sind sehr unterschiedlich und umfassen zusehends auch immer mehr Berufssparten, die bislang kaum von diesem technologischen Wandel berührt waren. Etwa im Bildungsbereich, Mediensektor, Gesundheits- und Sozialwesen, oder auch im Administrationsbereich unterschiedlicher Arbeitsfelder. Neben bereichsspezifischen Unterschieden wird immer deutlicher sichtbar, dass KI-Einsatz die Wissensarbeit, also nicht-körperliche, geistige und soziale Tätigkeiten verändert. Wissensarbeit gilt als wesentlicher Faktor für Innovation in modernen, technisierten Gesellschaften und hat eine stabilisierende Wirkung für betriebliches Management (Drucker 2001). In welcher Ausprägung sich diese Veränderungen auf Arbeitspraktiken und Arbeitsorganisation tatsächlich auswirken, ist dabei weitgehend offen und bedarf daher genauerer Betrachtung.

Das besondere Novum KI-basierter Technologien sind nicht nur ihre mittlerweile breiten Anwendungsvarianten, sondern vor allem die neuen Automatisierungsformen, die ihr Einsatz mit sich bringt. In manchen Arbeitsbereichen (etwa im Finanzsektor, Kundensupport, Transport- u. Logistikwesen etc.) prognostizieren Studien sogar mittelfristige Automatisierungsgrade von über 70 % (OECD 2018; IAB 2021). Zugleich sind solche Prognosen wenig aussagekräftig und einige Untersuchungen deuten darauf hin, dass der Einsatz von KI zu weniger großen, disruptiven Umbrüchen der gesamten Arbeitswelt führt als vermutet. Naheliegender sind kontinuierliche Weiterentwicklungen der Arbeitsorganisation, wodurch auch neue Tätigkeitsfelder entstehen (Schörpf et al. 2018; Giering 2021; Lane et al. 2023; Kornwachs 2023). Zudem hat Automatisierung kontextspezifisch unterschiedliche Bedeutungen und Wirkungen. Klassische Automatisierungsformen in Produktionsbetrieben und Fertigungsstätten betreffen bislang vordergründig körperliche Tätigkeiten, folgen anderen Logiken und haben andere Implikationen als KI-basierte Automatisierungsformen. Das wirft neue Fragen zur Rolle von Wissensarbeit auf, die in unterschiedlicher Form von KI-basierter Automatisierung betroffen ist.

Es ist nicht alles neu im KI-Hype, aber die Auswirkungen werden breiter

KI bringt neue Automatisierungsformen in der Wissensarbeit

Das im Rahmen des Digitalisierungsfonds der Arbeiterkammer Wien geförderte, rund 1,5 Jahre laufende Forschungsprojekt CAIL² hat daher untersucht, wie sich Arbeitspraktiken der Wissensarbeit durch den Einsatz KI-basierter Technologien und ihren neuen Automatisierungsformen verändern und welche Herausforderungen damit verbunden sind. Darauf aufbauend wurden Ansätze zur Stärkung einer Critical AI Literacy (CAIL) – also einer kritischen Kompetenz im Umgang mit KI – ausgearbeitet. Ausgangspunkt der Untersuchung waren drei Hauptthesen: 1) Mit dem Einsatz KI-basierter Technologien entstehen neue Automatisierungsformen, die sich zusehends auf Tätigkeiten der Wissensarbeit auswirken, die bislang Menschen vorbehalten waren. 2) Eine Folge dieser Automatisierung ist auch die Zunahme entsprechender Risiken, insbesondere von (Deep) Automation Bias. 3) Zur Bewältigung der damit verbundenen Herausforderungen ist die Etablierung neuer Kompetenzen im Umgang mit KI-basierten Technologien erforderlich.

Drei Hauptthesen der Untersuchung

Im Rahmen des Projekts wurden neben umfassender Analyse von Fachliteratur, Studien und empirischen Sekundärmaterialien, statistischen Erhebungen und Medienberichten zusätzlich 12 leitfadengestützte, qualitative Interviews mit Expert:innen aus verschiedenen Branchen und Arbeitsbereichen im Zeitraum von Juni 2023 bis März 2024 durchgeführt (u. a. aus den Bereichen Journalismus und Medien, Informationstechnologie und Informatik, Finanzwesen, Medizin und Gesundheitswesen, betriebliche Interessensvertretung und Bildungswesen). Die Zwischenergebnisse des Projekts wurden zudem in zwei Workshops mit Betriebsrät:innen und Arbeitsexpert:innen diskutiert, um unterschiedliche Sichtweisen aus der betrieblichen Praxis auf die Thematik zu erlangen. Der vorliegende Bericht fasst wesentliche aus der Projektarbeit gewonnene Erkenntnisse zusammen und fokussiert dabei u. a. auf folgende Fragestellungen:

- Welche Erwartungen haben Unternehmen an den KI-Einsatz?
- Was sind wesentliche Besonderheiten KI-basierter Automatisierung?
- Wie wirkt sich KI-Einsatz auf Arbeitspraktiken der Wissensarbeit aus?
- Welche Herausforderungen und Problemfelder sind damit verbunden und welche Zusammenhänge zum Risiko von (Deep) Automation Bias bestehen?
- Welche Einflussfaktoren sind für den konstruktiven KI-Einsatz in Unternehmen relevant?
- Welche Grundkompetenzen umfasst Critical AI Literacy und mit welchen Ansätzen ist das entsprechende Wissen vermittelbar?

In Abschnitt 1 wird nach dieser Einleitung kurz auf wesentliche Erwartungen, Narrative und aktuelle Nutzungszahlen von KI in Österreich eingegangen. Die Bedeutung und Funktionsweise von KI wird in Abschnitt 2 erläutert. In Abschnitt 3 wird die Frage behandelt, wie sich die Wissensarbeit verändert und anhand ausgewählter Anwendungsfelder gezeigt, welche Folgen die Technologie auf Arbeitspraktiken bereits hat. Abschnitt 4 geht auf wesentliche Besonderheiten KI-basierter Automatisierung ein und zeigt auf, warum (Deep) Automation Bias ein Meta-Risiko des KI-Einsatzes darstellt. In Abschnitt 5 werden zentrale Herausforderungen und Problemfelder identifiziert und diskutiert. Abschnitt 6 geht auf den wesentlichen Aspekt der Handlungsfähigkeit ein, erläutert die Bedeutung von CAIL und zeigt anhand des entwickelten CAIL-Frameworks, wie kritische KI-

Inhaltlicher Überblick

² Projekt-Nr. 5-349. Laufzeit: Jan. 2023 bis Okt. 2024.

Kompetenz in Betrieben vermittelt werden kann. Zur Orientierung bei der Planung und Gestaltung des KI-Einsatzes wurden zudem Handlungsempfehlungen für Unternehmen ausgearbeitet (Abschnitt 6.2). In Abschnitt 7 werden wesentliche Ergebnisse zusammenfassend dargestellt.

1.1 GÄNGIGE NARRATIVE UND ERWARTUNGEN AN KI

KI wirkt im Hype der letzten Jahre wie ein neues Phänomen. Nicht zuletzt dank der intensiven PR-Arbeit prominenter großer Tech-Unternehmen wie OpenAI, Tesla/xAI, Microsoft, Google/Alphabet, DeepMind und Co. und hoher medialer Aufmerksamkeit. Wenn etwa der Co-Gründer von DeepMind³ und CEO eines weiteren KI-Unternehmens im Time Magazine über die KI-Revolution⁴ schreibt, als wäre sie eine unaufhaltsam auf die Gesellschaft hereinbrechende Welle, dann erzeugt das sehr viel Aufmerksamkeit in der Öffentlichkeit. Gleiches gilt für diverse Warnungen seitens bekannter Tech-Unternehmer vor den Gefahren von KI für die gesamte Menschheit.⁵ Zudem tragen facettenreiche Berichte über generative KI-Anwendungen wie z. B. ChatGPT zur Bekanntheit dieser Tools bei. Über diese hohe Aufmerksamkeit für den Themenkomplex KI werden auch verschiedene Erwartungen mittransportiert. Zum einen gibt es eine ganze Reihe von Potenzialen, die KI zugesprochen werden und als gängige Narrative Verbreitung finden. Das weckt viele Hoffnungen und eine hohe Erwartungshaltung an den Nutzen KI-basierter Technologien. Zum anderen stehen den Potenzialen naturgemäß auch Ängste, Risiken und Gefahren gegenüber (Osoba/Welser 2017; HLEG 2019; Verdegem 2021; UNESCO 2022; Kornwachs 2023; Stowasser 2023). Geläufige Potenziale beziehen sich insbesondere auf die Steigerung von:

KI-Einsatz ist mit vielen Erwartungen und Ängsten verknüpft

- Produktivität und Qualität
- Effizienz und Transparenz von Arbeitsprozessen
- Kundenzufriedenheit
- Innovationsfähigkeit
- Wettbewerbsvorteilen
- Ressourceneinsatz und -schutz
- Arbeitsflexibilität durch Assistenz im Arbeitsprozess und damit auch
- Arbeitszufriedenheit durch Entlastung der Beschäftigten

Zu den Gefahren zählen u. a.:

- Arbeitsplatzverlust und Personalabbau
- Bias und steigende Diskriminierung
- Datenschutz- und Sicherheitsprobleme
- Verletzungen der Privatsphäre
- Zunehmende Überwachung und Kontrolle
- Copyright-Verletzungen und Verlust von Geschäftsgeheimnissen
- Intransparenz und Autonomieverlust
- Entmenschlichung von Arbeitskraft

³ Ein KI-Tochterunternehmen von Google.

⁴ time.com/6310115/ai-revolution-reshape-the-world/.

⁵ www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/elon-musk-warnt-vor-ki-bedrohung-und-sieht-dringliche-gefahr-19286238.html.

Unabhängig davon, wie realistisch diese Potenziale und Gefahren sind, zeigt sich eine tendenzielle Polarisierung beim Thema KI. Das betrifft nicht nur den öffentlichen Diskurs, sondern spiegelt sich teils auch auf der betrieblichen Ebene wider, wie auch aus Interviews hervor geht: Arbeitgeber tendieren eher dazu, den erweckten Erwartungen über die Potenziale der Technologien wie Kosteneinsparungen, Leistungssteigerungen und Produktivitätsgewinne etc. zu glauben. Arbeitnehmer:innen sind dagegen tendenziell besorgter über die negativen Auswirkungen (Interviews a-d, g-i). Dieses Bild bestätigt sich auch in Studien (siehe Abschnitt 1.2). Die Gründe sind vielschichtig: Abseits der hohen öffentlichen Aufmerksamkeit und Erwartungserhaltung besteht wirtschaftlicher Druck zur Digitalisierung. Zugleich herrscht Unwissen und Unsicherheit bezüglich der konkreten Ziele des Technologieeinsatzes und zu wenig Einbindung der Belegschaft bei der Planung und Einführung von KI-Anwendungen.

Diese polarisierte Sicht mit einerseits enormer Erwartungshaltung an die Leistungsfähigkeit von KI-basierten Technologien und andererseits diversen Sorgen und Ängsten erschwert einen realistischen Umgang und eine sozial nachhaltige Gestaltung des Technologie-Einsatzes. Hinzu kommt, dass es zur Bewältigung der diversen Risiken von KI zwar eine wachsende Vielfalt an Katalogen mit ethischen Richtlinien und Grundprinzipien für eine „verantwortungsvolle“ KI gibt (z. B. Floridi 2018, HLEG 2019, Hallersleben 2020; UNESCO 2022 etc.). Diese Ansätze sind ohne Zweifel wichtig, vor allem für den politischen und regulatorischen Umgang mit KI. Aber sie bleiben meist auf einer relativ abstrakten Ebene sehr grundsätzlicher Natur mit entsprechend geringer Praxistauglichkeit. Die Umsetzung dieser zahlreichen Prinzipien in Betrieben ist daher herausfordernd und aufgrund fehlenden Praxisbezugs auch oftmals kaum möglich. Zudem orientieren sich Unternehmen meist schon aufgrund ihrer wirtschaftlichen Sachzwänge primär an der Ausschöpfung von Potenzialen, wie sich auch in Untersuchungen zu den erwarteten Effekten zeigt.

Polarisierung und unterschiedliche Sichtweisen

Viele ethische Grundprinzipien aber teils mangelnde Praxistauglichkeit

1.2 ERWARTETE EFFEKTE UND NUTZUNGSVERHALTEN

Produktivitätssteigerungen, Automatisierung, Optimierung von Prozessen, Effizienz-, Qualitäts- und Flexibilitätsgewinne, Kostensenkungen und Vorteile bei Komplexitäts- und Wissensmanagement zählen zu den häufigsten Motiven für die Nutzung von KI in Unternehmen (Prem/Ruland 2019; Lane et al. 2023). Einer aktuellen OECD-Untersuchung zufolge ist die Hauptmotivation von Unternehmen für KI-Einsatz primär, Effizienz zu steigern und Personalkosten zu reduzieren. Das beruht auf der Grunderwartung, dass KI-Einsatz sowohl Produktivität als auch Arbeitsbedingungen verbessern kann (Lane et al. 2023). Beschäftigte im Finanz- und Produktionssektor, die bereits mit KI arbeiten, bestätigen positive Effekte für die Effizienz. Rund 30 % der Befragten gaben deutliche Leistungssteigerungen in diesen Bereichen an. Wesentlich differenzierter fällt das Bild bezüglich Auswirkungen auf Arbeitszufriedenheit, Gesundheit und Fairness aus: Hier werden deutlich weniger positive Effekte wahrgenommen und auffallend höhere Werte hinsichtlich Verschlechterungen für Beschäftigte. Die Untersuchung geht nicht näher auf spezifische Anwendungen und Tätigkeiten der KI-Automatisierung ein. Aus der Studie geht auch hervor, dass die Technologie-Einführung mit Sorgen und Unsicherheiten verbunden ist. Die Mehrheit der Beschäftigten gab

etwa an, erhöhten Leistungsdruck durch KI-bedingte Datenerfassungen zu verspüren, äußerten Sorgen vor überbordenden Datensammlungen, Eingriffen in die Privatsphäre, Benachteiligung und Diskriminierung. Die Ergebnisse zeigen auch Anzeichen für Unterschiede in der Einschätzung bezüglich des Bildungsstands: Arbeitnehmer:innen mit geringerem Bildungsniveau nehmen die Effekte von KI auf Leistungssteigerung und Arbeitsbedingungen tendenziell negativer wahr als jene mit höherem Bildungsniveau (Universitätsabschluss). Das kann verschiedene Gründe haben, etwa generelle Unterschiede in den Arbeitsbereichen je nach Qualifikation (ebd.). Es verdeutlicht aber auch, dass Bildung und Expertise eine wesentliche Rolle haben, ob und wie konstruktiver Einsatz von KI-basierten Technologien möglich ist.

Die Erwartung, Produktivität zu steigern, ist naheliegend. Aber inwieweit ist sie auch erfüllbar? Die Frage nach dem Mehrwert durch KI-Einsatz wird u. a. in Abschnitt 5 genauer diskutiert. Relativ eindeutig sind die Befunde hinsichtlich Veränderungen in der Arbeitsorganisation. Verschiedene Untersuchungen deuten darauf hin, dass KI in vielen Bereichen zu einer Neuorganisation von Aufgaben führt (u. a. Giering 2021; Kornwachs 2023; Lane et al. 2023). Das ergibt sich insbesondere aus dem Umstand, dass KI-Einsatz zu einer Automatisierung von Tätigkeiten führt, die bislang Menschen bearbeitet haben. Im Finanzsektor sagen das laut Lane et al. (2023) zwei Drittel der Arbeitgeber (66 %), im Produktionssektor sogar 72 %.

*Hauptmotive:
Mehr Produktivität,
aber kaum positive
Effekte für
Arbeitsverbesserung*

1.2.1 NUTZUNGSVERHALTEN

In Österreich nutzen nach aktuellem Stand laut Statistik Austria Erhebung 20 % der Unternehmen KI-Technologien (Statistik Austria 2024). Das ist ein deutlicher Sprung im Vergleich zum Vorjahr. 2023 waren es 11 % der österreichischen Unternehmen (Statistik Austria 2023). Der Zuwachs ist anteilmäßig vor allem auf größere Unternehmen zurückzuführen (Statistik Austria 2024): Der Anstieg bei Unternehmen mit 250 oder mehr Beschäftigten beträgt rund 15 %, bei Firmen mit 50-249 rund 12 %. In kleineren Unternehmen kommt KI dagegen deutlich seltener zum Einsatz. Das Nutzungsverhalten ist eher breit, d. h. nicht unbedingt an eine spezifische Art wirtschaftlicher Tätigkeit gebunden und im Dienstleistungssektor (vor allem mit IKT-Bezug) derzeit mit 23 % stärker ausgeprägt als im produzierenden Bereich (15 %). Die häufigsten Nutzungszwecke sind Texterkennung und -verarbeitung (65 %), Sprachgenerierung (41 %), Datenanalyse (34 %), Spracherkennung (29 %) und Prozessautomatisierung bzw. als Entscheidungshilfe (24 %), Bilderkennung und -verarbeitung (17 %) sowie KI-Systeme in autonom steuerbaren Maschinen oder Fahrzeugen (6 %).

*Steigende
Nutzungszahlen ...*

Im Vergleich zum Vorjahr erscheint der Anstieg groß. Allerdings relativiert sich das Bild bei Betrachtung der Nutzungszwecke beider Jahre. Hier zeigt sich, dass der Anstieg vor allem auf KI-Tools zur Texterkennung und -verarbeitung zurückzuführen ist. Hier gibt es einen Zuwachs – 2023 waren es 54 %, nun gaben 65 % der befragten Unternehmen diesen Nutzungszweck an. In anderen Bereichen gibt es dagegen sogar weniger Nutzung: etwa in der Prozessautomatisierung und als Entscheidungshilfe. 2023 waren das noch 32 %, 2024 sind es nur noch 24 %. Bei Automatisierten Datenanalysen mit maschinellem Lernen waren es 2023 noch 43 %. 2024 gaben nur 34 % an, KI zur Datenanalyse zu nutzen (Statistik Austria 2023/2024).

1.2.2 GRÜNDE GEGEN DIE NUTZUNG

Bei beiden Erhebungen wurde auch nach Gründen gegen die KI-Nutzung gefragt. Fast 80 % der österreichischen Unternehmen (79,1 %) nutzen keine KI-Technologien. Rund 20 % ziehen KI-Nutzung derzeit auch nicht in Erwägung. Die Gründe sind unterschiedlich: Am häufigsten wurde fehlendes internes Fachwissen genannt, gefolgt von rechtlichen Unklarheiten, Datenschutzbedenken, Datenvoraussetzungen, d. h. Schwierigkeiten bei Verfügbarkeit und Qualität von Daten, technischen Inkompatibilitäten, zu hohen Kosten, fehlendem Nutzen und ethischen Bedenken. Hier gibt es bei allen genannten Gründen 2024⁶ leichte Anstiege im Vergleich zum Vorjahr (Statistik Austria 2023), wie Tabelle 1 zeigt:

*... aber auch
wachsende Bedenken*

Tabelle 1: Gründe gegen KI-Nutzung (Quelle: Statistik Austria 2023/2024)

Gründe gegen KI-Nutzung	2023	2024
Fehlendes internes Fachwissen	7,2 %	12 %
Rechtliche Unklarheiten	5,2 %	9,8 %
Datenschutzbedenken	4,6 %	9,1 %
Datenvoraussetzungen	5 %	8,4 %
Technische Inkompatibilitäten	5 %	7,4 %
zu hohe Kosten	3,7 %	5,5 %
fehlender Nutzen	3 %	5,4 %
ethische Bedenken	2,5 %	4,6 %

Wie aus diesen Untersuchungen hervor geht, gibt es hohe Erwartungen an den Technologie-Einsatz und zugleich unterschiedliche Gründe für die bewusste Nicht-Nutzung. Diese sind zwar nicht stark ausgeprägt, deuten aber bereits auf einige Probleme und Herausforderungen hin, auf die in weiterer Folge noch genauer eingegangen wird (siehe Abschnitt 5).

⁶ www.statistik.at/fileadmin/pages/285/IKTEinsatzinUnternehmen16102024DE.ods.

2 GRUNDLAGEN UND FUNKTIONSWEISE

2.1 ZUM BEGRIFF KI

Der starke mediale Fokus auf dem Themenkomplex KI erweckt den Eindruck, es handle sich um ein neuartiges Phänomen. Tatsächlich hat KI bereits eine lange Geschichte. Der erste Chatbot namens ELIZA wurde bereits 1966 von Joseph Weizenbaum entwickelt, der damit auch auf die Grenzen und Probleme von KI-Systemen hinweisen wollte (Weizenbaum 1976). Die Ursprünge von KI reichen aber noch viel weiter zurück. Als „Geburtsstunde“ gilt ein Workshop am Dartmouth College in New Hampshire (USA) Mitte der 1950er Jahre. Dort wurde von renommierten Computerwissenschaftlern, u. a. den Mathematikern John McCarthy und Claude Shannon, KI offiziell als Forschungsdisziplin begründet⁷. Die Kernidee war es, Aspekte menschlicher Intelligenz wie Lernen, logisches Denken, Problemlösung und Entscheidungsfindung mit mathematischen Methoden formal abzubilden, zu simulieren und zu imitieren (McCarthy et al. 1955; Buchanan 2005; Verdegem 2021). Damals gängige Definitionen von KI beschrieben es als „science and engineering of making intelligent machines, especially intelligent computer programs“ (McCarthy 2007), das mit der Entwicklung intelligenter Computersysteme befasst ist, also Systemen, die über Merkmale verfügen, die mit intelligentem menschlichem Verhalten assoziiert werden (Barr et al. 1981).

Die ursprünglichen Visionen von der Entwicklung denkender Maschinen und Systemen mit Intelligenz, die bis heute nicht realisierbar sind, waren stark von deterministischen Weltanschauungen geprägt und wurden von KI-Forschern selbst immer wieder als ideologisch kritisiert (vgl. u. a. Weizenbaum 1976; McDermott 1981; Dreyfus 1992). Informatik-Pionier Alan Turing bezeichnete die Frage, ob Maschinen je denken können schon früh als zu irrelevant, um sie ernsthaft zu diskutieren (Turing 1950). Die klassischen Visionen zu künstlicher Intelligenz sind zwar nicht verschwunden, wurden im Lauf der Jahre jedoch weitgehend abgelöst von pragmatischeren Zugängen. Hierbei spielen vor allem maschinelle Lernverfahren zum Aufbereiten und Analysieren großer Datenmengen eine zentrale Rolle (siehe dazu genauer Abschnitt 2.3).

Zeitgemäßer lässt sich KI kurz gefasst als Teilgebiet der Informatik verstehen, die durch den Einsatz digitaler Systeme und Anwendungen zum Ziel hat, menschliche Fähigkeiten (wie Informationsverarbeitung, logisches Denken, Lernen, Planen etc.) mit mathematischen und statistischen Verfahren zu imitieren und automatisiert zu unterstützen. In der Praxis ist KI also ein Sammelbegriff für Technologien, die Menschen bei der Durchführung von unterschiedlichen Tätigkeiten, die kognitive Fähigkeiten und Denkleistung erfordern, automatisiert nutzen können. Die internationale Organisation für Normung (ISO) definiert KI als „a technical and scientific field devoted to the engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives“ (ISO 2022). Die Definition der KI-Verordnung

KI hat eine lange Geschichte

Klassische KI-Entwicklung war deterministisch geprägt

Zeitgemäße Definitionen von KI

⁷ www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

der EU⁸ knüpft hier an, erwähnt jedoch auch explizit Autonomie und Beeinflussung: „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können“ (Art 3 Z 1 EU-KI-VO).

Bevor die Funktionsweise von KI in Abschnitt 2.3 genauer erläutert wird, bietet der nächste Abschnitt einen kurzen Überblick darüber, wie die gesellschaftliche Bedeutung von KI eingeordnet werden kann.

2.2 WIE VERÄNDERT KI DIE GESELLSCHAFT?

Dass der Einsatz von KI-Systemen auf längere Sicht erhebliches Transformationspotenzial mit sich bringt, das Arbeit verändert und breitenwirksam unterschiedlichste Gesellschaftsbereiche durchdringt, ist unbestritten. Wie weitreichend diese Veränderungen im Konkreten tatsächlich ausfallen ist die Gretchenfrage vieler Untersuchungen. Zur Einordnung des soziotechnischen Phänomens KI spielt auch die soziotechnische Genese, also ihre Entstehungsgeschichte und damit verbundene Entwicklungen eine wichtige Rolle. KI ist kein neues Phänomen, sondern bezeichnet seit einigen Jahren ein breit sichtbar gewordenes Konvolut von verschiedenen Technologien als Teil der Digitalisierung. Abseits der über 70-jährigen Geschichte von KI als Forschungsfeld haben spätestens ab dem Jahr 2013 – mit dem Ausruf des Paradigmas von Big Data und Datafizierung⁹ – viele der heute mit KI konnotierten Themen bereits an Relevanz gewonnen. Algorithmen und datengetriebene Geschäftsmodelle sind seitdem auch in der breiteren Öffentlichkeit zum Thema geworden. Die Neuartigkeit des Begriffs für Laien kann mit ein Grund für die teilweise Verklärung der Bedeutung algorithmischer Systeme sein. Einfach formuliert ist Algorithmus schlicht ein Fachbegriff für eine Abfolge formalisierter Arbeitsschritte. KI-Systeme nutzen komplexe Algorithmen, aber nicht jeder Algorithmus hat etwas mit KI zu tun. Ein Algorithmus ist daher auch nicht mit KI gleichzusetzen, sondern schlicht eine mittels mathematischer Methoden formalisierter Ablauf von Tätigkeiten, um ein bestimmtes Ziel zu erreichen. Auch eine Ampelschaltung oder ein Kochrezept lassen sich als Algorithmus abbilden und algorithmische Systeme formalisieren die Verarbeitung von Daten auf unterschiedliche Weise. Ein einfaches Beispiel ist ein Sortiervorgang von Daten nach bestimmten Kriterien. Die Frage nach den konkreten Auswirkungen algorithmischer Systeme auf bestimmte Bereiche ist daher letztlich eine Frage der expliziten Formalisierbarkeit von Arbeitspraktiken über digitale Daten, ihren Grenzen, und den damit einher gehenden soziotechnischen Effekten.

KI ist kein neues Phänomen, sondern Teil des digitalen Wandels der Gesellschaft

⁸ eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=OJ:L_202401689.

⁹ de.wikipedia.org/wiki/Datafizierung.

Ähnliches trifft auf die gesellschaftliche Auseinandersetzung mit KI-basierten Technologien als soziotechnisches Phänomen zu. Die Frage der Formalisierbarkeit gilt umso mehr auch für KI, allerdings mit anderen Automatisierungsgraden (Strauß 2018). Der Begriff KI an sich ist widersprüchlich und suggeriert, es gäbe so etwas wie intelligente Computersysteme, die über vergleichbare oder sogar höhere Intelligenz verfügen könnten wie Menschen. Eine Diskussion über die Realisierbarkeit dieser alten und immer noch weitgehend ungeklärten Frage nach dem Wesen von Intelligenz würde den Rahmen dieser Studie weit sprengen. Dennoch sei angemerkt, dass die Mystifizierung des Begriffs KI einen seriösen Umgang mit den Auswirkungen KI-basierter Technologien erschweren, die letztlich neue Formen der Automatisierung in allen möglichen Gesellschaftsbereichen und Arbeitsformen implizieren. Im Sinne der Genese der angeführten technologischen Entwicklungen kommt mit KI eine weitere Ebene der Automatisierung hinzu (siehe Abschnitt 4). Das kann ein Teilschritt in einem Arbeitsprozess ebenso sein, wie ein automatisiertes System zur Entscheidungsunterstützung. Neuartig sind sowohl die Automatisierungsgrade, als auch die davon betroffenen Arbeitsformen. Wie eingangs erwähnt, hat der Einsatz KI-basierter Technologien zusehends Folgen für die Wissensarbeit (siehe Abschnitt 3), hier breit verstanden als Tätigkeiten, die nicht primär körperliche, sondern kognitive, geistige Anstrengungen erfordern. Die Effekte von KI-Systemen auf Arbeitsprozesse hängen wesentlich von ihrer Funktionsweise ab, die im nächsten Abschnitt erläutert wird.

Neue Fragen zur Formalisierbarkeit von Arbeitspraktiken durch KI-Einsatz

Die Mystifizierung von KI erschwert einen seriösen Umgang mit der Technologie

2.3 FUNKTIONSWEISE VON KI

Es gibt verschiedene Anwendungsfelder und Formen von KI, die in unzähligen Bereichen bereits zum Einsatz kommen. Die klassischen Bereiche von KI sind Computerlinguistik (Natural Language Processing – NLP), Wissens- Modellierung und -Repräsentation, Logik und automatisierte Entscheidungsfindung, maschinelles Lernen, Bild- und Sprachverarbeitung und Robotik (vgl. Russell/ Norvig 2010). Heute kommen diese Formen auf unterschiedliche Weise zum Einsatz. Zudem kommt mit sogenannter generativer KI (genAI) noch ein Bereich hinzu, der aufgrund der Bekanntheit (vor allem durch intensive Medienberichte zu diversen Tools) aktuell sehr weit verbreitet ist. Auch genAI und ihre technischen Grundlagen (Generative Adversarial Networks – GANs) sind nichts Neues. Aber in den letzten Jahren gab es deutliche Verbesserungen ihrer Funktionalität und Anwendbarkeit (Leible et al. 2024). Das ist mit ein Grund für den mittlerweile hohen Bekanntheitsgrad. Generative KI-Anwendungen können auf unterschiedliche Weise Daten und Inhalte (va. Text, Bild, Audio, Video) aufbereiten und in neuartiger Form darstellen. Typischerweise bieten sie mithilfe großer Sprachmodelle (large language models – LLMs) die Möglichkeit zur Interaktion in natürlicher Sprache in Form von integrierten Chatbots. D. h. die Eingabe (Prompts) kann ohne technisches Fachwissen und Programmierkenntnisse erfolgen. Dadurch sind diese KI-Anwendungen sehr niedrigschwellig in der Nutzung. Die hohe Frequenz von Medienberichten und die anhaltende Aufmerksamkeit auf dem massiv umworbenen Produkt ChatGPT von OpenAI verstellt etwas den Blick darauf,

Generative KI ist nur eines unter vielen Anwendungsfeldern

dass es viele verschiedene generative KI-Anwendungen gibt.¹⁰ Zudem ist das trotz hoher Bekanntheit auch nur eine unter vielen unterschiedlichen Anwendungen (siehe genauer dazu Abschnitt 3).

KI-basierte Technologien können je nach Anwendungsfeld sehr unterschiedlich sein. Im Wesentlichen bedeutet KI-Einsatz aber – kurzgefasst – die automatisierte Aufbereitung und Analyse von Daten. Grundsätzlich eignen sich KI-basierte Technologien besonders zur Mustererkennung, zur strukturierten Darstellung von Zusammenhängen sowie der Generierung von Inhalten auf Basis neuartiger Kombinationen von bestehenden Daten. Trotz vieler Varianten und Besonderheiten lässt sich im Kern ihre grundlegende Funktionsweise stark vereinfacht als typisches Prozess-Modell beschreiben:

Grundlegende Funktionsweise als Prozess erklärt

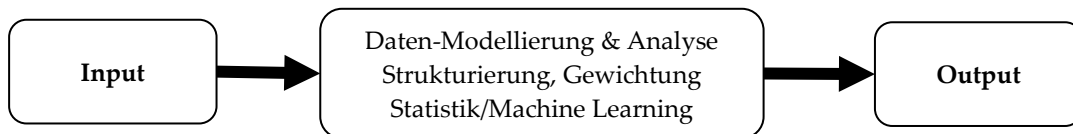


Abbildung 1: KI als vereinfachtes Prozess-Modell

Der Input umfasst eine bestimmte Menge an Daten (Eingabedaten), die dann mit maschinellen Lernverfahren aufbereitet, also modelliert, strukturiert, gewichtet und analysiert werden. „Lernen“ bedeutet hier automatisierte Datenanalysen mittels mathematisch-statistischer Verfahren, etwa um Prozesse zu optimieren oder Muster in großen Datenmengen zu erkennen und diese strukturiert darzustellen. Der Output, also das von KI produzierte Ergebnis ist also die anhand der Datenanalyse aufbereitete Information (auf Basis der Ausgabedaten). Diese kann je nach Anwendungsfeld unterschiedlich sein (z. B. eine Chatbot-generierte Antwort, ein KI-generiertes Bild, eine Daten-Visualisierung etc.).

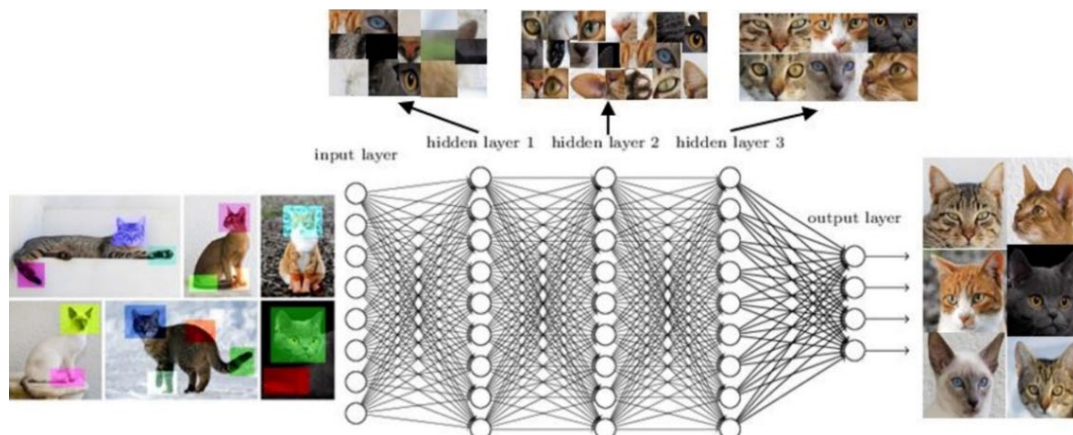


Abbildung 2: Beispiel einer KI-Bildererkennung mit Deep Learning (Strauß 2018)

Zur Veranschaulichung zeigt Abbildung 2 den relativ einfachen Ablauf einer KI-Anwendung zur Erkennung von Katzen, die auf Deep Learning (DL), einer hoch-

¹⁰ Z. B. Mistral, Claude, Perplexity, Huggingface, Gemini, Stable Diffusion oder auch diverse Bildgeneratoren wie Craiyon, Midjourney, Firefly und noch viele mehr.

komplexen ML-Variante basiert (siehe genauer Abschnitt 2.3.1), wobei künstliche neuronale Netzwerke zum Einsatz kommen. Dieses Verfahren zerlegt die Input-Daten (hier Katzenbilder) in verschiedene Abstraktionsebenen (am Bild die „hidden layers“), um daraus schrittweise Muster abzuleiten, die dann im Ergebnis eine qualitativ verwertbare Aufbereitung der Daten ermöglichen. Am Beispiel „lernt“ der DL-Algorithmus, verschiedene Katzenrassen voneinander zu unterscheiden.

2.3.1 ÜBERBLICK ZU MACHINE LEARNING VERFAHREN

KI basiert auf dem Einsatz komplexer Algorithmen, um Daten zu erfassen, zu analysieren und so aufzubereiten, dass daraus Erkenntnisse abgeleitet werden können. Das sind sogenannte Machine Learning (ML)-Verfahren, die heute in unterschiedlicher Form in vielen KI-Systemen verwendet werden. Als Forschungsfeld sucht ML u. a. nach Ansätzen, um die Leistung von KI-Systemen automatisiert zu verbessern sowie universalen Lern-Regeln für bestimmte Vorgänge (Mitchell 2006; Russell/Norvig 2010; Domingos 2012; LeCun 2015; Goodfellow et al. 2016). Im Idealfall „lernt“ ein KI-System durch Training des Datenmodells aus Beispieldaten, auch neue, nicht bereits trainierte Daten korrekt einzuordnen. Der Begriff des Lernens ist hier aber – ähnlich wie jener von künstlicher Intelligenz – irreführend und nicht mit menschlichem Lernen vergleichbar. Maschinelle Lern-Algorithmen sind mathematisch-statistische Verfahren, die Daten gewichten und Wahrscheinlichkeiten berechnen.

Es gibt eine Reihe unterschiedlicher ML-Verfahren (vgl. Mitchell 2006; Russell/Norvig 2010; Domingos 2012; Chibanguza et al. 2022), die oftmals miteinander kombiniert werden. Wichtig ist zunächst der Unterschied zwischen a) *überwachtem Lernen* (supervised learning) und b) *unüberwachtem Lernen* (unsupervised learning). Bei a) sind die Eingangs- und Ausgangsdaten bekannt, d. h. der Lern-Algorithmus basiert auf vorher festgelegten Kriterien und bekannten Trainingsdaten. Das zu erreichende Ziel des Lernverfahrens ist vorab festgelegt. Überwacht wird hier, inwieweit die produzierten Ergebnisse mit bekannten validen Daten übereinstimmen. Auf dieser Grundlage kann dann ein für einen bestimmten Anwendungskontext nutzbares Datenmodell erzeugt werden. Ein praktisches Beispiel wäre die Erkennung Handschriften historischer Dokumente: Das KI-System lernt mit validen Trainingsdaten, Handschriftmuster zu erkennen, wodurch sich in Folge dann Dokumente über diese Muster automatisiert auslesen lassen.

Bei b), unüberwachtem Lernen gibt es keine vorher festgelegten Kriterien (wie z. B. nach Kategorien vorstrukturierte Datensätze oder Label) und Lern-Ziele. Dieses Verfahren dient u. a. zur Analyse großer Datenmengen, um weitgehend selbstständig Muster zu erkennen, Datensätze zu strukturieren und Merkmale zuzuweisen. Das bringt gegenüber a) Effizienzgewinne und lässt sich auch auf unstrukturierte Datenbestände anwenden. Wesentlicher Nachteil ist jedoch die mitunter geringere Qualität und Überprüfbarkeit der Ergebnisse. Unüberwachtes Lernen wird u. a. genutzt, um Anomalien und Auffälligkeiten in großen Datensätzen zu erkennen. Das kann z. B. zur Erkennung von Betrug oder Angriffen auf IT-Systeme eingesetzt werden oder in der Medizin oder Genetik zur Eingrenzung von Krankheitsmustern und vieles mehr.

Deep learning (DL) ist eine Spezialform von b) und das bislang komplexeste ML-Verfahren. DL basiert auf künstlichen neuronalen Netzen (KNNs) mit mehreren Abstraktionsebenen. Das Konzept von KNNs orientiert sich am

Machine Learning Algorithmen sind mathematisch-statistische Verfahren

Wichtige Unterschiede, Stärken und Schwächen gängiger ML-Verfahren

biologischen Konzept von Neuronen und imitiert deren Struktur (LeCun et al. 2015; Goodfellow et al. 2016). DL-Verfahren sind eine Variante des unüberwachten Lernens, die u. a. weitgehend selbstständig Merkmale aus Rohdaten ableiten können (wie z. B. oben veranschaulicht, um aus unstrukturierten Katzenbildern, Katzen nach Merkmalen zu sortieren). DL basiert auf modularen Strukturen, die ein hohes Maß an Flexibilität bieten und Vorteile bei der Verarbeitung großer unstrukturierter Datensätze bringen. Dementsprechend dynamisch ist auch das Verhalten von DL-basierten KI-Systemen. DL wird u. a. für Vorhersagen und Klassifikationen eingesetzt. Beispielsweise in der Bild- oder Spracherkennung oder auch für komplexe Vorhersagen wie künftig voraussichtlich auch in der Meteorologie¹¹ (LeCun 2015; Goodfellow et al. 2016). DL braucht aber sehr große Datenmengen, viel Rechenleistung und hat Nachteile wie hohe Abhängigkeit zu Datenqualität, mangelnde Interpretierbarkeit und hohe Intransparenz, weil DL-Algorithmen aufgrund ihrer Dynamik auch ihre Funktionsweise auf nicht nachvollziehbare Weise verändern können (Goertzel 2015; Strauß 2018/2021a).

Eine weitere gängige ML-Varainte ist das *Reinforcement learning* (RL). RL bedeutet verstärkendes oder bestärkendes Lernen und ist eine Art semi-überwachter Ansatz. Mit Bestärkung ist gemeint, dass der Algorithmus aus Zustandsbeobachtungen und aus Feedback lernt. D. h., es fließen Daten aus der Umgebung und aus Interaktionsmustern des KI-Systems in den Datenaufbereitungsprozess über Feedbackschleifen ein. Den Rahmen bildet zunächst eine simulierte Umgebung. Diese besteht zwar aus festgelegten Regeln bzw. Parametern, die definieren, welche Aktionen mehr oder weniger vorteilhaft sind, um z. B. ein konkretes Ziel (z. B. Bearbeitung einer Anfrage) zu erreichen. Die Aktionen werden aber vom Software-Agenten selbst gesetzt und folgen keinem eindeutigen Schema. Sondern das KI-System „tastet“ sich quasi vor, um aus dem Feedback schrittweise Fehler zu reduzieren und zu besseren Ergebnissen zu kommen. Der Ansatz basiert auf Versuch und Irrtum, d. h. der Output basiert nicht auf vorab klaren Regeln für eine möglichst korrektes Ergebnis. Ein typisches Beispiel sind generative KIs (z. B. ChatGPT oder Perplexity), die über Texteingaben (Prompts) in natürlicher Sprache Anfragen bearbeiten. Hier werden u. a. RL-Verfahren genutzt, um aus Interaktionen mit Nutzer:innen zu lernen. Auch in der Robotik wird RL häufig eingesetzt, ebenso wie beim autonomen Fahren oder in der Logistik. RL-Verfahren bieten Vorteile für Aufgaben mit klarem Ziel, aber mehreren Möglichkeiten zur Erreichung wie etwa Wegoptimierung oder Prozesssteuerung. Sie benötigen in der Reinform wenig Daten, sondern lernen in einer simulierten Umgebung mit klaren Parametern aus Aktion und Reaktion (z. B. Nutzereingaben). Ein simples Beispiel wäre die Suche nach dem Weg aus einem Labyrinth. Nachteile sind u. a. hoher Zeit- und Rechenaufwand sowie Fehleranfälligkeit. Nicht geeignet sind RL-Verfahren u. a. bei sicherheitskritischen Anwendungen, wo Trial-and-Error kaum sinnvoll sind. In der Praxis gibt es unterschiedliche Ansätze (z. B. Q-Learning oder Deep RL) und RL wird meist mit anderen ML-Verfahren kombiniert (bei generativen KIs z. B. u. a. DL und RL). (Russell/Norvig 2010; Sutton/Barto 2015; Chibanguza et al. 2022).

Jedes ML-Verfahren ist unterschiedlich, auch hinsichtlich der automatisierten Datenverarbeitung. Überwachtes Lernen hat einen geringeren Automatisierungsgrad als unüberwachtes, verstärkendes oder Deep Learning. Letzteres bietet

Jedes ML-Verfahren verarbeitet Daten anders und hat unterschiedliche Automatisierungsgrade

¹¹ www.zamg.ac.at/cms/de/geophysik/news/kuenstliche-intelligenz-verbessert-vorhersage-von-sonnenstuermen.

derzeit den höchsten Automatisierungsgrad, da es über selbstoptimierende Algorithmen oder Ähnliches verfügen kann. Reinforcement Learning interagiert in der Regel dynamisch mit der Umgebung, z. B. durch Sensoren und Feedback, um Daten zu sammeln. Deep Learning erfordert meist große Datenmengen und verfügt über höhere Feature-Engineering-Fähigkeiten, d. h. es kann auch mit Rohdaten arbeiten, die auf unterschiedliche Weise integrierbar sind. Beide Ansätze sind sehr dynamisch und können daher dazu neigen, unvorhersehbares Verhalten zu verstärken. Sie sind daher im Vergleich zu Methoden des überwachten Lernens, das auf vorher festgelegten Kriterien und relativ stabilen Trainingsdaten basiert, noch unberechenbarer. Ein zentraler Aspekt ist, dass nicht jeder ML-Ansatz für jeden Anwendungskontext geeignet ist. Dies ist Entwickler:innen und technischen Expert:innen meist klar, jedoch nicht unbedingt Anwender:innen oder Entscheidungsträger:innen, die mit KI-Systemen interagieren oder ihren Einsatz in Unternehmen planen. Um Probleme zu vermeiden, ist es daher auch wichtig, die Entscheidungsträger und Personen, die mit KI-Systemen interagieren, für die generelle Funktionsweise und die Besonderheiten von KI-Systemen und ML-Verfahren und damit verbundene Probleme in bestimmten Kontexten zu sensibilisieren (Strauß 2021a/b).

2.3.2 NUTZEN UND GRENZEN KI-GENERierter ERGEBNISSE

KI-basierte Technologien können auf vielfältige Weise Nutzen in unterschiedlichen Anwendungsfeldern bringen. Etwa bei der (teil-)automatisierten Gestaltung, Bearbeitung und Optimierung von Prozessen, der Modellierung, Strukturierung, Analyse und Aufbereitung von Daten, der Identifikation von Mustern und Anomalien in großen Datenmengen, zur Kombination unterschiedlicher Daten und Generierung zusätzlicher Informationen, zum Aufzeigen von Zusammenhängen aus großen Datenbeständen, um komplexe Probleme zu bewältigen und die Entscheidungsfindung unterstützen und vieles mehr. Ein grundlegender Nutzen ist es, die Datenaufbereitung und -Analyse effizienter und rascher zu gestalten. Datenbasierte Vorgänge können also erheblich beschleunigt werden. Das ermöglicht auch immer neue Anwendungen, nicht zuletzt in der Forschung, wo KI/ML teils schon seit längerem genutzt wird (siehe auch Abschnitt 3). Zusammenfassend lässt sich KI daher als exploratives Werkzeug zur Problembewältigung begreifen, um die Komplexität von Arbeitsprozessen beherrschbarer zu machen.

Allerdings stehen dem Nutzen auch Grenzen gegenüber. Denn der KI-Einsatz bringt auch mehr Komplexität mit sich. Sowohl aufgrund des zusätzlichen Aufwands für die Datenpflege, aber auch zur Interpretation, Überprüfung und ggfs. Korrektur von Fehlern. Die Aussagekraft KI-generierter Ergebnisse ist begrenzt. Die Interpretation und Nutzung erfordert daher ein Grundverständnis dafür, wie KI-Systeme ihre Ergebnisse produzieren. Die Qualität des Inputs hat naheliegendermaßen wesentlichen Einfluss auf die Qualität des Outputs und die Funktionsfähigkeit des KI-Systems. „Schlechte“, also falsche, qualitativ geringwertige, irreführende oder für einen Anwendungsfall schlicht unpassende Daten führen daher auch zu falschen oder fragwürdigen Ergebnissen. Das hat je nach Anwendung unterschiedlich gravierende Auswirkungen. Noch relativ harmlos sind mittlerweile bekannte Unzulänglichkeiten generativer KIs wie ChatGPT, die häufig fehlerhafte oder frei erfundene Ergebnisse liefern. Das Produzieren falscher oder irreführender Ergebnisse von KI-Systemen wird auch als „KI-Halluzination“ bezeichnet (vgl. u. a. Leible et al. 2024). Problematisch wird es vor allem dann, wenn die von

*KI als exploratives
Werkzeug zur
Problembewältigung*

*Begrenzte
Aussagekraft
KI-generierter
Ergebnisse*

KI-Systemen generierten Ergebnisse unhinterfragt als Entscheidungsgrundlage eingesetzt werden. Fehler oder Unzulänglichkeiten haben daher letztlich auch negative Auswirkungen auf die Entscheidungsfindung, wenn sie zu schlechten oder Fehl-Entscheidungen führen.

Entgegen gängiger Annahmen ist die Korrektheit von Ergebnissen kein Kriterium, das KI-Anwendungen tatsächlich erfüllen können. Denn was als korrekt gilt, hängt wesentlich von Anwendungsfall in bestimmten Kontexten und der Anwendbarkeit der Technologie für konkrete Aufgaben ab. Vereinfacht ausgedrückt produziert KI schlicht Ergebnisse anhand komplexer Berechnungen auf Basis von Datenmodellen. Wie brauchbar die Ergebnisse sind, hängt maßgeblich von der Qualität der Daten und ihrer Aussagekraft für den konkreten Anwendungsfall ab. Ein Datenmodell lässt sich als abstrakte Abbildung eines kleinen Auszugs der Wirklichkeit verstehen. Die Abbildung selbst sagt aber noch nichts über die Brauchbarkeit der Daten in einem bestimmten Anwendungskontext aus. Zum Beispiel kann ein Datenmodell, wie oben in Abschnitt 2.3 gezeigt, aus einer großen Menge an rohen Bilddaten verschiedene Kategorien bilden und etwa dazu genutzt werden, Katzen zu erkennen. Wie gut das funktioniert, hängt aber auch davon ab, wie und womit das Datenmodell trainiert wurde, also von der Qualität und Brauchbarkeit der Trainingsdaten. Am Beispiel in Abschnitt 2.3 erklärt bedeutet das: Wenn das KI-System auch mit einer bestimmten Menge an Bildern von Fröschen trainiert wird, dann würde es auch Frösche für Katzen halten und umgekehrt. Denn KI-Systeme können zwar Daten schnell und auswerten und aufbereiten. Zu Qualität und Validität von Ergebnissen führt das aber nicht automatisch. Qualitätsanforderungen wie Korrektheit der Ergebnisse für einen bestimmten Anwendungskontext sind daher rein durch KI-Einsatz nicht erfüllbar. Um überhaupt einen Nutzen aus KI-Technologie ziehen zu können, sind menschliche Kompetenzen, Wissen und Erfahrung wichtiger denn je. Nicht zuletzt, weil die Anforderungen an die Qualitätskontrolle gerade durch KI-Einsatz deutlich erhöht werden (siehe auch Abschnitt 5).

KI-Technologie kann auch bei vielen simpel anmutenden Aufgaben ineffizient und fehleranfällig sein. Beispielsweise bei der korrekten Erkennung von analog dargestellten Messwerten wie Uhrzeiten oder visuell dargestellten Daten (Zhang et al. 2024; Kemper 2024). KI-Modelle haben offenbar Probleme damit, relevante visuelle Daten auszulesen und irrelevante auszufiltern (Bastian 2024). Auch bei der Verarbeitung einer größeren Menge von Dokumenten, etwa um daraus schlüssige Zusammenfassungen zu generieren gibt es deutliche Schwächen (Li et al. 2024). Damit verbunden ist das sogenannte „Lost in the Middle“-Phänomen, das die Schwierigkeit großer Sprachmodelle (LLMs) beschreibt, Informationen in der Mitte von Texten zu erfassen. KI-generierte Inhalte basieren daher meist hauptsächlich auf einer Zusammenfassung von Anfang und -Ende eines Textes (Liu et al. 2024). Ähnliche Probleme wurden auch bei der Bildverarbeitung aufgezeigt (Wu et al. 2024). Eine grundlegende Schwäche von KI-Systemen ist daher die bislang mangelhafte Kontextualisierung – also die korrekte Erkennung und Einordnung von inhaltlichen Zusammenhängen. Daten werden also häufig falsch oder auf für einen Anwendungskontext nicht passende Weise ausgewertet oder dargestellt.

Qualität und Korrektheit sind keine Kriterien, die KI an sich erfüllen kann

Probleme der Datenaufbereitung und „Lost in the Middle“-Phänomen

Mangelhafte Kontextualisierung

Die Grenzen und derzeitigen Schwächen von KI-Technologien verdeutlichen, dass ihre Nutzung auch einige neue Anforderungen mit sich bringt, die erfüllt sein müssen, damit die Technologien funktionsfähig und zuverlässig einsetzbar sind. Das wirft auch Fragen hinsichtlich des erzielbaren Mehrwerts durch KI-Einsatz auf, insbesondere wenn KI-generierte Ergebnisse fehlerhaft und intransparent sind. Die korrekte Interpretation und Überprüfung von KI-generierten Ergebnissen ist für jede KI-Anwendung, die Output generiert, relevant und gewinnt daher mit wachsendem Technologie-Einsatz künftig noch stärker an Bedeutung.

*Interpretation
und Überprüfung
KI-generierter
Ergebnisse immer
wichtiger*

3 WISSENSARBEIT UND KI: EIN WIDERSPRUCH?

Die Schaffung, Verwaltung und Verbreitung von neuem Wissen sind zentrale Bestandteile von Wissensarbeit. Sie ist ein wesentlicher Faktor für Produktivität und Innovation und hat eine stabilisierende Wirkung auf betriebliches Management (Drucker 1999; 2001). In der Wissensarbeit wird relevantes Wissen als Ressource nicht nur genutzt, sondern ggfs. auch revidiert und laufend weiterentwickelt (Willke 1998). Sie umfasst also dynamische, sich veränderte Tätigkeiten, deren Durchführung insbesondere Verantwortung, Autonomie, kontinuierliches Lernen mit sich bringen (Drucker 1999). Die Tätigkeiten implizieren Lernen, Aneignen und Vermitteln von neuem Wissen. Insofern erscheint zunächst plausibel, dass KI erheblichen Nutzen bringen kann, um Wissensarbeit zu unterstützen. Inwieweit das zutrifft und welche Herausforderungen damit verbunden sind, wird im Folgenden näher diskutiert.

Als drei Hauptdimensionen von Wissensarbeit gelten Komplexität, Neuartigkeit und Autonomie (Kelter et al. 2009). Die Bewältigung komplexer, oft neuartiger Aufgaben impliziert Flexibilität und relativ autonome Arbeitspraktiken. Wissensarbeiter:innen haben daher i. d. R. Prozesshoheit und können die Formen und Inhalte von Arbeitstätigkeiten und deren Umsetzung weitgehend selbst gestalten. Diese Arbeitsautonomie und Gestaltungsfreiheit ist nicht nur Merkmal, sondern auch Erfordernis von Wissensarbeit. Denn sie umfasst vor allem das Bewältigen von komplexen Aufgaben, deren Anforderungen unterschiedlich sein können, und die daher nicht immer nach dem gleichen Schema bearbeitbar sind. Ein zentraler Wesenszug von Wissensarbeit ist daher, dass sie viele komplexe Aufgaben umfasst, die nicht ohne weiteres formalisierbar sind. Wissensarbeit ist weitgehend ergebnisoffen und kaum standardisierbar. Das steht im Gegensatz zu Routinetätigkeiten, bei denen etablierte Lösungswege bereits weitgehend optimiert sind und wiederholt für vorab bekannte Aufgaben angewendet werden. Problemlösungskompetenz ist deshalb eine wichtige Grundvoraussetzung für Wissensarbeit (Drucker 1999, Hube 2005; Kelter et al. 2009).

Allerdings zeigt sich, dass KI primär repetitive Tätigkeiten durchführen kann und zumindest derzeit nicht, oder nur sehr eingeschränkt für komplexe Aufgaben geeignet ist. Vielmehr können KI-Technologien Routineaufgaben unterstützen, die sich formal abbilden und somit KI-automatisiert bearbeiten lassen (u. a. Albrecht/Kellermann 2020; Chibanguza 2022; Lane et al. 2023; OECD 2023; Stowasser 2023; Interviews a-j). Gerade bei komplexeren Aufgaben ist Formalisierbarkeit nur schwer möglich. Zudem kann der KI-Einsatz die Komplexität sogar zusätzlich erhöhen. Das gilt insbesondere für den zusätzlich entstehenden Arbeitsaufwand zur Überprüfung und Qualitätskontrolle KI-generierter Ergebnisse. Erhoffte Vorteile von KI gerade in der Wissensarbeit sind daher nicht so einfach realisierbar, wie erwartet wird.

Routineaufgaben sind Tätigkeiten, deren Ablauf weitgehend gleichbleibend ist. Komplexe Aufgaben haben dagegen keine klaren, vordefinierten Lösungswege. Die Annahme, Routineaufgaben wären mit KI automatisierbar, klingt zunächst schlüssig. Allerdings impliziert das eine formalisierte und isolierte Sicht auf Arbeitsabläufe, die es der Praxis in dieser Form selten gibt, gerade wenn es um eine

*Wissensarbeit
als Stabilisator
im betrieblichen
Management*

*Teil von Wissensarbeit
ist das Bewältigen
komplexer Aufgaben*

*KI eignet sich für
repetitive Tätigkeiten,
weniger für komplexe*

*Der Wert menschlichen
Erfahrungswissens*

Kombination aus einfacheren und komplexeren Tätigkeiten geht (vgl. Albrecht/Kellermann 2020). Diese Kombination ist nicht trivial und auch ganz einfache Routinetätigkeiten sind in Arbeitsprozesse und interorganisationale Arbeitspraktiken integriert, die nicht ohne weiteres gänzlich oder teil-automatisierbar sind. Das gilt besonders für die Wissensarbeit, die primär viele komplexere Tätigkeiten umfasst, die eben nicht standardisierbar und damit auch nicht automatisierbar sind. Die dabei auch anfallenden Routinetätigkeiten sind schwer isoliert von anderen Aufgaben trennbar. Auch deshalb, weil hier auch die wertvolle Ressource des menschlichen Erfahrungswissens ein zentrales Bindeglied zwischen unterschiedlichen Arbeitsabläufen und Arbeitsorganisation darstellt.

Eine klassische Herausforderung von Wissensarbeit ist es, das Zusammenspiel von personalem und organisationalem Wissen zu verstehen und auszubalancieren (vgl. Willke 1998). D. h. bislang waren die beiden Hauptakteure in diesem Zusammenspiel zum einen die individuellen Arbeitskräfte und zum anderen die Arbeitsorganisation. Der Einsatz von KI verändert dieses Verhältnis und macht es zunächst komplexer, weil eine Art soziotechnischer Entität dazu kommt, deren Funktion in vielen Fällen noch nicht klar definiert ist. Bei KI-Einsatz ist mittelfristig mit Änderungen und Neustrukturierungen in der Arbeitsorganisation zu rechnen. Das wiederum hängt sehr stark davon ab, ob und wie KI überhaupt in Arbeitsabläufe integrierbar ist (siehe unten). Das kann zu betrieblichen Spannungen führen und der erzielbare Nutzen hängt auch stark davon ab, welche Erwartungen an den KI-Einsatz bestehen. Eine realistische Erwartung ist es, KI als mögliches Unterstützungswerkzeug zu begreifen, dessen Nützlichkeit wesentlich vom Anwendungsfall abhängt. KI-Anwendungen können bei der Bearbeitung komplexerer Aufgaben z. B. nützlich sein, um etwa alternative Lösungswege bestehender Probleme aufzuzeigen oder auch neue Zusammenhänge und Muster in großen Datenmengen zu entdecken und vieles mehr. Demnach lässt sich die Technologie als Werkzeug für datenbasierte Entscheidungsunterstützung betrachten, dass bei Verknüpfung, Generierung und Bereitstellung von Wissen unterstützen kann. Allerdings als exploratives Werkzeug, das z. B. kreative Stimulanz bieten kann, nicht jedoch unbedingt ein korrektes Ergebnis oder einen optimalen Lösungsweg offenbart. Denn die Produktivität von Wissensarbeit hängt nicht nur vom Output, also den Ergebnissen ab, sondern auch von der Qualität der Ergebnisse (Drucker 1999). Dasselbe gilt für den produktiven Nutzen von KI-basierten Technologien.

Bei genauerer Betrachtung der Frage, wie KI die Wissensarbeit verändert, offenbart sich ein zentraler Widerspruch bzw. ein Paradoxon: Wenn sich Wissensarbeit vollständig automatisieren ließe, dann würde sie zu einer Routinetätigkeit werden und wäre folglich keine Wissensarbeit mehr. Zudem setzt Wissensarbeit auch einen gewissen Grad an Autonomie und Selbstbestimmung in der Arbeitsgestaltung voraus. Wie bereits ausgeführt, ist Wissensarbeit daher nicht in Form von eindeutig festlegbaren Arbeitsschritten programmierbar. Das gilt nach Drucker (1999, S. 85) auch umgekehrt: „(...) in knowledge work, the task does not program the worker“. Allerdings stellt der Einsatz von KI diese Aussage heute auf den Prüfstand. Es stellt sich die Frage, ob der Einsatz von KI-Systemen nicht auch dazu führen kann, dass Wissensarbeiter:innen durch algorithmisch vordefinierte Abläufe bis zu einem gewissen Grad „programmiert“ werden, wodurch sie an Handlungsfreiheit verlieren. KI-Automatisierung kann daher bei Überformalisierung von Tätigkeiten auch zu einer Einschränkung der Autonomie und

*Knackpunkt:
Sinnvolle
Integrierbarkeit von
KI in Arbeitsabläufe*

*Widerspruch zwischen
Automatisierung und
Wissensarbeit*

Handlungsfähigkeit¹² führen. Das wirft wiederum neue Fragen zur stabilisierenden und innovationsfördernden Wirkung von Wissensarbeit in Unternehmen auf.

Der Umgang mit KI-basierten Technologien selbst und den damit produzierten Ergebnissen ist mit einer Reihe von Problemen und Herausforderungen verbunden, die sich auch als „wicked problem“ beschreiben lassen (Strauß 2021b). Nachdem ein Wesenszug von Wissensarbeit das Bewältigen komplexer Probleme ist, sind die Aufgaben im Umgang mit den Herausforderungen von KI auch Bestandteile der Wissensarbeit. D. h. KI-basierte Technologien verändern Wissensarbeit insofern, als sich die Anforderungen im Umgang damit in der Wissensarbeit verändern: Wissensarbeiter:innen brauchen demnach auch neues Wissen darüber, ob und wie KI für bestimmte Aufgaben nutzbar ist (oder nicht). Die Schaffung und Aneignung dieses Wissens ist eine zentrale Herausforderung im Umgang mit KI-basierten Technologien in Wirtschaft und Gesellschaft.

Bei der Wissensproduktion können KI-Anwendungen in vielerlei Hinsicht unterstützen. Allerdings nicht durch Generierung neuen Wissens. Denn folgt man der Definition von Davenport et al. (1998, S. 43): „Knowledge is information combined with experience, context, interpretation and reflection“, dann ist Wissen die Kombination aus Information, Erfahrung, Kontext, Interpretation und Reflexion. Demnach kann KI nie Wissen generieren, weil es weder über Erfahrung, noch Reflexionsfähigkeiten verfügt und die Interpretation nur nach formalisierten, algorithmisch abbildbaren Kriterien möglich ist, nicht aber nach Sinnhaftigkeit und kontextueller Bedeutung. Es kann also nur Information in bestimmten Kontexten liefern, jedoch ohne Aussagekraft bezüglich Korrektheit, Sinnhaftigkeit und Nutzbarkeit dieser Information. Das Verstehen und Interpretieren KI-generierter Information sowie ihre Einbettung in soziotechnische Anwendungskontexte wird daher Teil des Aufgabenspektrums von Wissensarbeit. Kurz gesagt: Moderne Wissensarbeit mit KI umfasst zusätzlich zur „klassischen“ Problemlösungskompetenz nun auch die Kompetenz zu erkennen, ob und wie KI-Einsatz bei einer Tätigkeit sinnvoll ist oder eben nicht (siehe Abschnitt 6).

Stabilität und Zuverlässigkeit sind wesentliche Anforderungen an KI-basierte Technologien (vgl. auch Hube 2005). Gerade hier ergeben sich aber noch erhebliche weitere Herausforderungen. Damit KI tatsächlich die Wissensarbeit unterstützen kann, muss es zuverlässige Ergebnisse produzieren. Genau das ist aber häufig nicht der Fall. Das hängt stark vom Anwendungsfeld ab. Im nächsten Abschnitt wird auf ausgewählte Anwendungsfelder näher eingegangen.

KI verändert die Anforderungen in der Wissensarbeit

KI kann kein neues Wissen schaffen, nur bestehendes anders aufbereiten

Kritische KI-Kompetenz wird zum Bestandteil moderner Wissensarbeit

¹² Im Englischen wird hierfür der Begriff Agency verwendet.

3.1 ÜBERSICHT ZU AUSGEWÄHLTEN ANWENDUNGSBEREICHEN

Es gibt unzählige Anwendungsfelder von KI und die Technologie findet sukzessive Einzug in unterschiedlichen Branchen. Aufgrund der Fülle kann hier nur ein kleiner Überblick geboten werden. Auf ausgewählte Bereiche wird im Anschluss genauer eingegangen.

Abseits des Hypes um generative KIs wird die Technologie schon vielfach seit längerem eingesetzt, etwa in Smartphones und diversen integrierten KI-Anwendungen, über ML-Verfahren in der Robotik und verschiedenen Fahrassistenzsystemen für Maschinen oder selbstfahrende Fahrzeuge, in Form von Chatbots und Sprachassistenten im Support-Bereich in Callcentern, oder vielfältigen Einsatzfeldern im IT-Sektor bei der Softwareentwicklung, Prozessoptimierung etc. Auch in der Fertigungsindustrie gibt es bereits seit längerem zahlreiche Anwendungen, z. B. in Produktionsbetrieben wie der Textilindustrie, wo KI die Qualitätskontrolle für technische Textilien oder bestimmter Bauteile unterstützen kann. Es gibt u. a. KI-Anwendungen, um Anomalien und Abweichungen von vordefinierten Qualitätsstandards zu erkennen. Auch zur Wartung und Instandhaltung bietet KI viele Möglichkeiten zur frühzeitigen Erkennung von Verschleiß und Wartungsbedarfen bei Maschinen und Anlagen (Chibanguza et al. 2022; OECD 2023; Stowasser 2023). Solche Anwendungen werden auch vorausschauende Instandhaltung (predictive maintenance) genannt und u. a. auch bei den Wiener Linien eingesetzt.¹³

KI kann effizient und schnell Daten analysieren und aufbereiten, Muster identifizieren und vieles mehr. Gerade dadurch hat die Technologie vielfältige Einsatzmöglichkeiten in Wissenschaft und Forschung. Ein weiterer Bereich, wo KI schon länger verwendet wird, ist die Akustik. Neben Anwendungen zur Klangoptimierung, Spracherkennung, -Synthese, Signalanalysen u. dgl. sind moderne Hörsysteme/Hörgeräte schon länger mit KI/ML-Funktionalität¹⁴ ausgestattet, etwa um Umgebungsgeräusche herausfiltern oder Stimmen zu verstärken. Zudem werden in der Schallforschung ML-Methoden vielfältig genutzt, um akustische Signale zu messen. Die Anwendungen reichen von Musikproduktion, Tontechnik, Sonografie bis zur Materialprüfung, um über akustische Sensorsysteme z. B. Verschleiß frühzeitig zu erkennen und kritische Bauteile zu überwachen¹⁵ (Tschoepe et al. 2017; Gerlach/Eisele 2022).

In der Physik werden seit langem ML-Verfahren genutzt, etwa für komplexe Datenanalysen und Simulationen in der Teilchenforschung¹⁶, oder in der Quantenphysik. In der Materialwissenschaft werden z. B. anhand KI-basierter Modelle Materialstrukturen und spezifische Eigenschaften untersucht, um neue

*Vielfältige
Anwendungsfelder*

*In Wissenschaft
und Forschung wird
KI teils seit langem
genutzt*

*Enorme Potenziale
in der Medizin*

¹³ [zukunftfindetstadt.blog/2021/01/26/predictive-maintenance-wie-wir-fahrweg-und-fahrzeug-zusammenbringen-wollen/](https://www.zukunftfindetstadt.blog/2021/01/26/predictive-maintenance-wie-wir-fahrweg-und-fahrzeug-zusammenbringen-wollen/).

¹⁴ www.prosurdis.ch/kuenstliche-intelligenz-ki-in-der-hoerakustik-eine-revolution-fuer-besseres-hoeren/.

¹⁵ www.vision.fraunhofer.de/de/technologien-anwendungen/technologien/akustische-qualitaetspruefung/intelligente-qualitaetssicherung-ki-analyse-akustische-sig-nale.html.

¹⁶ z. B. am ÖAW-Institut für Hochenergiephysik www.oew.ac.at/hephy/forschung/machine-learning.

Materialen zu entwickeln (Materialinformatik).¹⁷ In den digitalen Geisteswissenschaften (digital humanities¹⁸) werden mittels ML u. a. historische und literarische Dokumente analysiert, Handschriften erfasst, das kulturelle Erbe digital abgebildet¹⁹ sowie in unterschiedlichen Fachbereichen (z. B. Literatur- Sprach- und Musikwissenschaft) erforscht, um neue wissenschaftliche Erkenntnisse zu gewinnen. In der Archäologie unterstützt KI²⁰ u. a. bei der Analyse archäologischer Funde wie der Auswertung antiker Schrifttafeln, oder auch von Luftaufnahmen zur Erforschung von Mustern an der Erdoberfläche. Kürzlich wurden damit etwa weitere Nazca-Geolyphen²¹ in Peru entdeckt. In Biologie und Medizin hilft KI z. B. bei der Erforschung von Zellstrukturen, Proteinen und der Entwicklung neuer Wirkstoffe. 2020 gelang es z. B. mithilfe eines speziellen KI-Systems, neue Proteinstrukturen zu entdecken.²² 2024 wurde diese Entdeckung mit dem Chemie-Nobelpreis ausgezeichnet. Zudem gibt es unzählige Forschungsaktivitäten zu KI in der Medizin, auch in Österreich. Zum Beispiel wird an der Medizinischen Universität Wien²³ seit vielen Jahren u. a. an der Entwicklung von KI/ML Verfahren mit Schwerpunkten auf Tumorcharakterisierung, Radiomics und Holomics (speziellen Bildanalyseformen) geforscht. Hierbei werden Daten aus bildgebenden Verfahren (z. B. PET, CT und MRT)²⁴ mit anderen medizinischen Daten kombiniert, um Vorhersagemodelle für die Präzisionsmedizin zu entwickeln. 2024 hat die österreichische Akademie der Wissenschaften mit AITHYRA ein eigenes Institut für KI in der Biomedizin gegründet, um komplexe biomedizinische Zusammenhänge und Krankheiten zu erforschen, die zur Entwicklung neuer Ansätze für zuverlässigere Diagnosen und Therapieformen dienen.²⁵ Das sind nur einige Beispiele der vielfältigen Anwendungsformen, wo KI erheblichen Nutzen bringen kann. Im Folgenden werden ausgewählte Anwendungsbereiche näher erläutert.

3.1.1 MEDIZIN UND GESUNDHEIT

Hier sind KI-Anwendungen besonders vielversprechend. Das Spektrum reicht von KI-Unterstützung in der medizinischen Diagnostik, z. B. durch Datenanalysen zur frühzeitigen Erkennung von Krebs und anderen Krankheiten wie Alzheimer, KI-gestützten Labordiagnosen, Entwicklung von Medikamenten, automatisierte Unterstützung in der Administration, Dokumentation und Prozessmanagement in Gesundheitseinrichtungen, bis zur Vision personalisierter Präventivmedizin (Interviews f, j; Bhagwat 2018; Davenport/Kalakota 2019; Tschandl et

¹⁷ www.iwm.fraunhofer.de/de/geschaeftsfelder/fertigungsprozesse/materialinformatik.html.

¹⁸ z. B. am ÖAW-Institut für Digital Humanities
www.oew.ac.at/de/acdh/acdh-ch-home.

¹⁹ blog.dnb.de/ki-und-digital-humanities-in-bibliotheken-ein-erfahrungsaustausch-auf-werkstattebene/.

²⁰ www.archaeologie-online.de/blog/innovative-forschung-ki-in-der-archaeologie-5050/.

²¹ www.heise.de/news/KI-gestuetzte-Durchmusterung-entdeckt-Hunderte-neue-Nazca-Geolyphen-9958382.html.

²² rudolphina.univie.ac.at/revolutioniert-kuenstliche-intelligenz-die-forschung.

²³ www.meduniwien.ac.at/web/forschung/artificial-intelligence-und-machine-learning/.

²⁴ PET: Positronen-Emissions-Tomografie, CT: Computertomografie, MRT: Magnetresonanztomografie

²⁵ www.oew.ac.at/news/oew-und-boehringer-ingelheim-stiftung-etablieren-neues-institut-fuer-kuenstliche-intelligenz-in-der-biomedizin-in-wien-1.

al. 2020; Funer 2022; Pfannstiel 2022; Hiltawsky 2024). Erwartet und erhofft werden von KI u. a.: mehr Präzision bei der Diagnose und Behandlung von Krankheiten, maßgeschneiderte Medikationen und Therapieformen mit weniger Nebenwirkungen, bessere Krankheitsvorsorge und infolge weniger Kosten im Gesundheitswesen (vgl. u. a. Funer 2022; Pfannstiel 2022).

In der Praxis ist hier noch sehr vieles im Entwicklungsstadium, aber einige Anwendungen sind bereits experimentell im Einsatz. In manchen Bereichen, wie bei bildgebenden Verfahren teils auch schon darüber hinaus und mit klarem Praxisbezug (siehe unten). Vereinfacht gesagt lässt sich der Einsatz von KI in der Medizin meist in zwei Formen unterscheiden: Virtuell und physisch (Hamet/Tremblay 2017; Malik et al. 2019). Der erste Teil umfasst z. B. ML-Anwendungen zur Analyse und Aufbereitung von medizinischen Daten zur Unterstützung bei Diagnosen und Behandlungsentscheidungen. Die physische Form bezieht sich auf den Einsatz KI-basierter Technologien zur Überwachung und Analyse des Patientenzustands; aber auch der Robotik zur Unterstützung bei Operationen, bei körperlichen Therapieformen, intelligenten Prothesen und in der Altenpflege. Etwa gibt es bereits verschiedene Pilotprojekte²⁶ wie dem Einsatz von Exoskeletten, KI-basierter Sensoren zur Verletzungsprävention oder sprachgesteuerter Pflegedokumentation (Hiltawsky 2024).

Große Hoffnung liegt auf der effizienteren Gestaltung von Arbeitsabläufen in der medizinischen Praxis. Im Idealfall führt der Einsatz von klinischen Support-Systemen zu einer Optimierung und Entlastung der behandelnden Ärzte, um mehr Zeit für Patient:innen zu gewinnen. Ein experimentelles Beispiel aus der Intensivmedizin ist die Erkennung von Sepsis (Blutvergiftung): Sepsis ist unter den häufigsten Todesursachen. Diese lässt sich bei Intensivpatienten durch Überwachung und Analyse von Vitaldaten in Echtzeit bereits vor Eintreten der sichtbaren Symptome erkennen und damit frühzeitiger behandeln (Interview f; Garnacho-Montero/Martín-Loeches 2020). Eine solche Anwendung wurde auch an der Technischen Universität Wien entwickelt²⁷ Es gibt noch unzählige weitere Beispiele wie etwa KI-gestützte Analysen in der Pathologie²⁸ zur Laborunterstützung bei Biopsien, KI-Algorithmen zur frühzeitigen Erkennung von Vorhofflimmern in der Kardiologie²⁹ u. v. m.

Abseits des vielseitigen Anwendungsspektrums kann KI in der medizinischen Diagnostik aus praxisorientierter Sicht auf zwei zentrale Arten eingesetzt werden: Zum einen, um Ärzt:innen in einer Form zu unterstützen, die zu einer effizienteren Gestaltung ihrer Arbeitsabläufe führt. Zum anderen, damit verbunden, zusätzliche Informationen zu erhalten, die ohne KI-Einsatz nicht direkt verfügbar wären (z. B. Zusatzinformationen über Nebenwirkungen, Krankheitsverläufe, mögliche andere Krankheitsbilder etc.). Zum Beispiel bei einer Röntgenaufnahme: Wenn eine KI-Anwendung nicht nur das Bild, sondern bereits weitere, untersuchungsrelevante Auffälligkeiten im Bild markiert oder zusätzliche,

*Unterschiede
zwischen physischem
und virtuellem
KI-Einsatz*

*Arbeitsentlastung
und zusätzliche
Fachinformation
als Ziele*

²⁶ Pilotprojekte in Pflege: Exoskelette und KI
www.orf.at/stories/3273317/.

²⁷ www.tuwien.at/tu-wien/aktuelles/news/news/besser-als-menschen-kuenstliche-intelligenz-in-der-intensivmedizin.

²⁸ www.springermedizin.de/kuenstliche-intelligenz/pathologie/kuenstliche-intelligenz-in-der-pathologie-wie-wo-und-warum/26846538.

²⁹ www.heise.de/news/Neuronales-Netz-erkennt-Herzinfarkte-so-gut-wie-erfahrene-Kardiologen-4107042.html.

diagnoserelevante Informationen anzeigt, die aus einem klinischen Befund allein nicht ableitbar sind (Interview j).

Besonders vielversprechend sind KI-Anwendungen bei bildgebenden Verfahren als Teil der medizinischen Untersuchung. Diese gibt es bereits seit mehreren Jahren, ursprünglich begonnen in der Radiologie mit Lungenläsionsdetektion. Daraus sind weitere Anwendungen im Bereich Radiomics entstanden, einem Teilgebiet der medizinischen Bildverarbeitung und der Radiologie, das sich mit der Analyse von Bilddaten wie z. B. Röntgenaufnahmen befasst. Der Einsatz von maschinellem Lernen ist hier besonders aussichtsreich (vgl. Kocak et al. 2019; Interview j). Bildgebende Verfahren eignen sich auch als Praxisbeispiel, um zu verdeutlichen, was KI in der Diagnostik leisten kann: Etwa, um Vorgänge wie das Finden und Markieren von Auffälligkeiten in Aufnahmen automatisiert zu unterstützen. Ein KI-System kann dazu genutzt werden, Bilder beim Lungenröntgen effizienter auszuwerten, indem auffällige Muster erkannt und visuell markiert werden. Bislang muss dieser Vorgang noch oftmals händisch durchgeführt werden. D. h. ein Arzt markiert per Hand oder im Bildbearbeitungsprogramm nach eigenem Ermessen eine Auffälligkeit (z. B. bei Verdacht auf Geschwüre odgl.). KI kann diesen Vorgang unterstützen und beschleunigen, weil Erkennen und Markieren damit automatisierbar wird. Mediziner:innen erhalten dann zusätzlich ein vormarkiertes Bild, und werden so in ihrer Expertise unterstützt. Im Idealfall wird die Bildauswertung effizienter gestaltet und damit auch Diagnose und Behandlungsformen. Studien zufolge lässt sich mit KI-Anwendungen bei der Bildanalyse eine relativ hohe Genauigkeit erzielen, vergleichbar mit Fachpersonal, jedoch deutlich schneller (Liu et al. 2019).

Bei diagnostischen Untersuchungen kommen sogenannte Picture Archiving and Communication Systems (PACS), zum Einsatz. Die ersten PACS wurden bereits in den 1970er Jahren entwickelt, sind seit den 1990er Jahren zu einer Standardtechnologie geworden, die mit KI eine deutliche Weiterentwicklung erfahren³⁰. PACS sind Computersysteme zur Bildarchivierung, Bearbeitung und Kommunikation von Daten aus bildgebenden Verfahren, die vor allem in Radiologie und Nuklearmedizin, aber auch vielen anderen Bereichen wie Kardiologie, Endoskopie, Pathologie oder Mikrobiologie verwendet werden. Über PACS kann medizinisches Personal auf alle Bilddaten zugreifen und Arbeitslisten zu Patient:innen, etwa um zusätzliche Daten wie CT-Detailaufnahmen einer Patient:in genauer zu untersuchen. Ohne KI-Anwendung müssen dabei ca. 200 Bilder manuell gesichtet und analysiert werden. KI kann diesen Vorgang deutlich unterstützen, zum Beispiel bei der Krebsuntersuchung. Ein Teil der Untersuchung ist die Segmentierung des Tumors, z. B. eines Lungenkarzioms. Bei der Segmentierung werden inhaltlich zusammenhängende Regionen in der Bildgebung sichtbar gemacht. KI-Algorithmen können diesen Vorgang bereits im Hintergrund so automatisieren, dass der Arzt bereits den segmentierten Tumor sieht und im Detail analysieren kann.

Ein wesentlicher Vorteil von KI-Anwendungen kann also die Komplexitätsreduktion bei der Diagnose und in Folge der Behandlung sein. Etwa, wenn die KI-Anwendung das Analysefeld eines Lungenröntgens präzise eingrenzt. Dadurch können im Idealfall Effizienzgewinne erzielt, sowie Komplexität und der zeitliche Aufwand solcher Bildauswertungen reduziert werden, sodass Fachpersonal mehr

*Aussichtsreiche
KI-Unterstützung
bei Diagnosen mit
bildgebenden Verfahren*

*Komplexitätsreduktion
möglich, weil
Bilddaten mit KI
schneller auswertbar*

*Entscheidung bleibt
immer beim Menschen*

³⁰ www.postdicom.com/de/blog/what-is-pacs-and-how-does-it-work.

Zeit für die Interpretation der Diagnoseergebnisse und für Entscheidungen zur weiteren Behandlung gewinnt. Ob das gelingt, hängt wiederum vom Zusammenspiel menschlicher und technischer Einflussfaktoren ab (genauer dazu siehe Abschnitt 4). Die Entscheidung, wie mit den Ergebnissen verfahren wird, und was daraus für die Behandlung folgt, liegt aber auch bei KI-gestützter Auswertung immer bei den behandelnden Ärzt:innen. D. h. KI-Systeme entscheiden grundsätzlich nichts. Sie zeigen nur, ob und was z. B. in Bilddaten gefunden wurde. D. h. es gibt immer die Möglichkeit, KI-Auswertungen zu hinterfragen und weitere Analysen durchzuführen. Ein wesentlicher Vorteil von bildgebenden Verfahren ist hierbei die visuelle Überprüfbarkeit.

3.1.2 JOURNALISMUS UND MEDIEN

Datenanalysen und digitale Tools zur Unterstützung werden im Bereich Journalismus und Medien seit vielen Jahren genutzt. Vor dem KI-Hype noch unter synonymen Schlagworten wie „robot journalism“, „automated journalism“ oder Datenjournalismus (Dörr 2016; Dupuis 2016). Heute gilt KI gilt letztlich als weitere, sehr komplexe Variante des datengetriebenen Journalismus mit zusätzlichem Automatisierungsgrad (Schell 2022). KI wird hier als datenbasiertes Werkzeug genutzt, um Journalist:innen darin zu unterstützen, Daten zu strukturieren und darauf aufbauend journalistische Inhalte aufzubereiten.

Entlang der Wertschöpfungskette im Medienbereich können die Hauptanwendungen von KI in drei Bereiche unterteilt werden: Sourcing, Produktion und Distribution/Publishing. Sourcing umfasst u. a. verschiedene Tätigkeiten zur KI-basierten Datenaufbereitung, wie z. B. Analyse von Texten oder multimedialer Inhalte, um relevante Begriffe zu erkennen und einzuordnen, Transkription (speech-to-text), Sprachübersetzung, Dokumentenklassifikation oder auch Themenmonitoring. In der Produktion kann KI u. a. zur Unterstützung bei der Generierung von Texten (data-to-text), Bildern oder anderen Medienformaten, beim Editieren, Visualisieren und Anreichern von Inhalten, zur Förderung von Inklusion etc. verwendet werden. Der Bereich Distribution bezieht sich auf automatisierte Empfehlungen u. Personalisierung, Targeting, Platzierung und Verknüpfung von thematisch ähnlichen Inhalten, Verbreitung von Medieninhalten über soziale Media und andere Plattformen, Moderation und Medienbeobachtung (Krawarik et al. 2021; Bailer et al. 2022; BBC 2024). In allen drei Bereichen sind KI-Anwendungen bereits im Einsatz, aber vieles davon wird noch erprobt. KI-Automatisierungen in der Produktion oder Content-Generierung haben bislang noch eher experimentellen Charakter. Zudem bestehen hier auch mehr Vorbehalte aufgrund funktionaler Schwächen wie Fehleranfälligkeit, ethischer Bedenken und vieler ungeklärter Fragen (Schell 2023; Newman 2024; Interviews c, d, e).

Die Medienbranche macht bereits vielfältig Gebrauch von den Möglichkeiten, die KI bietet. In Deutschland erprobt u. a. der Axel Springer-Verlag den KI-Einsatz in journalistischen Prozessen. Dort soll KI etwa künftig redaktionelle Arbeit der Bildzeitung unterstützen, z. B. bei der Layoutierung, der Transkription von Interviews oder der Aufbereitung und Anpassung von Texten, Videos (z. B. schneiden, texten, kürzen, vertonen) oder anderen medialen Inhalten auch für die Distribution über soziale Medien (Wilkens 2023). In Österreich hat insbeson-

KI als Variante des Datenjournalismus

KI-Anwendungen entlang der medialen Wertschöpfungskette

Redaktionelle Arbeit wird zusehends mit KI unterstützt

dere die APA (Austria Presse Agentur) seit 2019 die Implementierung eines breiten Spektrums an automatisierten Tools kontinuierlich ausgebaut.³¹ Die Einsatzgebiete reichen von Datenanalysen bei der anlassbezogenen Berichterstattung, Textassistenten für routinemäßige Aufgaben bei der Inhaltserstellung, automatisierten Textvorschlägen für Kurzmeldungen, bis hin zu datenbasierten Berichten über verschiedene Ereignisse (Schell 2022). Für einen verantwortungsvollen und transparenten Umgang mit KI hat die APA auch eigene Leitlinie ausgearbeitet, die insbesondere auf ethische Grundsätze wie Achtung menschlicher Autonomie, Schadensverhütung, Fairness und Erklärbarkeit Bezug nimmt (APA 2023). Im ORF (Österreichischer Rundfunk) wurde die kürzlich ausgezeichnete³² Software AIDitor entwickelt, die verschiedene Funktionen wie Audiotranskriptionen, Zitatenerkennung, Optimierung von Audioaufnahmen, und Generierung von Inhalten in andere Medienformate sowie Übersetzung in mehr als 40 Sprachen ermöglicht (Die Presse 2023; ORF 2024). International gibt es viele verschiedene Aktivitäten. Große Newsagenturen und Medienhäuser wie etwa Thomson Reuters, BBC, New York Times³³, Washington Post und weitere erforschen die verschiedenen Einsatzmöglichkeiten von KI im Medienbetrieb und entwickeln auch selbst Anwendungen, um deren Praxistauglichkeit zu testen. Etwa für visuelles Storytelling, automatisierte Übersetzungen oder KI-gestützte Aufbereitung multimedialer Inhalte in verschiedene Formate (z. B. text-to-audio, audio-to-video), Highlighting von Passagen in Live Streams, aber auch Anwendungen zur Erkennung von „deepfakes“ und manipulativen KI-Inhalten u. v. m. (Springer 2023; BBC 2024; NYT 2024; Frank 2024; Reuters 2024; Newman 2024). Neben KI-Anwendungen zur Unterstützung von Medieninhalten besteht insbesondere bei letzterem großer Bedarf an Werkzeugen im Umgang mit manipulativen KI-Inhalten, auch um KI-Einsatz nachvollziehbarer zu machen. Das ist nicht nur für Medienunternehmen, sondern auch für Rezipierende bzw. die Öffentlichkeit relevant (Newman 2021; Krawarik et al. 2021; Anlen/Llorente 2024). In Österreich werden etwa im Forschungsprojekt defalsif-AI³⁴ medienforensische Werkzeuge auf KI-Basis entwickelt, die Anwendern eine erste Einschätzung der Glaubwürdigkeit von Text-, Bild-, Video- oder Audiomaterial im Internet erlauben.

Neben journalistischen Anwendungsfeldern gibt es auch in der Werbe- und Marketing-Industrie schon seit längerem diverse Anwendungen u. a. für die Produktion und Verbreitung von Inhalten. KI bietet gerade hier vielfältige Möglichkeiten, kostengünstig unterschiedliche multimediale Inhalte (Texte, Bilder, Sprache, Videos etc.) aufzubereiten. Große Textilunternehmen nutzen etwa zusehends KI-generierte Bilder statt Fotos von echten Menschen als Models. Ein wachsender Bereich im Marketing ist zudem das sogenannte Influencer Marketing. Dabei werden KI-generierte, fiktive Personen („AI influencer“), insbesondere in sozialen Medien für unterschiedliche Praktiken zur Werbung und Beeinflussung eingesetzt (Looi/Kahlor 2024). Diese Entwicklungen werfen weitere Fragen zum Umgang mit KI-generierten Medienformaten und den Grenzen zwischen authentischen, fiktionalen und manipulativen Inhalten auf. Die Zuverlässigkeit von Inhalten ist ein wesentliches Erfordernis für den konstruktiven KI-Einsatz in allen Bereichen. Das gilt vor allem auch im Bildungssektor.

Umgang mit Deepfakes und manipulativen KI-Inhalten als wichtige Herausforderung

KI in Werbung und Influencer-Marketing

³¹ apa.at/produkt/ki-tools/.

³² orf.at/stories/3360846/.

³³ www.nytc.com/press/introducing-the-a-i-initiatives-team/.

³⁴ science.apa.at/project/defalsifai/.

3.1.3 SCHULE UND BILDUNGSWESEN

Auch im Bildungssektor bestehen große Erwartungen an KI und noch viele offene Fragen im Umgang damit. Anders als in den oben genannten Bereichen, wo sehr fachspezifische Anforderungen gelten, ist im Bildungsbereich der KI-Einsatz derzeit sehr unspezifisch mit variierenden Anwendungsformen und ungenauen Anwendungskontexten. Schulen sind einerseits damit konfrontiert, dass Schüler:innen niedrigschwellige generative KI Tools etwa zur Bearbeitung von Hausaufgaben und vielem mehr längst nutzen. Gleiches gilt an Universitäten, wo Lehrende davon ausgehen, dass die Mehrheit der Studierenden Seminar- und Abschlussarbeiten mithilfe von KI verfasst. Im Rahmen einer bundesweiten Studie in Deutschland wurde das KI-Nutzungsverhalten von über 6.300 Studierenden erfasst. Die überwiegende Mehrheit (63,2 %) haben KI-basierte Tools bereits im Rahmen des Studiums genutzt. Die Nutzungszwecke sind dabei vor allem: Klärung von Verständnisfragen und Erläuterung fachspezifischer Kontexte (35,6 %), Literaturrecherchen (28,6 %), Übersetzungen (26,6 %), sowie zur Analyse, Verarbeitung und Erstellung von Texten (24,8 %). Die Anwendungsformen sind in manchen Fächern unterschiedlich: In technischen Studien wie Informatik, Mathematik und Naturwissenschaften wird KI vor allem zur Programmierungsunterstützung genutzt. In sozialwissenschaftlichen Fächern sind es eher Aufgaben zur Text-Analyse und -Bearbeitung. Im künstlerischen Bereich zudem als kreatives Werkzeug etwa zur Konzeptentwicklung (Garrel et al. 2023). Aus diesem Nutzungsverhalten lässt sich einiges davon ableiten, wie KI das Lernen unterstützen kann bzw. das auch bereits stattfindet. Allerdings derzeit oft im Verborgenen und mit sehr unterschiedlichen Zugängen und Kompetenzen. Denn es fehlt vielerorts an klaren Strategien im Umgang damit. Einer Untersuchung der UNESCO zufolge hatten 2023 über 90 % der befragten Bildungseinrichtungen (weltweit über 450 Schulen und Universitäten) keinerlei Richtlinien oder Leitfäden für einen bewussten und informierten Umgang mit KI-basierten Technologien (UNESCO 2024). In manchen Ländern, u. a. in Deutschland und Österreich gibt es mittlerweile an vielen Universitäten und anderen Bildungseinrichtungen zumindest diverse Leitfäden (z. B. TU Wien³⁵, Universität Wien³⁶, Innsbruck³⁷, Uni Paderborn,³⁸ Saarland³⁹ u. v. m.). An Schulen gerichtet hat u. a. das österreichische Bildungsministerium einen öffentlichen Leitfaden zur Verwendung KI-basierter Tools veröffentlicht (BMBWF 2023a). Zudem erprobt das Bildungsministerium im Rahmen einer KI-Initiative⁴⁰ seit kurzem in 100 österreichischen Pilotschulen den praktischen Einsatz von KI-Anwendungen, um herauszufinden, wie sich KI an Schulen sinnvoll einsetzen lässt.⁴¹ Abseits der bereits stattfindenden Nutzung diverser Tools von Lernenden gibt es vielfältige Möglichkeiten, KI-Technologie

Viele offene Fragen im Bildungssektor

Mangel an klaren Strategien im Umgang mit KI in der Bildung

³⁵ www.tuwien.at/tu-wien/ueber-die-tuw/qualitaetsmanagement-an-der-tu-wien/news/news/leitlinien-fuer-den-verantwortungsoollen-einsatz-generativer-ki-in-der-forschung.

³⁶ phaidra.univie.ac.at/detail/o:2092606.

³⁷ www.uibk.ac.at/de/betriebswirtschaft/studies/umgang-mit-ki/.

³⁸ www.uni-paderborn.de/fileadmin/lehre/Digitale_Lehre_2023/Leitlinie_Einsatz_texterstellender_KI-Werkzeuge_2024-02.pdf.

³⁹ www.uni-saarland.de/fileadmin/upload/dezernat/ls/Handreichung_ChatGPT.pdf.
⁴⁰ education.at/community/ki-initiative-des-bm.

⁴¹ vorarlberg.orf.at/stories/3255102/.

im Bildungsbereich einzusetzen, u. a.: zur Erstellung und Bereitstellung von multimedialen Lehrmaterialien (etwa über KI-basierte Lernplattformen, Edutheken u.dgl.), als interaktiver Lernassistent zur individuellen Unterstützung bei verschiedenen Aufgabestellungen und Problemlösungen, Erstellung und Anpassung von Texten und Aufgaben je nach Lern- und Förderbedarf, für Fremdsprachentraining⁴², zur Fehlerkorrektur oder zur Förderung von Kreativität durch Generieren von Ideen, zur Sprachförderung und Abbau von sozialen Ungleichheiten und interkultureller Zusammenarbeit z. B. durch Übersetzungen in unterschiedliche Sprachen etc. Auch zur Fortbildung von Lehrpersonal gibt es viel Potenzial, etwa über Webinare (BMBWF 2023b; Shah 2023).

KI bietet also auch im Bildungsbereich vielfältige Anwendungsmöglichkeiten, die aber auch mehr Klarheit erfordern, um daraus realen Nutzen zu ziehen. Das setzt auch voraus, KI als Werkzeug zu begreifen, dessen Einsatz nicht überall sinnvoll ist. Etwa können KI-Tools zwar Recherchearbeit, analytisches oder wissenschaftliches Arbeiten für manches unterstützen aber nicht ersetzen. Der Umgang mit KI-Werkzeugen wird künftig ein noch stärkerer Teil von Lehrinhalten und dazu sind auch alternative Bildungskonzepte erforderlich (Garrel et al. 2023; Shah 2023). Ein zentraler Aspekt für den Bildungsbereich ist daher nicht primär die Anwendung von KI-Technologie, sondern die Vermittlung von Wissen für einen informierten, kritischen Umgang damit. Das erfordert auch eine interdisziplinäre Auseinandersetzung mit den Erkenntnissen aus Pädagogik, Didaktik und Entwicklungspsychologie, um Lernprozesse weiterhin – analog und digital – im Sinne des Bildungsauftrags der Schule erfolgreich zu gestalten. Etwa können die Schwächen von KI (siehe Abschnitt 2.2) auch so genutzt werden, um KI-generierte Inhalte gezielt kritisch im Unterricht auf Fehler hin zu analysieren. So könnte auch kritische Medienkompetenz, die immer mehr an Bedeutung gewinnt, mit neuen Ansätzen vermittelt und gestärkt werden.

Eine wesentliche Herausforderung für das Bildungswesen liegt aber in der sich verändernden Landschaft des Arbeitsmarktes und damit in der Vorbereitung von Schüler:innen, Studierenden und Pädagog:innen auf diesen Wandel. Hierbei geht es weniger um technische Expertise, sondern um den wachsenden Bedarf zur Vermittlung „klassischer“ Kernkompetenzen wie logisches und analytisches Denken, Kommunikationsfähigkeiten, Teamarbeit, Kreativität und eigenständiges Verstehen und Lösen von Problemen etc. (Interviews c, k, l). Gerade diese Fähigkeiten sind auch im Umgang mit KI-basierten Technologien von zentraler Bedeutung und werden als Grundvoraussetzung für eine sinnvolle Nutzung sogar umso wichtiger.

3.1.4 BEREICHSÜBERGREIFENDER KI-EINSATZ IM BERUFSALLTAG

Neben zahlreichen Anwendungen in unterschiedlichen Arbeitsbereichen zeichnet sich ab, dass KI in den nächsten Jahren auch schrittweise Einzug in bereichsübergreifende Arbeitsformen und „klassische“ Büroarbeit hält. Zum einen wird – ähnlich wie in Schulen – über diverse Spielarten generativer KIs deren Nutzen meist beiläufig experimentell für unterschiedliche Tätigkeiten genutzt. Auch hier kommt es ohne klaren Fachbezug i. d. R. eher zu unspezifischen Anwendungsformen und lose oder unklarem Anwendungskontext. Etwa, um kreativen Sti-

*KI als
Unterstützungs-
Werkzeug nutzen –
nicht als Ersatz für die
Wissensvermittlung*

*Vermittlung
von Wissen und
„klassischer“
Kernkompetenz
gewinnt an Bedeutung*

*KI hält langsam
Einzug in Büroarbeit*

⁴² www.heise.de/tests/Drei-KI-Sprechtrainer-im-Test-Fremdsprache-mit-KI-lernen-9676010.html.

mulus zu erhalten, Textbausteine, Bilder oder andere Medienformate zu generieren oder umzugestalten, Dokumente zusammenzufassen, oder sonstige Inhalte KI-unterstützt zu erstellen. Auch hier findet KI-Nutzung also meist nebenbei statt ohne klare Bezugspunkte zu konkreten Tätigkeiten und mit unterschiedlichen Kompetenzniveaus. Einer (nicht repräsentativen) Schweizer Studie zufolge nutzen rund 60 % der befragten Angestellten generative KIs im Berufsalltag. 47 % nutzen KI zur Textbearbeitung, 26 % zur Bildbearbeitung und 24 % für Programmiertätigkeiten (Deloitte 2023). Viele Unternehmen haben auch hier meist keine klare Strategie, ob und wie KI-basierte Technologien genutzt werden sollen und für welche Zwecke.

Es gibt viele Möglichkeiten, die Produktion von Inhalten und nicht-spezialisierten Tätigkeiten im Arbeitsalltag zu unterstützen. Seit längerem können diverse Tools automatisiert Transkripte erstellen, etwa zur Dokumentation von Meetings oder Texte und E-Mails zusammenfassen (allerdings meist mit vielen Fehlern und fragwürdiger Qualität). Schon kleine Werkzeuge wie automatisierte Sprachübersetzungsdienste (z. B. DeepL) oder direkt in Anwendungen wie Webbrowser integrierte KI-Features können den Arbeitsalltag erleichtern. Als vielversprechend gelten vor allem unterstützende Tools bei der Inhaltsbearbeitung. Zum Beispiel bei der Erzeugung, Bearbeitung Anpassung von Bildern, wie Formatänderungen und vielem mehr. Das ist insbesondere für Tätigkeiten nützlich, die mit der Gestaltung von Inhalten und Medienformaten arbeiten (z. B. Grafik, Design, Layoutierung etc.). Weit verbreitete Standard-Software zur Bearbeitung und Verwaltung von Dokumenten und Bildern setzt bereits verstärkt auf integrierte KI-Anwendungen (z. B. über die Integration von Adobe-Firefly in Photoshop und PDF-Programme). Zudem bieten einige Anwendungen auch ganz neue Möglichkeiten zur Unterstützung künstlerisch-kreativer Arbeitsschritte, z. B. bei der Bildbearbeitung oder für Videoinhalte, um aus Textprompts und Bildern Videosequenzen zu generieren (Lohmann 2023; Nickel 2024b). Ein Vorteil grafischer Tätigkeiten ist die unmittelbare visuelle Überprüfbarkeit KI-generierter Ergebnisse. Im Gegensatz zu Texten, wo die Überprüfung schon aufgrund des Korrekturlesens deutlich aufwändiger ist.

Trotz vieler Möglichkeiten sind gerade im Office-Bereich viele Anwendungen oftmals noch nicht ausgereift, haben zahlreiche Schwächen und produzieren Fehler, wie aus praktischen Testberichten hervor geht (Maier/Hoffmann 2024; Nickel 2024a; IOZ 2024; Janssen 2024). Manche Kommentatoren sehen in den KI-Ambitionen des marktdominierenden Unternehmens Microsoft sogar nur eine Art Neu-Auflage des gescheiterten digitalen Office-Assistenten „Karl Klammer“ (Cunnigham 2023), der Anfang 2000 aufgrund starker Kritik und großer Unzufriedenheit bei Anwender:innen endgültig aus Office-Anwendungen entfernt wurde.⁴³ Ob auch gegenwärtig neuen KI-Anwendungen ein ähnliches Schicksal droht, ist derzeit noch nicht absehbar.

Gerade generative KI-Anwendungen sind eher kleine Bürohelfer mit Schwächen, die kleinere Tätigkeiten mit sehr variierender Qualität unterstützen können. Zuverlässigkeit ist bei diesen Anwendungen nicht zu erwarten. Im Gegenteil ist viel manuelle Korrekturarbeit nötig, um überhaupt zu brauchbaren Ergebnissen zu kommen (vgl. u. a. Große 2024; Hohenwalde et al. 2024). Schwierigkeiten

*Viele
Unterstützungstools
zur Erleichterung
kleinerer Tätigkeiten*

*Deutliche Schwächen
und Fehler im
Office-Bereich*

*KI als kleine
Bürohelfer mit
Schwächen und
geringer Qualität*

⁴³ Im Englischen bekannt als „Clipper“ [de.wikipedia.org/wiki/Karl_Klammer_\(Microsoft\)](https://de.wikipedia.org/wiki/Karl_Klammer_(Microsoft)).

können vor allem entstehen, wenn es keinen bewussten Umgang mit diesen Schwächen gibt und unklar bleibt, wenn KI für unternehmensrelevante Entscheidungen eingesetzt wurde. Diese Intransparenz der Nutzung kann insbesondere entstehen und sich verstärken, wenn die Einführung von KI in Unternehmen ohne konkrete Planung und Ziele stattfindet. Das kann gerade bei weit verbreiteten IT-Systemen der Fall sein. Denn oftmals werden KI-Anwendungen über diesen Weg quasi beiläufig in den Arbeitsalltag integriert. Das geschieht meist nicht unbedingt proaktiv, sondern über Änderungen in Standardsoftware wie etwa sehr weit verbreiteten Betriebssystemen und damit verbundenen Softwareprodukten, insbesondere Office-Anwendungen. In den nächsten Jahren ist damit zu rechnen, dass KI noch stärker als bislang zum fixen Bestandteil von Standard-Software wird. Der Microsoft-Konzern will etwa seine KI-Anwendung MS-Copilot künftig stärker in seine Produkte integrieren (Cunningham 2023; Nickel 2024a; Standard 2024a). Der Aspekt der Integration von KI über Standardsoftware wird in Abschnitt 5.4 genauer diskutiert.

4 BESONDERHEITEN KI-BASIERTER AUTOMATION

Wie aus der Übersicht hervor geht, gibt es bereits vielfältige Einsatzgebiete für KI-basierte Technologien. Das Anwendungsspektrum wird stetig breiter und betrifft alle möglichen Branchen und Tätigkeitsfelder. Trotz vieler, bereichsspezifischer Unterschiede gibt es eine wesentliche Gemeinsamkeit: Der wachsende Einsatz von KI-Technologien bringt neue Formen der Automatisierung von Tätigkeiten, die bislang Menschen vorbehalten waren. In diesem Abschnitt wird erläutert, warum diese neuartigen Automatisierungsformen andere Implikationen haben als herkömmliche.

Im klassischen Verständnis meint Automatisierung die Gestaltung von Arbeitsprozessen derart, dass bestimmte Arbeitsabläufe durch Einsatz von Maschinen bzw. Technologien durchgeführt werden. Menschen organisieren und beaufsichtigen diese Prozesse, müssen aber die einzelnen Arbeitsschritte nicht selbst durchführen. Auch der Einsatz von KI impliziert die Automation bestimmter Arbeitsprozesse. Allerdings sind davon nicht mehr nur vorwiegend mechanische bzw. manuelle Tätigkeiten, sondern geistige Tätigkeiten und Wissensarbeit im weiteren Sinne (siehe Abschnitt 3), betroffen. Zudem gibt es einige wesentliche Besonderheiten KI-basierter Automatisierung.

KI-Automation bedeutet nicht zwingend einen Ersatz menschlicher Arbeit, sondern eine Ergänzung, wo bestimmte Tätigkeiten eines Arbeitsprozesses automatisiert werden. Manche Autoren beschreiben daher auch Automation und Augmentation (Erweiterung) als zwei Hauptanwendungsfelder von KI (vgl. Raisch/Krakowski 2021; von Richthofen et al. 2023). Augmentation meint hier die Kombination menschlicher und maschineller Arbeit in einem Tätigkeitsbereich. Allerdings wird auch hier eine Tätigkeit zumindest teil-automatisiert. Daher ist Automatisierung der primäre Effekt von KI-Einsatz und Augmentation als sekundärer Effekt bezieht sich auf den Automatisierungsgrad. Einfache Beispiele für KI-Automatisierung sind automatisierte Bild-, Text- und Sprachverarbeitung wie sie seit einigen Jahren mit Chatbots und Sprachassistenten im Kundenservice u. dgl. vermehrt zum Einsatz kommen. Auch hier ersetzt die Technologie nicht vollständig eine Aufgabe, aber sie verändert sie durch die Automatisierung einer bestimmten Tätigkeit (hier z. B. Annahme und Weiterleitung von Kundenanfragen).

In der Literatur gibt es verschiedene Konzepte zur Kategorisierung unterschiedlicher Automatisierungsgrade, insbesondere in der Fertigung (Frohm et al. 2008; Stowasser 2023). Eine gängige Unterscheidung im Bereich der autonomen Fahrzeuge umfasst beispielsweise fünf bis sechs Stufen, die häufig auch auf andere Bereiche KI-basierter Automatisierung übertragen werden. Zum Beispiel (Edwards et al. 2020) für Verarbeitung natürlicher Sprache über große Sprachmodelle. Solche Konzepte sind nützlich, um KI-Automatisierung allgemein zu systematisieren. Sie beschreiben aber nur die grundsätzlich angestrebte Möglichkeit, nicht die Realisierbarkeit im Konkreten. Die wesentliche Frage zu welchem Zweck, in welcher Form automatisiert wird und welchen Nutzen das bringt, bleibt hier außen vor. Dabei wird auch die Bedeutung menschlicher Interventionsmöglichkeiten und Überprüfbarkeit der mit KI automatisierten Prozesse ver-

*KI-basierte
Automatisierung hat
weitreichende Folgen*

*Unterschiedliche
Automatisierungsgrade*

nachlässigt. Gerade diese Aspekte sind aber zentral und hängen nicht nur vom grundsätzlichen Automatisierungsgrad, sondern vor allem von der Art und Funktionsweise eines KI-Systems ab. Wie in Abschnitt 2.3 erläutert, gibt es unterschiedliche Machine Learning-Verfahren, deren Funktionsweisen auch die Automatisierbarkeit von Tätigkeiten beeinflussen. Jedes Verfahren hat Besonderheiten hinsichtlich Datenmodellen und Art der Aufbereitung und Analyse von Daten. Unterschiedliche Ansätze können daher unterschiedliche Auswirkungen auf den Automatisierungsgrad und damit auch auf die damit verbundenen Risiken haben. Daher ist es auch für nicht-technische Nutzer:innen wichtig, die Implikationen der Automatisierung zu verstehen, einschließlich des allgemeinen ML-Ansatzes des Systems, mit dem sie interagieren. Gerade aufgrund zunehmender KI-Automatisierung gewinnt die Frage nach menschlicher Kontrolle weiter an Bedeutung. Durch die wachsende Verbreitung von KI-Anwendungen gibt es mehr Bedarf, die grundsätzliche Funktionsweise und Auswirkung von KI-Automatisierung zu verstehen, auch für Menschen ohne technisches Fachwissen.

*Menschliche
Intervention ist zentral*

4.1 HÖHERE DYNAMIK UND VOLATILITÄT IM SYSTEMVERHALTEN

Ein zentraler Unterschied zwischen herkömmlichen Automatisierungstechnologien und KI-basierter Automatisierung liegt in der Prozesslogik: klassische Automatisierung ist regelbasiert. D. h. es gibt klar definierte Abläufe und Regeln in Form von klaren Bedingungen (z. B. wenn Zustand x dann Aktion y) für deren Bearbeitung. Ein sehr einfaches Beispiel für ein regelbasiertes System ist eine Ampelschaltung. Regelbasierte Systeme wie moderne Computersysteme sind zwar komplex, aber grundsätzlich sind ihre Funktionsweise und damit verbundene Automatisierungsprozesse noch überprüfbar, sofern die Regeln und Bearbeitungsschritte nachvollziehbar sind. Moderne KI-Systeme sind dagegen nicht immer nur regelbasiert, sondern können – je nach ML-Verfahren – auch teil- oder ganz automatisiert ihre Funktionsweisen adaptieren, also verändern. Das hat eine höhere Dynamik und Volatilität (Schwankung/Wechselhaftigkeit) in Funktionsweise und Systemverhalten zur Folge, die auch zu höherer Intransparenz und geringerer Kontrollierbarkeit sowie Korrigierbarkeit führt. Zudem steigen dadurch auch die Anforderungen an die Qualitätskontrolle in Arbeitsprozessen und es entstehen neue Fragen hinsichtlich Zuverlässigkeit, Aktualität, Wartbarkeit etc.

*KI-Systeme
funktionieren nicht
nur regelbasiert*

Das Problem der Eigendynamik und Volatilität KI-basierter Systeme lässt sich anhand des Beispiels eines Autopilotensystems veranschaulichen: Der Autopilot eines Flugzeugs ist ein hochkomplexes System. Es ist regelbasiert, d. h. die Funktionsweise ist zwar komplex, aber klar definiert und hat den Zweck, menschliche Piloten derart zu unterstützen, dass diese das Flugzeug nicht ständig manuell steuern müssen. Als regelbasiertes System bleibt die Funktionsweise und Zuverlässigkeit des Systems für Piloten und andere Expert:innen überprüfbar und ggfs. korrigierbar. Der Autopilot steuert das Flugzeug anhand vom Piloten vorgegebenen Werten zu Flughöhe, Geschwindigkeit und weiteren Parametern. Ein solches System benötigt zudem Daten und ein plausibles Modell des Flugzeugs und seiner Umgebung mit zuverlässigen Informationen etwa zu Triebwerk, Topografie, GPS-Daten zum Standort, Wetter usw. das sich u. a. aus verschiedenen Sensoren speist. Das ermöglicht zum einen dem Autopiloten, die Route zu halten.

*Hohe Eigendynamik
und Volatilität
erschweren
Kontrollierbarkeit
von KI-Systemen*

Gleichzeitig dienen diese Informationen aber nicht primär dem Autopiloten zur autonomen Steuerung, sondern dem Menschen, um den Autopiloten ggfs. zu korrigieren. D. h. menschliche Intervention ist hier ein zentraler Bestandteil des Systems. Autopilotensysteme sind also keine vollständig autonomen Systeme.⁴⁴ Ein Autopilotensystem, das mit einem Deep Learning Verfahren laufend neue Rohdaten über Sensoren aus der Umgebung vollautomatisiert verarbeiten würde, um die Flugbahn ohne menschliche Intervention nach unklaren Regeln verändern würde, wäre unvorhersehbar, unkontrollierbar im Verhalten und damit ein enormes Sicherheitsrisiko. Der menschliche Pilot könnte immer nur korrigierend eingreifen und wäre aus Mangel an Kontrollmöglichkeiten letztlich dem System unterlegen, was im schlimmsten Fall zur Eskalation führt, wenn automatisierte KI-Entscheidung und menschliche Intervention miteinander kollidieren (Strauß 2021a). Dieses Beispiel verdeutlicht das Risiko, dass KI-basierte Automatisierung menschliche Autonomie bzw. präziser – menschliche Handlungsfähigkeit beschränken kann. Im nächsten Abschnitt wird dieses Problem als Teilaspekt von Deep Automation Bias als Meta-Risiko genauer erläutert.

*Gefahr der „Kollision“
maschineller
Entscheidung und
menschlicher
Intervention*

4.2 (DEEP) AUTOMATION BIAS: EIN ZENTRALES RISIKO DES KI-EINSATZES

Automation Bias bedeutet wörtlich automatisierte Verzerrung oder Befangenheit und bezeichnet das grundlegende Risiko, dass von automatisierten Technologien generierte Ergebnisse unhinterfragt akzeptiert werden. Es führt zu Fehleinschätzungen über die Funktionsweise von automatisierten Systemen und die Aussagekraft automatisiert erzeugter Information. AB wird grundsätzlich begünstigt, wenn der Anwendungskontext einer Technologie nicht, oder unzureichend mit der Arbeitspraxis übereinstimmt (Goddard et al. 2012/2014; Lyell/Coiera 2016; Strauß 2021a). Auf Anwender:innen-Seite erhöhen falsche Vorstellungen von Funktionalität und Grenzen automatisierter Technologien dieses Risiko. Mit KI-Einsatz steigen AB-bedingte Risiken noch weiter, weil sie die Komplexität und Intransparenz automatisierter Prozesse potenziell noch weiter erhöhen. Wie oben erläutert, unterscheiden sich KI-Systeme in ihrer Funktionsweise von klassischen Automatisierungstechnologien. Letztere sind regelbasierte Systeme, die zwar komplex sind, aber in ihrem Verhalten grundsätzlich vorhersehbar und dadurch sind sie auch überprüf- und kontrollierbar. KI-Technologien sind dagegen zunehmend komplexer und undurchschaubarer.

In Bezug auf KI-basierte Technologien wird dieses Risiko aufgrund der oben genannten Besonderheiten von KI-Automatisierung hier daher als Deep⁴⁵ Automation Bias (DAB) – (quasi tiefgreifende automatisierter Verzerrung) – bezeichnet. Denn es ist davon auszugehen, dass zunehmender KI-Einsatz AB begünstigt und verstärken kann. Aufgrund ihrer höheren und zunehmenden Komplexität, Dynamik und Volatilität sind KI-Systeme auch intransparenter, und ihr Systemverhalten unberechenbarer. Dadurch sind sie potenziell schwerer kontrollierbar (Strauß

*Erhöhtes Risiko
durch intransparentes
Systemverhalten und
falschen Vorstellungen
zur Funktionsweise*

⁴⁴ magazin.passengersfriend.com/autopilot-im-flugzeug-was-macht-der-pilot-noch-selbst.

⁴⁵ In Anlehnung an Deep Learning, dem bislang komplexesten ML-Verfahren (siehe Abschnitt 2.3).

2021a/b). Wie sich ein KI-System technisch verhält, hängt zunächst maßgeblich von mehreren Faktoren ab, insbesondere von:

- Qualität der zugrundeliegenden Datenmodelle
- Leistungsfähigkeit des maschinellen Lernens
- Benutzerfreundlichkeit der Schnittstellen
- Verantwortlichkeitsstrukturen und
- Transparenz- und Kontrollmechanismen im System

Die Kombination dieser Elemente wirkt sich auf die Zuverlässigkeit, Vorhersehbarkeit und Überprüfbarkeit des KI-Systems aus. Auch auf Anwenderseite gibt es Wirkfaktoren, wie:

- Fachwissen und technische Skills
- Verfügbare Ressourcen (auch zeitlich) und Handlungsspielraum
- Praxis-Erfahrung und Arbeitsbelastung
- Interpretationsfähigkeit
- Problembewusstsein und Lösungskompetenz

Abbildung 3 veranschaulicht die verschiedenen Einflussfaktoren von DAB entlang von zwei miteinander zusammenhängenden Dimensionen: Oben das „klassische“ Risiko des übermäßigen Vertrauens in das Verhalten eines KI-Systems und der Fehlinterpretation. Unten das zusätzliche Risiko der begrenzten Kontrollierbarkeit und der fehlenden Handlungsfähigkeit. Beide Risikodimensionen sind eng miteinander verwoben und werden von verschiedenen, sich gegenseitig beeinflussenden Faktoren bestimmt. Der Schlüsselfaktor, der beide Dimensionen verbindet, ist die Unsicherheit. Diese wird sowohl von technischen als auch von sozialen und organisatorischen Faktoren beeinflusst, die sich wechselseitig verstärken können.

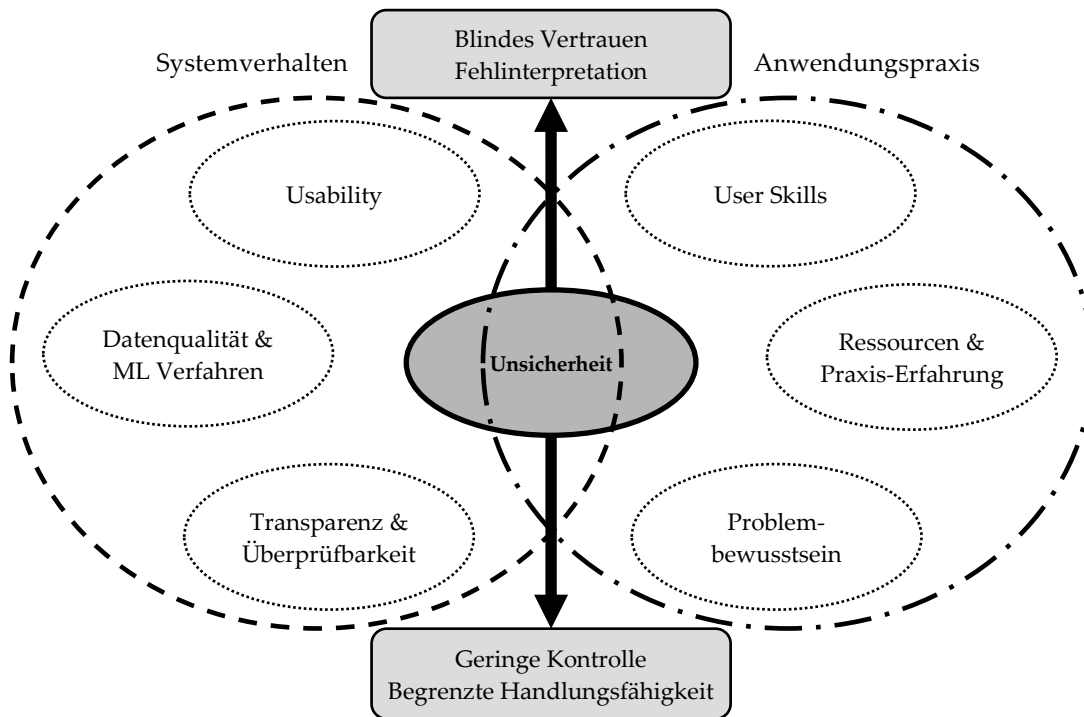


Abbildung 3: Einflussfaktoren von Deep Automation Bias (Quelle: adaptiert von Strauß 2021a)

Unterschiedliche Studien zu Automation Bias zeigen (u. a. Goddard et al. 2014; Lyell/Coiera 2016), dass die Möglichkeiten zur Interpretation und Prüfung automatisierter Verfahren durch das Zusammenspiel dieser Faktoren erheblich beeinflusst werden. Wissen und Fähigkeiten, praktische Erfahrung, verfügbare Ressourcen, Handlungsdruck, Arbeitsbelastung und vor allem auch Zeit zur Bewältigung von Aufgaben bestimmen maßgeblich, wie viel Handlungsspielraum und Handlungsfähigkeit vorhanden ist, um KI-Systeme effektiv und kontrolliert zu nutzen. Zugleich erschwert intransparentes und unberechenbares Systemverhalten das harmonische Zusammenspiel dieser Faktoren. Das beschränkt die Möglichkeiten, die adäquate Funktionsweise eines KI-Systems zu überprüfen. Das verringert auch die Kontrollierbarkeit, die Handlungsfähigkeit und Interventionsmöglichkeiten, um in die automatisierte Entscheidungsfindung einzugreifen. Der Schweregrad des DAB-Risikos und damit verbundener Risiken hängt vom Zusammenspiel all dieser Faktoren ab. Bestehen etwa wenig Möglichkeiten, um das Systemverhalten zu überprüfen, zu hinterfragen und ggfs. mit Intervention steuernd oder korrigierend einzugreifen, ist das Risiko höher. Gleiches gilt für Faktoren auf Anwender:innenseite: geringes Problembewusstsein, zu wenig Wissen, Ressourcen oder Zeit schränken die Fähigkeit zur Überprüfung und Erkennung von Fehlern oder Korrekturbedarf ein. Damit sinkt auch die individuelle Handlungsfähigkeit (Strauß 2021a).

Das Zusammenspiel zwischen Systemverhalten und Anwendungspraxis in konkreten Arbeitskontexten ist daher entscheidend für den konstruktiven KI-Einsatz. Dazu muss einerseits das KI-System den Anforderungen gerecht werden, die für bestimmte Tätigkeiten gelten. Andererseits müssen Anwender:innen das System auch zum Kontext passend nutzen. D. h. seine Funktionalität und Eignung für diese Tätigkeiten auch richtig einschätzen und zuordnen können. Kommt es

Zusammenspiel der Faktoren beeinflusst Handlungsfähigkeit

Ausgewogenes Verhältnis zwischen Systemverhalten und Arbeitspraxis entscheidend

hier zu Missverhältnissen zwischen Systemverhalten und Anwendungspraxis, kann das zu Problemen führen. Problematisch ist etwa, wenn die Leistungsfähigkeit des KI-Systems falsch eingeschätzt oder wenn Unklarheit darüber herrscht, welche Funktionalität und welche diesbezüglichen Grenzen das System hat. Diese Problematik kann durch verschiedene Umstände entstehen. Ein grundlegendes Problem ist zum einen die Neigung von Menschen, Technologie tendenziell eher „blind“ zu vertrauen und Ergebnisse nicht zu hinterfragen. Im Zusammenhang mit sprachbasierten Systemen wie Chatbots gibt es hier auch Bezüge zum sogenannten Eliza-Effekt, der die Neigung beschreibt, KI-Systemen menschliche Eigenschaften zuzuschreiben, sie also zu „vermenschlichen“ (vgl. Hofstadter 1996; Strauß 2018). Das macht Fehleinschätzungen über die Funktionalität noch wahrscheinlicher. Zum anderen kann die Komplexität der Technologien dazu führen, dass Anwender:innen nicht mehr die Fähigkeit haben, technologisch-gestützte Prozesse und Ergebnisse zu verstehen und kritisch zu hinterfragen. Das kann neben mangelndem Wissen und Erfahrung auch schlicht ein Problem mangelnder Ressourcen und Zeitmangel sein. Werden keine Maßnahmen zur Verringerung der Risiken getroffen, kann es in betrieblichen Arbeitsabläufen zu einer Art Teufelskreis kommen: Blindes Vertrauen in KI-Systeme, kombiniert mit intransparentem, unkontrollierbarem Systemverhalten. Kommen dann noch hoher Reaktionsdruck in Arbeitsabläufen und Ressourcenknappheit dazu – im Sinne eingeschränkter Möglichkeiten zum Verstehen und Interpretieren KI-generierter Ergebnisse und ggfs. zur Erkennung und Korrektur von Fehlern – dann wird auch das Risiko von unentdeckten Systemfehlern und letztlich von Fehlentscheidungen erhöht.

Auf längere Sicht kann das auch dazu führen, dass etablierte Arbeitspraktiken trotz Schwächen und Problemen an die Vorgaben von KI-Systemen angepasst und ihnen gewissermaßen unterworfen werden. Das kann dann die Entstehung paralleler Arbeitspraktiken und Umgehungsstrategien befördern, weil etablierte betriebliche Abläufe nicht gefährdet werden sollen und zugleich die technisch-organisatorischen Vorgaben des KI-Systems erfüllt sein müssen. Das birgt daher erhebliches innerbetriebliches Konfliktpotenzial, das zu Überlastung, Demotivation und sinkender Mitarbeiterzufriedenheit führen kann, weil zusätzlich zur Arbeitsbewältigung auch noch Parallelstrukturen aufrechterhalten werden müssen. So kann im schlimmsten Fall aus dem angestrebten Ziel der Entlastung von Arbeitsabläufen durch KI-Einsatz eine erhebliche Mehrbelastung werden, die sich negativ auf die Belegschaft und den Gesamtbetrieb auswirken kann (Strauß/Udrea 2024). DAB-Risiken können mit KI-Einsatz weiter zunehmen. Die Interviewpartner:innen bestätigten in jeweils unterschiedlichen Bereichen, dass das Problem von Fehleinschätzungen von KI-generierten Ergebnissen besteht (Interviews a-l). Dieser Problematik kann auf unterschiedliche Weise begegnet werden. Ein wichtiger Faktor zur Risikominderung ist das Problembewusstsein hinsichtlich des realen Nutzens und der Grenzen von KI-Anwendungen in bestimmten Arbeitskontexten als Bestandteil einer Critical AI Literacy. Diesbezügliche Ansätze werden in Abschnitt 6 genauer behandelt. Im folgenden Abschnitt werden zunächst zentrale Herausforderungen und Problemfelder aufgezeigt, die für den Umgang mit KI-basierten Technologien relevant sind.

Missverhältnis erhöht Risiko von Fehlentscheidungen und birgt mittelfristig Konfliktpotenzial

5 ZENTRALE HERAUSFORDERUNGEN UND PROBLEMFELDER

Trotz zahlreicher Unterschiede in den Anwendungsfeldern verschiedener Branchen gibt es grundlegende Gemeinsamkeiten bei Problemfeldern und Herausforderungen im Umgang mit KI-basierten Technologien, auf die im Folgenden genauer eingegangen wird.

5.1 WELCHEN MEHRWERT KANN KI BRINGEN?

Die Frage nach dem Mehrwert ist zentral, ist aber bei genauerer Betrachtung der Funktionsweise KI-basierter Technologien nicht einfach zu beantworten. Viele der grundlegenden Probleme beim Einsatz KI-basierter Technologien sind auf falsche Grundannahmen und Erwartungen zurückzuführen (siehe Abschnitt 1). Denn Effizienz- und Produktivitätssteigerungen werden zwar erwartet, sind jedoch nur bedingt realisierbar. Die weit verbreitete Prämisse, dass KI grundsätzlich valide Ergebnisse produziert und Effizienz steigert, ist schlicht nicht haltbar.

Dennoch ist diese Prämisse gängig, auch weil KI nicht nur ein technisches, sondern vor allem ein ökonomisches Phänomen ist. Das zeigt schon der enorme PR-Aufwand um generative KIs, die deshalb zwar hohe Bekanntheit genießen, aber deren praktischer Nutzen im Vergleich zu den zahlreichen Problemen im Detail noch sehr vage ist. OpenAI, das Unternehmen hinter ChatGPT, ist besonders aktiv nicht nur in der technischen Entwicklung, sondern vor allem im strategischen Marketing, um ihr KI-System medial breit zu bewerben (Haoyuan 2024). Unternehmen, die mit der Entscheidung konfrontiert sind, KI einzuführen, sollten sich daher im Vorfeld konkret mit der Frage nach dem realen Mehrwert im eigenen Arbeitsbereich auseinandersetzen. Denn der erzielbare Mehrwert hängt sehr stark von Anwendungsart, betroffenen Arbeitsabläufen und den Umgang mit steigender Komplexität und Kompliziertheit aufgrund des mit KI-Einsatz entstehenden Mehraufwands ab. Dass KI nicht nur Effizienzgewinne verspricht, sondern ganz konkret mehr Arbeitsaufwand mit sich bringt, wird kaum thematisiert. Die Frage nach dem Mehrwert ist demnach auch eine Frage nach dem Mehraufwand durch KI. Erst dadurch kann ein realistisches Bild darüber gewonnen werden, ob und inwiefern sich der Einsatz lohnt oder nicht.

Unbestritten bieten KI-Systeme viele Vorteile (siehe Abschnitte 2.3.2 und 3), von der Mustererkennung und Analyse großer Datenmengen bis zur Automatisierung diverser Tätigkeiten in unterschiedlichsten Bereichen. Allerdings primär bei gleichbleibenden, repetitiven Aufgaben, nicht unbedingt bei komplexeren. KI kann die Bewältigung komplexer Aufgaben insofern unterstützen, wie sie es erleichtert, kleinere Arbeitsschritte automatisiert zu kombinieren. Diese Kombination ist aber selten automatisierbar und mit viel Arbeitsaufwand verbunden. Allgemein ausgedrückt versprechen KI-basierte Technologien also primär einen Mehrwert bei Tätigkeiten, die repetitiv sind, deren Abläufe eindeutig sind und die einen hohen Formalisierungsgrad aufweisen – also formal und logisch abbildbar.

Viele Probleme auf falsche Erwartungen zurückzuführen

Frage nach Mehrwert mit KI auch eine Frage des Mehraufwands durch KI

Naheliegender erscheinen hier Aufgaben in der Softwareentwicklung oder anderen IT-Berufen, in praktisch jeder Branche in irgendeiner Form benötigt werden. Aus einigen Untersuchungen geht auch hervor, dass diverse KI-Tools, auch generative wie ChatGPT zu Arbeitserleichterungen in der Programmierung führen, um z. B. Software-Code zu generieren (Barenkamp 2020; Ozkaya 2023). Betrachtet man aber auch den Mehraufwand, werden negative Effekte auf die Qualität der Softwareentwicklung durch den Einsatz von KI-Assistenten sichtbar (Hosbach 2024; Harding/Kloster 2024). KI-Tools können zwar bei der Erzeugung einzelner Softwareelemente wie kleinerer Programmteile helfen und Tätigkeiten der Softwareentwicklung dadurch beschleunigen. Aber weder ist dadurch das automatisierte Generieren vollständig nutzbarer Anwendungen möglich, noch das Erzeugen fehlerfreier Software. Im Gegenteil steigt hierbei der Mehraufwand zur Überprüfung von Ergebnissen, Fehlerkorrektur und Qualitätskontrolle. D. h., die Qualitätsanforderungen steigen selbst in diesem Bereich an, wo grundsätzlich deutlich mehr Nutzen durch KI-Einsatz erzielbar ist (Strauß/Udrea 2024). Wird dieser Aspekt außer Acht gelassen, kann Software zwar schneller entwickelt werden, aber bei sinkender Qualität. Das zeigt sich „im Kleinen“ schon bei den diversen Unzulänglichkeiten generativer KIs, die nur selten tatsächlich valide Ergebnisse generieren. Im größeren Kontext ist KI-Einsatz mit vielen zusätzlichen, nicht-automatisierbaren Aufgaben verbunden. Insbesondere zur Datenpflege, Überprüfung, Systemwartung und Qualitätskontrolle. Um einen Mehrwert zu erzielen, sind insbesondere Fachwissen und Problembewusstsein Grundvoraussetzungen (siehe Abschnitt 4.2). Die Annahme, KI steigere automatisch die Effizienz von Tätigkeiten ist daher ein Trugschluss. Ohne Beachtung des Mehraufwands kann KI-Einsatz sogar zu Qualitätsverschlechterungen führen. Die Berücksichtigung dieser Aspekte ist zentral für Entscheidungen darüber, ob, was (also welche Tätigkeiten) und wie konkret mit KI automatisiert unterstützt werden sollen oder nicht. Zudem müssen im Vorfeld der Einführung die Ziele und Anforderungen klar sein, die der KI-Einsatz erfüllen soll. Das gilt für jede Art von Technologieeinführung in Betrieben, wird in der Praxis aber mitunter verabsäumt (Strauß/Udrea 2024). Das begünstigt Probleme bei der Integration der Technologie in Arbeitsabläufe und in weiterer Folge Probleme bei der Nutzung.

Das lässt sich beispielhaft anhand des Chatbots „Berufsinfomat“ vom österreichischen Arbeitsmarktservice (AMS) verdeutlichen (Proschofsky 2024; Mey 2024). Die Anwendung sollte Arbeitssuchende unterstützen, hatte aber erhebliche Mängel, die schon kurz nach der Veröffentlichung sichtbar wurden. Etwa das Reproduzieren geschlechterspezifischer Stereotype und Vorurteile: Berufsempfehlungen waren nach Geschlecht unterschiedlich aufbereitet. Zudem wurden auch einige technische Probleme der auf ChatGPT basierenden Anwendung deutlich. Im Berufsinfomat wurden somit auch die Schwächen dieses generativen KI-Produkts repliziert wie fehlerhafte oder irreführende Ergebnisse und Datenschutzprobleme. Hier gab es offenbar keine Schutzmaßnahmen in der Entwicklung. Etwa wurde auf der Webseite des „Infomat“ sogar vor der Eingabe personenbezogener Daten gewarnt. Viele Probleme wären bereits in der Anwendungsentwicklung vermeidbar gewesen wie etwa Maßnahmen gegen die Verstärkung von Stereotypen und Genderbias. Das AMS beteuerte, die Probleme beheben und die Anwendung verbessern zu wollen (Proschofsky 2024; Weiss 2024; Strauß/Udrea 2024). Das Beispiel zeigt, welche Probleme mit KI-Anwendungen entstehen können, wenn es an klaren Vorgaben an das Produkt mangelt. Es verdeutlicht, wie das Missverhältnis zwischen Systemverhalten und Nutzungspraxis zu praktischen Problemen führt:

*Arbeitserleichterungen
in der Programmierung
aber zulasten der
Qualität*

*Ohne Rücksicht auf
Mehraufwand kann KI
die Qualität sogar
verschlechtern*

Denn Probleme in der Nutzung liegen nahe, wenn die Funktionalität der Anwendung nicht mit den Anforderungen an den Einsatz in der Praxis konform geht.

Dieser Aspekt gilt bereichsübergreifend für jeden Anwendungsfall von KI. Auch in der Medizin, einem Feld mit erheblichem Potenzial, gibt es enorme Herausforderungen zu bewältigen, um praktischen Nutzen mit KI-Systemen zu realisieren. Zum einen bezüglich Transparenz und Nachvollziehbarkeit von KI-generierten Ergebnissen. Fehler und Fehlinterpretation sind hier besonders gravierend, daher erfordert der Einsatz von KI zur Unterstützung bei der Diagnostik und klinischen Entscheidungen umfassende Möglichkeiten, zur Überprüfung KI-generierter Resultate (Funer 2022). Wichtig ist zudem, inwieweit sich menschliche und maschinelle Fähigkeiten und Fehleranfälligkeiten unterscheiden und wie sich die Technologie einsetzen lässt, um einen echten Mehrwert zu schaffen. Es geht dabei nicht nur um die Produktion empirisch genauer und zuverlässiger Ergebnisse, sondern auch um deren Interpretierbarkeit. Das ist eine zentrale Voraussetzung dafür, dass KI-generierte Informationen und deren Genese im jeweiligen Anwendungskontext verstehbar ist, um zu entscheiden, ob diese akzeptiert, modifiziert oder als unzureichend abgelehnt werden (ebd.). Interpretierbarkeit ist zudem an Faktoren wie Funktionsweise und Zuverlässigkeit des Systems, Fachexpertise und kritisches Problembewusstsein von Anwender:innen geknüpft.

KI-Einsatz bringt nur dann einen Nutzen, wenn die Form der Entscheidungsunterstützung der KI-Anwendung in Einklang mit der konkreten Tätigkeit steht, wie u. a. eine Studie über den Einsatz von KI zur Hautkrebserkennung aufzeigt (Tschandl et al. 2020). Das gilt eher bei einfacheren Tätigkeiten, die klar definierbar sind aber nicht bei komplexeren, diagnostischen Problemen. Denn fehlerhafte KI-generierte Ergebnisse können letztlich auch medizinischen Expert:innen in die Irre führen. Eine zentrale Empfehlung der Studie ist daher, KI-generierten Ergebnisse nur unter Vorbehalt und im Zweifel gar nicht zu trauen und grundsätzlich immer der eigenen Einschätzung mehr (Tschandl et al. 2020). KI-Anwendungen haben zwar Potenzial, Menschen bei Aufgaben in der medizinischen Diagnostik zu unterstützen und zu entlasten. Aber die größte Herausforderung liegt nach wie vor darin, Umsetzungsfragen zu klären. D. h. auch hier: welche Tätigkeiten in welcher Form mittels KI-Einsatz automatisiert unterstützbar sind und sein sollen (Davenport/Kalakota 2019). Das ist nicht nur bei Anwendungen im Gesundheitswesen relevant, sondern gilt grundsätzlich für alle KI-Anwendungen.

Tatsächlich ist die simple und für Wissensarbeit zentrale Frage „what is the task?“ (Drucker 1999) aktueller denn je: D. h. was genau ist die Aufgabe, die ein KI-System bearbeiten oder unterstützen soll? Klare Antworten auf diese Frage sind zentral, um KI überhaupt sinnvoll einsetzen zu können. Laut Experteneinschätzungen gib es hier oft Defizite in der Praxis und viele offene Fragen, wie KI tatsächlich produktiv nutzbar ist in Unternehmen, unter welchen Voraussetzungen, und mit welchem Aufwand. Zwar gibt es in praktisch jeder Branche enorme Potenziale in verschiedenen Anwendungsfeldern, aber die Ausgestaltung variiert sehr stark. Das liegt auch daran, dass oftmals unklar ist, welche Erwartungen eigentlich bestehen, was KI-Technologie in konkreten Arbeitskontexten überhaupt leisten soll und ob das auch realisierbar ist (Interviews a, b, g, h, i, k). Die Klärung von Zielen, Anforderungen und des damit verbundenen Aufwands sind daher wichtige Grundbedingungen, um mit KI-Einsatz überhaupt einen Mehrwert erzielen zu können.

Funktionalität der KI-Anwendung muss Anforderungen in der Praxis gerecht werden

Nutzen bei der Entscheidungsunterstützung bei klar definierten Tätigkeiten

Entscheidend: Welche Aufgaben soll KI-Einsatz unterstützen?

5.2 KLUFT ZWISCHEN KI-NUTZUNG ALS EXPERTENSYSTEM UND „NEBENBEI“-TECHNOLOGIE

Wenngleich generalisierte Aussagen über verschiedene Branchen hinweg nur eingeschränkt sein können, zeigt sich doch ein klares Bild, dass Anwendungskontext und die Art der Nutzung von KI-Systemen maßgeblich beeinflusst, ob und welchen Mehrwert der Technologie-Einsatz in der Praxis tatsächlich bringt. Hier besteht ein Spannungsverhältnis zwischen der Anwendung von KI als bewusstes Werkzeug eingebettet in ein Expertensystem, und der scheinbar beiläufigen Anwendung für verschiedene, unkonkrete Tätigkeiten. Bei letzterem wird KI oftmals von Softwareanbietern in Standardsoftware wie z. B. weit verbreiteten Betriebssystemen integriert. Dadurch können KI-Anwendungen schrittweise und mitunter auch unbemerkt den Arbeitsalltag verändern (siehe auch Abschnitte 3.1.4 und 5.5). Die Auswirkungen auf die Arbeitsorganisation hängen daher wesentlich von der Art der Nutzung von KI-Anwendungen und ihrer Integration in konkrete Tätigkeiten ab. Hierbei spielt auch die Form der Wissensarbeit und das damit verbundene Tätigkeitsspektrum eine maßgebliche Rolle. Wie in Abschnitt 4.2 aufgezeigt, stehen den technischen Anforderungen an KI-Systeme allem voran: Zuverlässigkeit, Nachvollziehbarkeit und Interpretierbarkeit auch Anforderungen auf Anwender:innenseite gegenüber. Dazu zählen insbesondere auch Fachwissen, Erfahrung und Problembewusstsein. Tendenziell sind diese Faktoren stärker ausgeprägt, wenn KI-Anwendungen als Expertensystem für klar definierte Aufgaben und mit entsprechender Fachexpertise genutzt werden.

Die Kluft zwischen der KI-Anwendung als Expertensystem und der Nutzung als Art „En passant“-Technologie entsteht vor allem, wenn es keinen klar definierten Aufgabenbezug gibt. D. h. je breiter und unspezifischer die KI-Nutzung ausfällt, desto geringer auch der potenzielle Mehrwert. In Umgebungen mit spezialisierten Arbeitsbereichen mit entsprechend hoher Fachexpertise wie z. B. in der Medizin, der Forschung oder auch im Journalismus wird KI tendenziell eher als Expertensystem genutzt. In solchen Arbeitsumgebungen bestehen schon aufgrund der Art der zu unterstützenden Tätigkeiten oftmals höhere Anforderungen und mehr Notwendigkeit, Zweck und Ziele des Technologieeinsatzes im Vorfeld zu klären. Zudem ist bei Arbeitsbereichen mit hohem Verantwortungsgrad das Problembewusstsein für Fehler, Risiken und Folgen unzuverlässiger KI-Systeme meist höher als etwa in Bereichen, wo der KI-Einsatz nicht für spezifische Tätigkeiten für konkrete Zwecke mit bestimmten Anforderungen genutzt wird. Das kann mit dem Berufsfeld stark variieren und etwa bei klassischen Bürotätigkeiten eher der Fall sein. Oder auch im Bildungsbereich, wo derzeit noch sehr viel Unklarheit herrscht, wie KI sinnvoll einsetzbar ist (siehe Abschnitt 3.1.3). Der Grund liegt aber nicht unbedingt an mangelnder Fachexpertise von Mitarbeiter:innen, sondern an der Art und Weise, wie KI-Technologie eingesetzt wird und für welche Tätigkeiten. In solchen Arbeitsumgebungen ist die Einbeziehung von Arbeitnehmer:innen und Betriebsräten besonders wichtig. Werden KI-Technologien gänzlich ohne Berücksichtigung der Sichtweisen und Interessen der Belegschaft eingeführt, ist mit Problemen in der Praxis zu rechnen. Das verringert auch den praktischen Nutzen für Unternehmen.

Spannungsfeld zwischen KI-Einsatz als bewusstes Werkzeug und unkonkreter Nutzung

Klar definierte Aufgaben, Fachexpertise und Verantwortung

Welchen Unterschied die KI-Nutzung als Expertensystem macht, lässt sich am Beispiel Medizin verdeutlichen. Wie in Abschnitt 3.1.1 gezeigt, können KI-Systeme im medizinischen Bereich vielfältig eingesetzt werden, um Arbeitsabläufe zu verbessern und Diagnosen zu unterstützen, vor allem in bildgebenden Verfahren, aber auch in vielen anderen Anwendungsbereichen. Die rechtlichen und technischen Anforderungen⁴⁶ für den Einsatz von Technologien in der Medizin und die damit verbundenen Zertifizierungs- und Zulassungsprozesse sind hoch (vgl. VDE 2024). Präzision und Zuverlässigkeit sind wesentliche Grunderfordernisse an die Technologien, nicht zuletzt deshalb, weil sonst auch Menschenleben gefährdet sein können, etwa durch Fehldiagnosen. Das Problembewusstsein für technische Unzulänglichkeiten ist in medizinischen Anwendungsfeldern daher ebenso relativ hoch. Dadurch werden KI-Systeme hier tendenziell eher vorsichtig und nicht leichtfertig eingesetzt, weil die Risiken von Fehldiagnosen eher bekannt sind. Der Grundsatz, dass die Entscheidung immer beim Menschen liegt und nie bei einer Maschine, ist einerseits rechtlich geboten (Pfannstiel 2022) und lässt sich hier auch aus dem Berufsethos⁴⁷ ableiten, weil Mediziner:innen dem Patientenwohl verpflichtet sind und diese Verantwortung mit Entscheidungen verbunden ist, die nicht ohne weiteres über Technologien ausgelagert werden kann.

*Medizin als
Beispiel für relativ
stark ausgeprägtes
Problembewusstsein*

Ähnliches trifft zum Teil auch auf den Bereich Journalismus zu. Hier gibt es rechtliche und ethische Grundsätze, aus denen sich im Ideal hohe Qualitätsanforderungen für journalistische Inhalte ergeben: Entsprechend der journalistischen Sorgfaltspflicht müssen journalistische Inhalte faktentreu, valide und überprüfbar sein. In Anbetracht der Probleme von KI-Systemen hinsichtlich Transparenz, Fakten, validen Ergebnissen, Fehleranfälligkeit etc. ergibt sich hier schon eine gewisse Notwendigkeit im Berufsfeld, KI nicht leichtfertig einzusetzen und Ergebnisse zu überprüfen. Das gilt auch für die Forschung, wo Grundsätze wie Integrität, Transparenz, Nachvollziehbarkeit und gute wissenschaftliche Praxis⁴⁸ grundsätzlich eher einen verantwortungsvollen KI-Einsatz nahelegen. Das bedeutet nicht, dass andere Branchen generell leichtfertiger mit KI-Technologien verfahren. Aber hier gibt es tendenziell andere Herausforderungen. Insbesondere, wenn es keine klaren Anwendungszwecke, Ziele und Qualitätsanforderungen gibt, die zum Problembewusstsein beitragen und zur Orientierung bei der KI-Nutzung dienen. D. h. neben der oft genannten Transparenz von KI-Systemen braucht es auch eine Transparenz der Nutzung. Das erfordert Klarheit hinsichtlich des Nutzungszwecks und der Ziele von KI-Anwendungen. Erfolgt die KI-Einführung nicht als bewusste Entscheidung aufgrund betrieblicher Ziele und zur Unterstützung konkreter Aufgaben, ist diese Nutzungs-Transparenz schon im Vorfeld erschwert. Das ist häufig der Fall, auch weil Unternehmen hohen Druck verspüren, auf aktuelle Marktentwicklungen zu reagieren (Interviews a-i). KI-Einsatz erfolgt daher nicht immer primär aufgrund rationaler Überlegungen,

*Es braucht auch
eine Transparenz der
KI-Nutzung*

⁴⁶ Hier gibt es z. B. die EU Medizinprodukteverordnung (MDR – medical device regulation).

⁴⁷ www.bpb.de/themen/umwelt/bioethik/174950/aerztliche-ethik/.

⁴⁸ BMBWF (2020), Praxisleitfaden für Integrität und Ethik in der Wissenschaft.

www.bmbwf.gv.at/dam/jcr:91cf68d5-511e-4413-81ed-d71896f16e7c/Praxisleitfaden.

sondern aufgrund von externen Vorgaben wie der Marktkonkurrenz. Die Hoffnung auf mehr Effizienz und Produktivitätsgewinne (siehe Abschnitt 1) wird damit indirekt verstärkt. Zudem findet KI-Einführung auch seitens großer Softwarehersteller mit hoher Marktposition statt, oftmals auch ohne direktes Zutun betroffener Unternehmen oder privater Anwender:innen. Zum Beispiel, wenn KI-Anwendungen über weit verbreitete Betriebssysteme oder Standardsoftware Einzug halten in Arbeitsprozesse. Dadurch entstehen auch neue technologische Abhängigkeiten, derer sich Unternehmen nicht immer bewusst sind (siehe auch Abschnitt 5.5). Diese Art der Technologie-Einführung erschwert auch zusätzlich die konstruktive Einbindung von Arbeitnehmer:innen und Interessensverbänden. Daher bestehen in Folge dann auch weniger Gestaltungsmöglichkeiten für die nutzbringende Integration von KI-Systemen in bestehende Arbeitsabläufe. Insgesamt wird es unter solchen Bedingungen erschwert, spezifische Anwendungszwecke und Ziele festzulegen, die ein konstruktiver KI-Einsatz erfordert. Damit verringert sich auch die Chance, KI-Technologie so in die betriebliche Praxis einzubetten, dass sie zu einem tatsächlichen Mehrwert führt. In diesem Zusammenhang ist laut Arbeitsexpert:innen (u. a. Interviews a, b, g, h) auch wichtig, Mitarbeiter:innen mit niedrigschwelligen Informationen zu sensibilisieren, und mit mehr Transparenz auch die Frage gemeinsam mit der Belegschaft zu behandeln, wie KI-Technologien so eingesetzt werden können, dass sie den Beschäftigten helfen und auch zu Arbeitserleichterungen führen.

*Nutzbringende
Integration in
Arbeitsabläufe
wesentlich*

5.3 AKZEPTANZ UND AKZEPTABILITÄT

Es gibt eine Reihe von Einflussfaktoren, die sich auf die Akzeptanz von KI auswirken. In der Literatur werden insbesondere Sicherheit, Nutzen, Kompatibilität, Privatsphäre und Vertrauen, aber auch die Organisationskultur und Jobsicherheit angeführt (Wilkens 2020; Dabbous et al. 2022; Hasija/Esper 2022; Mirbbaie et al. 2022). D. h., KI-Einsatz wird eher akzeptiert, wenn die Anwendungen sicher sind und nicht die Privatsphäre berühren, keine Arbeitsplätze gefährden und sie praktischen Nutzen bringen. Dabei ist besonders relevant, wie kompatibel die Anwendungen mit bestehenden Arbeitsabläufen sind und wie gut sie darin integrierbar sind.

Ein zentraler Fallstrick ist hierbei, dass gerade über diese wichtigen Faktoren oftmals Unklarheit herrscht. Es gibt eine Diskrepanz zwischen den hohen Erwartungen an KI und der Realisierbarkeit dieser Erwartungen in bestehende Arbeitsorganisation und Arbeitspraktiken. Diese Diskrepanz sichtbar zu machen ist ein wichtiger Aspekt, um zu mehr Klarheit über den real erzielbaren Nutzen zu gelangen. Zudem erzeugt KI-Einsatz auch Handlungsdruck, der sich auf unterschiedliche Weise (institutionell und individuell) manifestiert. Auf institutioneller Ebene besteht für Unternehmen Druck aufgrund wirtschaftlicher Bedingungen wie Wettbewerbsfähigkeit, Produktivität und Effizienz von Arbeitsprozessen. Entlang gängiger Narrative (siehe Abschnitt 1.1) versprechen KI-Technologien, die wirtschaftlichen Bedingungen für Unternehmen zu verbessern. Entscheidungsträger:innen sind daher mit entsprechenden Erwartungen konfrontiert, die auch entsprechend innerbetrieblichen Druck erzeugen, sich dazu zu verhalten. Dieser innerbetriebliche Druck überträgt sich dann auch auf individueller Ebene

*KI-Einsatz erzeugt
institutionellen und
individuellen
Handlungsdruck*

auf die Belegschaft: KI-Einsatz steht unter dem Primat der Leistungssteigerung. Dabei bleibt aber häufig weitgehend offen, wie diese Erwartung realisiert werden kann (Interviews a, b, d, g, h, i). Zudem erschweren der hohe Abstraktionsgrad von KI-Technologie und Unklarheiten über ihre Funktionsweise den fundierten Umgang damit auf individueller wie institutioneller Ebene. Das Ausschöpfen der Potenziale von KI erfordert aber klare Ziele und Strategien zur Einbettung der Technologie in bestehende Arbeitspraktiken in Einklang mit organisationalen und sozialen Bedingungen. Daher spielt auch die Vorgehensweise bei der Technologie-Einführung eine zentrale Rolle. Wie konfliktbeladen oder reibungslos der KI-Einsatz erfolgt, hängt nicht nur von der Akzeptanz der Technologie ab, sondern auch von der Akzeptabilität. Beide Faktoren sind eng miteinander verknüpft, haben aber unterschiedliche Bedeutungen. Akzeptanz bedeutet Einverständnis oder Nicht-Ablehnung, sagt als subjektiver Faktor für sich genommen jedoch nichts über die Annehmbarkeit einer Technologie in bestimmten Anwendungskontexten aus (Grunwald 2008; Taebi 2017). Akzeptabilität ist ein Maß für die Angemessenheit des Technologie-Einsatzes in Bezug auf den daraus resultierenden Nutzen und der Sozialverträglichkeit ihrer Anwendung. Eine Technologie kann zwar auf hohe Akzeptanz stoßen, weil sie etwa vielversprechende Potenziale hat oder benutzerfreundlich ist. Aber ihre Nutzung kann dennoch nicht konform mit betrieblichen Anforderungen gehen, keinen praktischen Nutzen bringen, ethische und soziale Probleme verschärfen, oder sogar gegen gesetzliche Bestimmungen verstoßen. Zur Vorbeugung dieser Probleme ist daher die nachvollziehbare und sozialverträgliche Gestaltung des Technologieeinsatzes bereits in der Planungsphase entscheidend. Das erleichtert die konstruktive Nutzung der Technologie und die Entwicklung wirksamer Problemlösungsstrategien. Ein weiterer wichtiger Faktor ist hierbei auch die Praktikabilität, also die Eignung des KI-Einsatzes in einem bestimmten Anwendungskontext.

Eine damit verbundene große Herausforderung betrifft auch den Faktor Komplexität. Zum einen bieten KI-basierte Technologien Möglichkeiten, die Komplexität von Arbeitsprozessen besser zu bewältigen. Zum Beispiel durch effiziente Unterstützung bei der Datenanalyse und Aufbereitung von Wissen. Zum anderen erhöht KI-Einsatz aber auch die Komplexität (siehe auch Abschnitt 2.3). Das betrifft vor allem auch den unterschätzten zusätzlichen Aufwand bei der Gestaltung und Nutzung von KI, etwa um die Ergebnisqualität zu sichern und Fehler zu vermeiden und ggfs. zu korrigieren. Diese Komplexitätserhöhung kann zu Hindernissen in der effizienten Gestaltung von Arbeitsabläufen führen. Etwa, wenn zusätzliche Arbeitsschritte erforderlich werden, um KI-Anwendungen überhaupt nutzen zu können oder hoher Aufwand zur Fehlerkorrektur von KI-generierten Ergebnissen (siehe Abschnitt 4.2, 5.1). Werden Arbeitsabläufe durch KI-Einsatz ohne ersichtliche Vorteile nur komplexer und komplizierter, wirkt sich das negativ auf die Akzeptanz der Technologie aus. Die Folge können dann Umgehungsstrategien und parallele Arbeitsstrukturen sein und längerfristig Überlastung und Demotivation von Mitarbeiter:innen.

Problematisch kann hierbei auch sein, wenn KI-basierte Technologien zu starren Arbeitsabläufen führen und dadurch Freiräume und Handlungsspielraum von Mitarbeiter:innen eingeschränkt werden (Interviews a-d, g-i). Denn der KI-Einsatz setzt ein gewisses Maß an Datafizierung und Formalisierung von Arbeitsabläufen voraus. Zum einen werden größere Datenmengen benötigt und generiert. Das kann Ängste und Sorgen von Mitarbeiter:innen vor Überwachung und Kontrolle verstärken. Zum anderen besteht auch die Gefahr, dass es zu einer

*Vorgehen bei
Technologie-Einführung
auch Frage der
Akzeptabilität*

*Zwischen
Komplexitätserhöhung
und -Bewältigung*

Überformalisierung von Abläufen kommt, die zwar kurzfristig mehr Effizienz versprechen, aber längerfristig zu mehr organisatorischem und prozeduralem Aufwand führen und die Mitarbeiterzufriedenheit senkt. In weiterer Folge hat das auch negative Folgen für die Produktivität von Unternehmen. Es ist daher auch wichtig abzuwägen, ob und inwieweit KI-Einsatz für bestimmte Arbeitsabläufe geeignet ist oder nicht. Andernfalls kann gefühlter oder realer Verlust von Kreativität und Gestaltungsfreiheit auch zu weniger Akzeptanz und Rückgang der Arbeitszufriedenheit führen. Das hat dann auch betriebliche Folgen wie Rückgang von Produktivität.

Es braucht deshalb auch Maßnahmen zur Bewältigung jener Komplexität, die aus dem KI-Einsatz resultiert. Vertrauen ist grundsätzlich ein wichtiger Aspekt im Umgang mit Komplexität und für Akzeptanz von Technologien. Vertrauen führt zur psychologischen Entlastung, weil es wichtige Bedürfnisse für Sicherheit und Kontrolle erfüllt (Chibanguza et al. 2022). Demnach würde das Vertrauen in die korrekte Funktionsweise von KI-Systemen die Akzeptanz in diese Systeme begünstigen. Die Krux ist hierbei aber, dass die Funktionsweise von KI oft intransparent, schwer nachvollziehbar und damit auch kaum kontrollierbar sein kann. Zudem ist gerade hier übermäßiges, „blindes“ Vertrauen in die Technologie ein Problemfaktor, der das Risiko von (Deep) Automation Bias und Fehler begünstigt (siehe Abschnitt 4). Besonders dann, wenn Vertrauen aus Unkenntnis oder mangelnden Möglichkeiten zur kritischen Reflexion der Funktionsweise einer Technologie resultiert. Einfache Annahmen wie etwa „Vertrauen stärkt die Akzeptanz“ sind bezogen auf Technologieeinsatz daher im Allgemeinen schon zu stark vereinfachend und bei KI-Einsatz im Speziellen zusätzlich problematisch. Ein wesentlicher Grund liegt an der grundlegend anderen Funktionsweise von KI-Systemen im Vergleich zu regelbasierten Systemen: KI-Systeme sind komplexer und verhalten sich tendenziell dynamischer, was sie auch schwerer überprüfbar und kontrollierbar macht. Wenn hier noch übermäßiges Vertrauen aufgrund von Unkenntnis hinzukommt, steigt die Komplexität im Umgang mit KI-Technologie noch weiter. Relevanter als Vertrauen ist daher der Faktor kritisches Bewusstsein. Anwender:Innen sollten möglichst in der Lage sein, beurteilen zu können, inwieweit ein KI-System zuverlässig funktioniert und ob die produzierten Ergebnisse verwendbar oder nicht verwendbar sind. In Bereichen, wo KI als Expertensystem genutzt wird bzw. in eine Expertensystemumgebung einbettet ist, mit entsprechenden Anforderungen an Prozessqualität und Fachexpertise von Nutzer:innen (siehe Abschnitt 5.1).

Eine wichtige Grundvoraussetzung für Vertrauen und Akzeptanz ist daher, die nachvollziehbare und sozialverträgliche Gestaltung des Technologieeinsatzes bereits in der Planungsphase. Denn Transparenz und Mitsprache bei der Technologieeinführung stärken auch die Akzeptanz und Akzeptabilität. Eine ergebnisoffene Auseinandersetzung mit KI bietet zudem auch Chancen für Arbeitgeber und Arbeitnehmer, gemeinsam die innerbetrieblichen Strukturen und Abläufe und Arbeitsorganisation zu reflektieren. Das kann auch den innerbetrieblichen Zusammenhalt stärken.

Probleme durch zu stark für KI-Einsatz formalisierte Arbeitsabläufe

Erfolgsfaktor transparente und sozialverträgliche Gestaltung

5.4 „VERDECKTE KI“: INTRANSPARENZ UND DATENANALYSEN IM HINTERGRUND

Innerbetrieblich entstehen durch Digitalisierung und KI-basierte Anwendungen weitere Herausforderungen. Transparenz ist ein bekanntes Erfordernis an KI-Systeme und Intransparenz ein vielgenanntes Problem. Neben dem Einsatz zur Unterstützung von Arbeitsprozessen können KI-Systeme auch anderen Zwecken dienen, die nicht ohne weiteres ersichtlich sind. Anwendungen werden immer vernetzter, dadurch werden auch Datenflüsse immer schwerer kontrollierbar. Immer häufiger kommt es zu Formen algorithmischer Erfassung und Auswertung großer Datenmengen im Hintergrund, etwa als Bestandteil von Standardsoftware wie großer Office-Anwendungen oder vertriebsunterstützender Systeme. Problematisch ist insbesondere, wenn personenbezogene Daten automatisiert im Verborgenen mitanalysiert und zur Mitarbeiterkontrolle genutzt werden. Teils gibt es Mikro-Anwendungen in größeren Systemen, die neben ihres Hauptverwendungszwecks scheinbar beiläufig Daten erfassen und auswerten (Interviews a, b, g, h, i). Bereits bestehende Datenschutz-Problematiken können sich so deutlich verschärfen. Expert:innen sehen hier Gefahren unzulässiger Datenerfassungen zur verdeckten Mitarbeiterüberwachung. In manchen Branchen ist das bereits ein evidentes Problem, wie etwa in der Transportlogistik und Lieferdiensten. Studien zeigen zudem wachsende Sorgen und Ängste von Mitarbeiter:innen vor zunehmender Überwachung am Arbeitsplatz infolge automatisierter Datensammlungen im Hintergrund und steigender Nutzung KI-basierter Technologien (Interviews a, b, g, h, i, k; Lane et al. 2023; IFES 2023; Deloitte 2024).

Der Einsatz dieser Technologien und der damit verbundene wachsende Bedarf an Datenerfassungen führen zu Verunsicherung in Unternehmen. Neben der Gefahr KI-basierter Mitarbeiterüberwachung können auch andere Probleme durch verdeckte Anwendungen entstehen. Etwa wenn KI-Systeme über integrierte Funktionen und zusätzliche Anwendungen zusätzliche Daten erfassen und analysieren, die über das eigentlich benötigte Maß für konkrete Nutzungszwecke hinaus gehen. Dieses Problem besteht weniger bei spezialisierten KI-Anwendungen, jedoch mehr bei Standardanwendungen und Betriebssystemsoftware. Standardsoftware kann hierbei in mehrfacher Hinsicht problematisch sein: wenn Hersteller KI in ihre Produkte so einführen, sodass betriebliche und private Anwender:innen diese Funktionen nicht mehr deaktivieren können. Das hat dann zur Folge, dass auf intransparente Weise unterschiedliche Daten von KI-Funktionen erfasst und weiterverarbeitet werden. Insbesondere KI-Anwendungen im Office-Bereich können sehr viele Daten erfassen, deren Zweck nicht immer nachvollziehbar ist. Das kann erhebliche Datenschutz- und Sicherheitsprobleme nach sich ziehen. Ganz besonders, wenn private/personenbezogene Daten oder sensible Unternehmensdaten betroffen sind. Hinzu kommt, dass aufgrund mangelnder Kontextualisierung und Zweckbindung von breiten KI-Anwendungen oftmals unklar ist, welche Daten tatsächlich für bestimmte Aufgaben nötig sind und welche darüber hinaus miterfasst werden und zu welchem Zweck. Hier zeigt sich langsam verschiebende Grenze zwischen privaten und öffentlichen bzw. zwischen zweckgebundenen und unweckmäßig erfassten Daten. Schon vor dem KI-Hype seit längerem umstritten sind Metriken und Datenanalysen über das Nutzerverhalten, wie etwa im Fall der Cloud-basierten Standardsoftware Microsoft Office 365.

Datenschutzprobleme durch verdeckte Datenauswertungen

Unterschätzter Problembereich: Beiläufige KI-Einführung über Standardsoftware

Der Fall ist beispielhaft für schleichenden Zwang zu proprietären Cloud-Anwendungen und verdeutlicht, welche Probleme damit einher gehen können. Im Fall Office 365 wurde in der Vergangenheit bereits mehrfach (u. a. in Deutschland, den Niederlanden und der EU) behördlich festgestellt, dass die Datensammelpraktiken der Software nicht rechtskonform sind und gegen die Datenschutzgrundverordnung verstoßen (Beiersmann 2018; Dinges 2020; Rosbach 2020; DSK 2022). Im März 2024 zeigte der EU-Datenschutzbeauftragte sogar bei der EU-Kommission Datenschutzverstöße aufgrund rechtswidriger Nutzung von Microsoft 365 auf.⁴⁹ Darüber hinaus gibt es erhebliche Datenschutzprobleme in Schulen, wo die Verarbeitung personenbezogener Daten besonders sensibel ist. Auch hier wurden Datenschutzverletzungen festgestellt und daher Verbote der Software an Schulen gefordert (Rosbach 2020; Krempel 2023; noyb 2024). Mit KI-Einsatz nimmt auch diese Problematik tendenziell noch weiter zu und Lehrerverbände fordern daher Rechtskonformität bei der KI-Nutzung in Schulen (Beer 2024).

Der Fall ist beispielhaft für ein größeres Problem: ein starkes Abhängigkeitsverhältnis zu Standardsoftwareprodukten. Unternehmen wie auch andere Organisationen haben Probleme, Standardsoftware zu vermeiden, selbst wenn ihre Funktionalität nicht rechtskonform ist. Außerdem fehlt es hier auch oftmals an Wissen und Bewusstsein über diese Problematik. Daher kann sich durch intransparenten Technologieeinsatz das Machtungleichgewicht am Arbeitsplatz verstärken. Zum einen, wie bereits ausgeführt, zwischen Arbeitgeber und Belegschaft, wenn KI-Systeme als Kontrollinstrument missbraucht werden. Zum anderen zwischen Technologiebetreibern und Unternehmen: Die meisten Unternehmen können KI-Systeme nicht selbst entwickeln, sondern sind auf externe Dienstleister und Technologiebetreiber angewiesen. Große Technologie-Unternehmen erhalten so noch mehr Marktmacht und wirtschaftlichen Einfluss. Bereits jetzt bestehen hier zahlreiche technologische Abhängigkeiten (Strauß/Bettin 2023), die mit steigender Verbreitung und Nutzung von KI-Systemen weiter zunehmen. Innerbetrieblich können dadurch technologische Lock-Ins entstehen bzw. sich weiter verfestigen – also die enge Bindung an bestimmte Herstellersysteme, die einen Wechsel zu alternativen Systemen erschwert. Das ist insbesondere auf Ebene von Betriebssystemen der Fall. Dieser Markt wird global dominiert von wenigen Anbietern: Etwa hat Microsoft mit seinem Betriebssystem Windows weltweit einen Marktanteil von über 70 %.⁵⁰ Diese monopolartige Marktstellung kann sich durch daran gekoppelte Softwareprodukte noch weiter verfestigen. Derzeit wird seitens des deutschen Kartellamts und der EU-Kommission geprüft, ob Microsoft mit seinen Produkten (wie Windows, Teams, Xbox oder auch Bing) gegen EU-Kartellvorschriften verstößt.⁵¹ Mit zusätzlichen KI-Anwendungen gibt es bereits weitere, stark umstrittene Entwicklungen. Zum Beispiel die KI-Funktion „Recall“, die wie ein fotografisches Gedächtnis für KI sämtliche Nutzeraktivitäten und Inhalte

*Mangelnde
Rechtskonformität*

*Steigende
technologischer
Abhängigkeiten zu
KI-Betreibern*

⁴⁹ www.heise.de/news/EU-Datenschutzbeauftragter-EU-Kommission-hat-Microsoft-365-rechtswidrig-genutzt-9651423.html.

⁵⁰ de.statista.com/statistik/daten/studie/157902/umfrage/marktanteil-der-genutzten-betriebssysteme-weltweit-seit-2009/.

⁵¹ www.golem.de/news/monopol-kartellamt-prueft-bedrohung-durch-microsoft-2303-173026.html
germany.representation.ec.europa.eu/news/eu-kartellvorschriften-kommission-ubermittelt-microsoft-mitteilung-der-beschwerdepunkte-wegen-2024-06-25_de.

speichern und auswerten will. Die Funktion ist als Teil des KI-Assistenten Copilot geplant, der schrittweise in alle Microsoft-Systeme Einzug halten soll. Sicherheitsforschende äußerten erhebliche Bedenken aufgrund gravierender Datenschutz- und Sicherheitsprobleme. Dennoch ist derzeit offen, ob die Funktion nicht trotzdem integraler und nicht mehr deaktivierbarer Systembestandteil wird (Standard 2024b; Joos 2024; Warren 2024). In Anbetracht der unterschiedlichen Probleme aufgrund proprietärer Standardsoftware zeichnet sich zudem Bedarf nach Open-Source Alternativen von KI-Systemen ab, die nicht an datengetriebene Geschäftsmodelle gekoppelt sind wie derzeitige marktdominierende Systeme.

Die Integration von KI-Systemen in zentrale Geschäftsprozesse ohne entsprechende Sicherheitsvorkehrungen kann auch Sicherheitsprobleme begünstigen und die Vulnerabilität von Unternehmen erhöhen. Insbesondere, wenn KI-Systeme zu einem Bestandteil der unternehmensinternen kritischen Infrastruktur werden. Dadurch werden diese Systeme auch lukrativer für Cyberangriffe. Es gibt bereits viele unterschiedliche Angriffsvarianten mit und auf KI⁵², etwa neue Formen von Phishing-Attacken, vereinfachte Schadcodeprogrammierung, automatisiertes Sammeln von Daten für gezielte Angriffe, neue Betrugsformen mit KI-generierten Inhalten wie Deep Fakes, oder auch neue Formen wie prompt injection⁵³ – einer neuartigen Angriffsform zur Manipulation von generativen KI-Systemen. Dabei kann u. a. Schadsoftware über herkömmlich wirkende Prompts eingeschleust werden, um dann an vertrauliche Daten zu gelangen oder das System zu beeinträchtigen.

Um diese Intransparenz und die damit verbundenen Probleme zu verringern, braucht es innerbetrieblich mehr Klarheit über die Funktionsweisen der eingesetzten Systeme und in welcher Form sie welche Mitarbeiterdaten verarbeiten, sowie Überprüfungsmöglichkeiten für Interessensvertretungen. Das ist für die Einhaltung arbeitsrechtlicher Bestimmungen wichtig und für die Stärkung von Mitarbeiterzufriedenheit. Entsprechende Maßnahmen zur Stärkung von Transparenz und Überprüfbarkeit haben auch positive Auswirkungen auf die Produktivität. Dass insgesamt grundlegender Bedarf nach mehr Wissen und konkreten Maßnahmen im Umgang mit KI besteht, zeigt auch der hohe Grad an Unwissenheit und Verunsicherung in der breiteren Öffentlichkeit. Wie etwa eine repräsentative Umfrage in Deutschland aufzeigt, gibt es trotz wachsender Bekanntheit mangelhafte Kenntnisse zu Algorithmen und KI. Die Studie betont daher wie wichtig Wissen- und Kompetenzaufbau sind, auch in der gesellschaftlichen Breite (Bertelsmann 2022).

Neben breiter Bewusstseinsbildung sind Maßnahmen in Unternehmen wesentlich: Innerbetrieblich spielen hier Instrumente zur Förderung von Transparenz und Überprüfbarkeit eine zentrale Rolle. Grundsätzlich bildet in Österreich das Arbeitsverfassungsgesetz⁵⁴ (ArbVG) hierbei eine wichtige Grundlage für Betriebsräte. Darin sind auch Einsichtsrechte festgelegt: Laut §91 ArbVG muss der Arbeitgeber von sich aus den Betriebsrat informieren, wenn „die wirtschaftlichen, sozialen, gesundheitlichen oder kulturellen Interessen“ der Arbeitnehmer:Innen berührt sind. (§91 Abs. 1 ArbVG). Auch über automationsgestützte Aufzeichnungen und Verarbeitungen personenbezogener Daten in einem System und deren

Bedarf nach Open Source Alternativen

Neue Sicherheitsprobleme und Vulnerabilität kritischer Infrastruktur

Bedarf nach Stärkung innerbetrieblicher Instrumente für Transparenz und Überprüfbarkeit

⁵² www.onlinesicherheit.gv.at/Services/News/KI-Cyberangriffe-Ueberblick.html;
www.onlinesicherheit.gv.at/Services/News/KI-Cyberattacken-Risiken-Trends.html.

⁵³ www.ibm.com/de-de/topics/prompt-injection.

⁵⁴ www.jusline.at/gesetz/arbvg.

Grundlagen sind Arbeitgeber:innen dem Betriebsrat gegenüber auskunftspflichtig und müssen Einsicht gewähren. Laut Arbeitsexpert:innen wird diese Bringschuld in der Praxis aber häufig übersehen und muss dann aktiv von Betriebsräten eingefordert werden (Interviews a, b, g, h). Instrumente wie diese sind sehr wichtig, um Transparenz über KI-Einsatz zu schaffen und die Belegschaft klar zu informieren, welche innerbetrieblichen Auswirkungen daraus resultieren. Hier verorten Expert:innen zusätzlichen Bedarf zur Stärkung solcher Instrumente über Regulierungen, die ergänzend zur KI-Verordnung auch die Arbeitsrechte im Umgang mit KI stärken.

5.5 NEUER QUALIFIKATIONSBEDARF ABER UNEINDEUTIGE TÄTIGKEITSBEREICHE

Mit steigendem KI-Einsatz nehmen auch Sorgen vor Abbau von Arbeitskräften zu. Diese Sorgen vor Arbeitsplatzverlust sind laut Einschätzung unterschiedlicher Expert:innen zwar nicht unberechtigt, aber zumindest momentan nicht als unmittelbare Folge von KI zu erwarten. Im Gegenteil erfordert konstruktiver Einsatz der Technologie tendenziell sogar teils zusätzliche Stellen und einige Tätigkeiten verlagern sich. Allerdings gibt es in manchen Branchen durchaus bereits Pläne zur Stellenreduktion. Z. B. beim deutschen Axel Springer Verlag, wo durch Technologieeinsatz auch Personal und Kosten im Boulevardsektor eingespart werden sollen (Wilkens 2023). Auch in der IT-Branche kursieren Pläne zur Personalreduktion innerhalb der nächsten Jahre, etwa bei Stellen ohne Kundenkontakt.⁵⁵ Gleichzeitig erhöht sich aber Bedarf nach neuen Stellen und neuen Kompetenzen auch in dieser Branche.⁵⁶

Aus einigen Studien geht hervor, dass KI-Einsatz die Arbeitswelt nicht derart verändert, dass menschliche Arbeitskräfte ersetzt werden. In manchen Bereichen zeichnen sich bereits eine Erweiterung des Tätigkeitsspektrums und die Entstehung neuer Jobs ab. Demnach gibt es derzeit zumindest keine eindeutigen empirischen Belege für Arbeitsplatzverluste, die unmittelbar auf KI-Einsatz rückführbar sind. Die Beschäftigungsniveaus scheinen trotz KI-Zunahme bislang weitgehend stabil zu sein. Allerdings ist ein etwas verlangsamtes Beschäftigungswachstum bereits sichtbar (Lane et al. 2023; OECD 2023). Inwieweit das eine direkte Folge von KI-Einsatz ist, lässt sich nicht eindeutig belegen. Vielmehr kann das eine zwischenzeitliche Entwicklung sein, weil die KI-Einführung mit zusätzlichen Fachkompetenzen und Spezialisierungen einher geht, die nicht unmittelbar verfügbar sind, sondern erst im Entstehen sind. Um die verschiedenen Möglichkeiten des Technologieeinsatzes auszuschöpfen, braucht es auch neue Qualifikationen und Kompetenzen (Albrecht/Kellermann 2020; Lane et al. 2023; OECD 2023; Stowasser 2023). Eine steigende Nachfrage an Berufen mit KI-Bezug ist bereits sichtbar. Beispielsweise gilt derzeit etwa Prompt Engineering⁵⁷ als neues Tätigkeitsfeld bei der Verwendung generativer KI-Anwendungen, um zu

*Abbau von
Arbeitskräften derzeit
keine direkte KI-Folge
aber Tätigkeiten
verlagern sich*

*Neue
Qualifikationsprofile
aber unkonkrete
Aufgabenbereiche*

⁵⁵ futurezone.at/digital-life/ibm-will-8000-jobs-durch-kuenstliche-intelligenz-ersetzen/402433035.

⁵⁶ www.computerwoche.de/article/3564046/ki-verandert-92-prozent-der-it-jobs.html.

⁵⁷ www.businessinsider.com/prompt-engineering-ai-chatgpt-jobs-explained-2023-3.

möglichst produktiven Ergebnissen zu kommen. Hier ist allerdings zweifelhaft, ob überlegtere Texteingaben schon ein eigenes Qualifikationsprofil erfordern. Ein Indikator für steigende Nachfrage an Berufen sind einschlägige Stellenausschreibungen. Auf Jobportalen sind aktuell⁵⁸ diverse Stellen mit KI-Bezug mit variierenden Schlagworten ausgeschrieben (z. B. AI/Machine Learning Engineer, AI Prompt Engineer, AI/ML Developer, AI Software Engineer, Data Analyst, AI Manager, AI Consultant, KI Enthusiast, Researcher for generative AI usw.). Ob hier AI als Aufhänger für bekannte Jobs in IT und Softwareentwicklung dient oder tatsächlich völlig neue Berufe entstehen, ist dabei nicht direkt ersichtlich. Die Ausschreibungen zeigen aber einen generellen Zuwachs an Fachstellen im IT-Sektor aufgrund steigender Bekanntheit von KI-basierten Technologien. Welche neuen Qualifikationen und Aufgabenbereiche im Konkreten entstehen, hängt zudem stark von den Anwendungsbereichen und ihren Anforderungen ab. Diese sind wiederum je nach Bereich häufig wage, weil oftmals noch gar nicht klar ist, für welche konkreten Tätigkeiten KI tatsächlich genutzt wird, und wie sich dadurch die Tätigkeitsbereiche und Qualifikationsprofile mittel- und längerfristig verändern.

Abseits des naheliegend steigenden Bedarfs an technischen Qualifikationen sind auch andere Kompetenzen erforderlich, die aus den neuen Möglichkeiten der Arbeitsgestaltung resultieren (siehe Abschnitt 6). Dabei entstehen auch viele Erleichterungen für Beschäftigte durch neuartige Arbeitsschritte. Etwa neue Möglichkeiten zur Datenauswertung und Qualitätssicherung durch KI-basierte Bilderkennungsverfahren, in der Sprachverarbeitung, der Anlagenverwaltung, oder bei der Instandhaltung und Wartung; z. B. durch die automatisierte Überwachung von Verschleißteilen zur frühzeitigen Wartung und vieles mehr (siehe auch Abschnitt 3). Mit dieser Verschiebung von Arbeitsschritten fallen daher nicht einfach Tätigkeiten weg, sondern es entstehen auch neue Beschäftigungsfelder in anderen Bereichen (Albrecht/Kellermann 2020; OECD 2023; Stowasser 2023). Es deutet daher mittelfristig einiges auf eine weitere Automatisierung und teilweise Reduktion von Routinetätigkeiten und eine Zunahme an Tätigkeiten wie etwa Qualitätsmanagement in KI-gestützten Prozessen hin.

Ein Teil von Qualitätssicherungsaufgaben umfasst aber auch in der Öffentlichkeit weniger bekannte und unterschätzte Tätigkeiten zur Datenbereinigung für KI-Systeme. Große Technologieunternehmen lagern diese Tätigkeiten häufig an Clickworker in Entwicklungsländer aus. Diese „Datenreinigung“ geschieht nicht automatisiert, sondern umfasst auch manuelle Arbeit, die meist von Billiglohnarbeiter:innen im globalen Süden durchgeführt werden.⁵⁹ Etwa verdienen Arbeiter:innen in Kenia 1,50 US-Dollar für die Datenbereinigung von ChatGPT⁶⁰. Ähnliches findet auch bei anderen Tech-Konzernen statt wie zB. Tesla, Facebook/Meta, Google, Microsoft etc. Die Arbeit wird meist nicht direkt über die Tech-Konzerne durchgeführt, sondern über Subaufträge an auf Datenbereinigung und Labeling spezialisierte Unternehmen wie z. B. Cloudfactory, Samasource, Appen oder Clickworker (Kretschmer 2023; Peer 2023; ifb 2024). Bei dieser Arbeit müssen Menschen große Mengen an fehlerhaften, teils auch illegalen, problematischen und verstörenden Inhalten (wie Hass, Gewaltdarstellungen, Pornografie usw.)

*Arbeiterleichterung
führt auch zu neuen
Beschäftigungsfeldern*

*Mittelfristig Zunahme
an Aufgaben zur
Qualitätssicherung KI-
gestützter Prozessen*

⁵⁸ U. a. bei Stepstone.at, Karriere.at, ai-jobs.at mit Stand 28.10.

⁵⁹ t3n.de/news/ki-clickworker-ausbeutung-1634433/ etc.

⁶⁰ www.derstandard.at/story/2000142768897/das-schmutzige-geheimnis-von-chat-gpt-sind-kenianische-billiglohnkraefte.

sichten, aussortieren, löschen oder kennzeichnen („labeling“)⁶¹. Diese weitgehend unsichtbaren, und oftmals prekären Arbeitsformen⁶² sind eine Begleiterscheinung der Gig Economy und Plattformökonomie, die mit KI noch weiter zunimmt.

Neben diesen ausgelagerten „Datenreinigungsarbeiten“ kann KI-Einsatz aber auch in anderen Bereichen zu einer Art Monotonisierung von Tätigkeiten der Wissensarbeit führen. Zum einen, weil Datenbereinigung und -pflege bei so gut wie allen KI-Anwendungen nötig ist. Diese Arbeit ist für Betriebe nicht ohne weiteres auslagerbar, muss jedenfalls mitberücksichtigt werden und kann dann innerbetrieblich zur unliebsamen, monotonen Tätigkeit werden. Zum anderen, wenn Tätigkeiten für den Zweck ihrer Durchführbarkeit mit KI so formalisiert werden, dass dies zu einer Einschränkung von kreativen und sinnstiftenden Handlungsspielräumen führt. Ein plakatives Beispiel wäre eine Grafikerin, die nur noch KI-generierte Inhalte erstellt und in ihrer Arbeit keine Möglichkeit mehr zur eigenständigen Kreativität hat. In der US-Filmbranche gab es bereits Gewerkschaftsproteste gegen den Einsatz von KI, um z. B. durch KI-generierte Doubles Kosten und Personal abzubauen. U. a. wurde erwirkt, dass biometrische Merkmale wie Gesicht, Körper und Stimme nicht ohne Zustimmung der betroffenen Schauspieler:innen für digitale Kopien verwendet werden dürfen (Wilhelm 2024).

Aber die Gefahr, dass KI längerfristig zu einer Monotonisierung führt, besteht nicht nur vordergründig bei kreativen Tätigkeiten. Sondern insbesondere dann, wenn KI nicht als Werkzeug zur Verbesserung von Arbeitsprozessen eingesetzt wird, sondern Arbeitsprozesse schlicht derart formalisiert werden, dass KI-Einsatz überhaupt möglich ist – KI-Einsatz also zum Selbstzweck wird. Andere, fiktive Beispiele wären etwa die Reduktion journalistischer Arbeit auf semi-automatisiertes Generieren von KI-Inhalten auf Basis von Social Media Content; oder Medizin:innen, deren Aufgaben auf das Übernehmen KI-generierter Diagnosen und Korrigieren von Fehlern reduziert würde. Gerade in Bereichen mit hoher Expertise besteht die Gefahr der Monotonisierung aber tendenziell weniger als in anderen Bereichen. Denn Tätigkeiten sind hier zumindest an einen bestimmten Qualitätsanspruch, Verantwortung und Sorgfaltspflichten geknüpft und KI wird in solchen Arbeitsbereichen eher als Teil eines Expertensystems eingesetzt (siehe auch Abschnitt 5.2).

Wo fachspezifischer Nutzen bereits absehbar ist, kann der Aufwand zur Neuorganisation von Arbeitsprozessen tendenziell geringer ausfallen als in Bereichen, wo KI-Anwendungen für Tätigkeiten genutzt werden, für die bislang keine technische Fachexpertise erforderlich war und die keiner klaren Ablaufstruktur folgen. Das heißt, der Einsatz von KI-Technologien erfordert zum einen ein höheres Kompetenzniveau, etwa in Bezug auf Fähigkeiten zur Analyse und Interpretation. Und zum anderen auch ein breiteres Spektrum an Kompetenzen etwa fachspezifisches Wissen zum Technologieeinsatz (OECD 2023). Dieser Kompetenzaufbau ist nicht nur wichtig, um KI-Systeme anzuwenden, sondern um Arbeitsabläufe so zu gestalten, dass mit KI tatsächlich auch mehr Raum für andere, nicht monotone Tätigkeiten entsteht und Handlungsfähigkeit gestärkt wird. Andernfalls droht mittel- u. längerfristig eine Monotonisierung von Arbeit, die letztlich den Faktor Arbeit nicht ent-, sondern belastet und auch negative Folgen für die Wirtschaftlichkeit von Betrieben haben kann. Insbesondere, wenn der

Gefahr der Monotonisierung von Arbeit und Einschränkung von Handlungsspielräumen

Mehr Bedarf nach Kompetenzen zur Analyse und Interpretation

KI kann Arbeit auch zusätzlich belasten

⁶¹ netzpolitik.org/2024/ki-arbeiter-in-kenia-die-arbeitsbedingungen-sind-erbaermlich/
www.nzz.ch/meinung/kuenstliche-intelligenz-geister-in-der-maschine-ld.1513894.

⁶² www.nzz.ch/meinung/kuenstliche-intelligenz-geister-in-der-maschine-ld.1513894.

Technologieeinsatz dazu führt, dass Wissensarbeiter:innen zu wenig Freiräume bei der Gestaltung ihrer Arbeit haben und KI-Systemen „zuarbeiten“. Im Idealfall führt KI-Einsatz zur Entlastung bei Routineaufgaben, wodurch mehr Zeit für qualitative Aufgaben zur Verfügung steht. Dazu braucht es aber unter Einbindung von Betriebsräten auch klare Regeln und ein klares Bekenntnis dazu, dass die Zeitersparnis konstruktiv zur Verbesserung der Arbeitsqualität, und nicht nur zu weiteren Kostensenkungen und Personalabbau genutzt wird (Interviews a-d, g-i; von Richthofen et al. 2023).

Es braucht daher neue, alternative Ansätze, wie KI-basierte Technologien konstruktiv und sinnstiftend in Arbeitsabläufe integriert werden können, um einen Mehrwert für Betriebe und Beschäftigte zu schaffen. Damit geht auch eine teilweise Umorganisation von Tätigkeiten und Arbeitsprozessen einher, zu der auch eine Klärung des konkreten Bedarfs an zusätzlichen Fähigkeiten und neuen Qualifikationen zählt. Es entstehen also neue Tätigkeiten und Rollen, die auch neue Fähigkeiten erfordern. Wie in Abschnitt 5.1 ausgeführt, ist eine wesentliche Grundvoraussetzung für den sinnvollen KI-Einsatz zunächst die Klärung des Verwendungszwecks. Also z. B.: Welche Tätigkeit oder die Lösung welches Problems sollen mit KI unterstützt werden?

Das impliziert Analysen des Ist-Zustands und des konkreten Bedarfs nach KI-Unterstützung (von Richthofen et al. 2023) in den innerbetrieblichen Arbeitsprozessen. Daraus leitet sich auch der Bedarf an konkreter Fachexpertise und Zusatzqualifikationen in den jeweiligen Arbeitsbereichen ab. Neben vielen spezifischen, von Branche zu Branche unterschiedlichen Anforderungen an Fachkompetenzen ist ein zentraler, übergreifender Faktor aber die Stärkung von Handlungsfähigkeit im Umgang mit KI-basierten Technologien. Dabei spielt der Aufbau einer kritischen KI-Kompetenz eine zentrale Rolle, die im nächsten Abschnitt genauer erläutert wird.

*Klares Bekenntnis zur
Arbeitsverbesserung
erforderlich*

*Ist-Zustandsanalysen
zur Klärung des
konkreten KI-Bedarfs*

6 HANDLUNGSFÄHIGKEIT MIT CRITICAL AI LITERACY STÄRKEN

Wie gezeigt wurde, gibt es eine Reihe von Herausforderungen und Problemfeldern im Umgang mit KI-basierten Technologien zu bewältigen. Im Zentrum steht dabei die menschliche Handlungsfähigkeit auf unterschiedlichen Ebenen. Hier wird bewusst nicht nur der Begriff Autonomie verwendet, der als ethisches Konzept der Selbstbestimmung zwar relevant, aber wenig praxistauglich ist. Fragen wie z. B.: „Schränkt KI meine Autonomie ein?“ sind in praktischen Anwendungsfällen kaum beantwortbar, weil sie auch subjektive Wahrnehmungen von Freiheit implizieren. Handlungsfähigkeit ist dagegen zutreffender, weil ihre Beeinträchtigung unmittelbarer sichtbar ist. Denn trotz vielfältiger Anwendungsfelder und beeindruckender neuer Möglichkeiten der Arbeitsunterstützung mit KI-Systemen ist ein nachhaltiger Mehrwert letztlich nur dann erzielbar, wenn der Einsatz dieser Technologien zu einer Stärkung der menschlichen Handlungsfähigkeit führt und nicht zu einer Einschränkung.

Aufgrund der Besonderheiten KI-basierter Automatisierung und dem daraus resultierenden DAB-Risiko (siehe Abschnitt 4) entstehen auch neue Fragen zur Gestaltung des Zusammenspiels zwischen individueller Handlungsfähigkeit und Arbeitsorganisation. KI-Einsatz verändert dieses Zusammenspiel, jedoch ist noch weitgehend offen, in welcher Form. Ein vielgenanntes Schlagwort für menschliche Kontrolle von KI ist Human-in-the-Loop (HITL). HITL-Konzepte gelten als zentral, um sicherzustellen, dass der Mensch immer die Kontrolle über KI-Systeme behält (Emmanouilidis et al. 2019). Wie praxistauglich diese Konzepte sind, hängt aber wiederum von ihrer Implementierung in konkreten Anwendungsbereichen, Arbeitsbedingungen und der Kompetenz der eingebundenen menschlichen Akteure ab. HITL kann auch als technisches Konzept verstanden werden, um Menschen als Bestandteil von Kontrollmechanismen in Arbeitsabläufe zu integrieren, die aber für KI-Systeme vorgesehen sind. In dieser Lesart von HITL wäre der Mensch demnach eher ein unterstützender „Zuarbeiter“ von KI. Im Sinne des Ziels, menschliche Handlungsfähigkeit zu stärken, ist das eine Herangehensweise die hinterfragbar ist.

Betriebe können nur von KI profitieren, wenn auch ein praktischer Nutzen gegeben ist. Der praktische Nutzen von KI-Systemen hängt maßgeblich von der Zuverlässigkeit und Überprüfbarkeit ihrer Funktionsweise, ihrer kontextuellen Einbettung in konkrete Arbeitsabläufe zur Unterstützung menschlicher Handlungen und der kritischen Fachexpertise der Anwender:innen ab. D. h. eine Grundbedingung für praktischen Nutzen ist es, KI-Systeme als Werkzeuge zur Unterstützung menschlicher Handlungsfähigkeit einzusetzen und Ansätze zu etablieren, damit Anwender:innen den Nutzen auch selbständig beurteilen können. Das setzt wiederum Wissen, Problembewusstsein und Fachkompetenzen voraus. Genau daran mangelt es aber bislang oftmals (Interviews a-j). Es gibt demnach einen grundlegenden Mangel an Wissen und kritischem Bewusstsein über den praktischen Nutzen und die Grenzen des Einsatzes von KI, der sich in vielen Anwendungsbereichen widerspiegelt. Zur Bewältigung dieses Mangels braucht es daher mehr kritische Technikkompetenz im Umgang mit KI-Systemen. Mit Critical AI Literacy (CAIL) ist hier eine kritische KI-Kompetenz angesprochen,

Neue Fragen der Balance zwischen Handlungsfähigkeit & Arbeitsorganisation

Der Mensch ist kein „Zuarbeiter“ von KI

Grundbedingung für praktischen Nutzen: KI-Systeme zur Stärkung von Handlungsfähigkeit

die die Fähigkeit beinhaltet, die Grundfunktionalität eines KI-basierten Systems und seiner Eignung oder Nicht-Eignung in bestimmten Anwendungskontexten zu verstehen (Strauß 2021a). Die Entwicklung technischer Fähigkeiten zur KI-Nutzung steht dabei nicht im Vordergrund. Technische Fähigkeiten wie etwa zur Entwicklung oder Anpassung von KI-Systemen nötig, bleiben auch weiterhin Bestandteil technischer Disziplinen wie Informatik oder anderen Berufen mit IT-Bezug. Vielmehr geht es darum, dass auch Anwender:innen und Entscheidungsträger:innen ohne spezifische technische Fachexpertise befähigt werden, KI-Systeme zu verstehen, zu nutzen und deren Ergebnisse einzuschätzen und hinterfragen zu können. Ein wesentlicher Bestandteil von CAIL ist daher der Aufbau von Bewusstsein, Wissen und Erfahrung im Umgang mit KI-basierten Technologien und den damit verbundenen Möglichkeiten, Grenzen und Herausforderungen. Das sind wichtige Voraussetzungen für den konstruktiven und sozialverträglichen KI-Einsatz für alle betrieblichen Akteure (d. h. Anwender:innen, Entwickler:innen, Entscheidungsträger:innen und Interessensvertretungen) und letztlich auch der von KI-Einsatz Betroffenen.

Grundsätzlich ist die Erweiterung von Kompetenzen im Umgang mit neuen Technologien und Medien immer zentral. Es gibt daher eine Reihe unterschiedlicher, aber miteinander verbundener Literacy-Konzepte (wie technological literacy, computer literacy, media literacy, ICT/digital literacy etc.). Zudem sind grundlegende Fähigkeiten wie Lese-, Schreib- und Sprachkompetenzen sowie logisches und kritisches Denken zentrale Fundamente jedes Ansatzes (vgl. Yi 2021). Das Besondere an KI-bezogener Literacy liegt an den neuen Anforderungen und Herausforderungen des Einsatzes KI-basierter Technologien. KI-Systeme bringen neue Automatisierungsformen, verhalten sich dynamischer, verändern sich schneller und sind daher auch schwerer versteh- und kontrollierbar als andere digitale Technologien (siehe Abschnitt 4). Daraus ergeben sich teils ganz andere Nutzungsanforderungen.

In der Literatur wird AI Literacy meist als Kompetenz beschrieben, die Individuen befähigt, KI-Technologien zu verstehen, um sie effektiv und verantwortungsvoll zu nutzen (Long/Magerok 2020; Ng et al. 2021). Die Betonung auf Critical AI Literacy (Strauß 2021a) ist im Rahmen dieser Untersuchung bewusst gewählt, um das Wissen im Umgang mit KI nicht auf die Nutzung der Technologie zu beschränken. Die gängigen Definitionen suggerieren einen notwendigen KI-Einsatz, der kaum hinterfragbar ist. Dadurch wird der Bewertungsrahmen eingeschränkt, um eine informierte Entscheidung darüber zu treffen, ob und wie die Technologie mit Mehrwert nutzbar ist. Es braucht daher nicht nur techniszentrierte, sondern auch grundlegende kritische Kompetenzen, um zu verstehen, in welchem Ausmaß KI-Einsatz überhaupt einen Mehrwert bringt. Denn wie auch aus den Herausforderungen hervor geht, ist trotz vieler Vorteile die Nutzung nicht in jedem Fall für jede Tätigkeit zielführend. CAIL ist eine wichtige Voraussetzung für die konstruktive Nutzbarkeit von KI. Sie umfasst insbesondere:

- *Basiswissen und Problembewusstsein* über den praktischen Nutzen und die Grenzen von KI-basierten Technologien
- *Interpretationsfähigkeit*: Das Verhalten von KI-Systemen, KI-generierte Inhalte und Ergebnisse zu erkennen und verstehen
 - d. h. ihre *Aussagekraft* und ihren situativen Kontext *einzuschätzen* (also die konkreten Umstände in einem Anwendungsfall)
 - *Überprüfen der Korrektheit* und *Plausibilität* von Ergebnissen

Auch Menschen ohne technische Expertise müssen KI-Systeme verstehen und hinterfragen können

Wesentliche Komponenten von Critical AI Literacy

- **Bewältigungsfähigkeit:** Problemlösungskompetenz im Umgang mit unklarem Systemverhalten, fehlerhaften, irreführenden oder fragwürdigen Ergebnissen
 - *Beurteilen des praktischen Nutzens und der Akzeptabilität* zum Beispiel: Ist KI-Ergebnis im konkreten Anwendungsfall brauchbar und akzeptabel oder nicht?
 - *Verwenden oder zu verwerfen* von Ergebnissen, Korrektur und ggfs. Veranlassung weiterer Überprüfung des KI-Systems als informierte, bewusste Entscheidung

D. h. im Idealfall werden Anwender:innen darin ermächtigt, zu beurteilen, inwieweit ein KI-System zuverlässig funktioniert, korrekte, plausible oder ausreichend genaue Ergebnisse produziert und unter welchen Voraussetzungen die Ergebnisse verwendbar oder nicht mehr verwendbar sind. CAIL lässt sich in Anbetracht der langfristigen Auswirkungen von KI auch als Baustein zur Futures Literacy⁶³ einordnen, d. h. die Fähigkeit, Veränderungen zu verstehen und lösungsorientiert damit umzugehen.

Ermächtigung von Anwender:innen

⁶³ www.unesco.org/en/futures-literacy.

6.1 CAIL-FRAMEWORK

Um CAIL innerbetrieblich zu stärken, kann der folgende Bewertungsrahmen (Strauß 2021a) zur Unterstützung herangezogen werden:

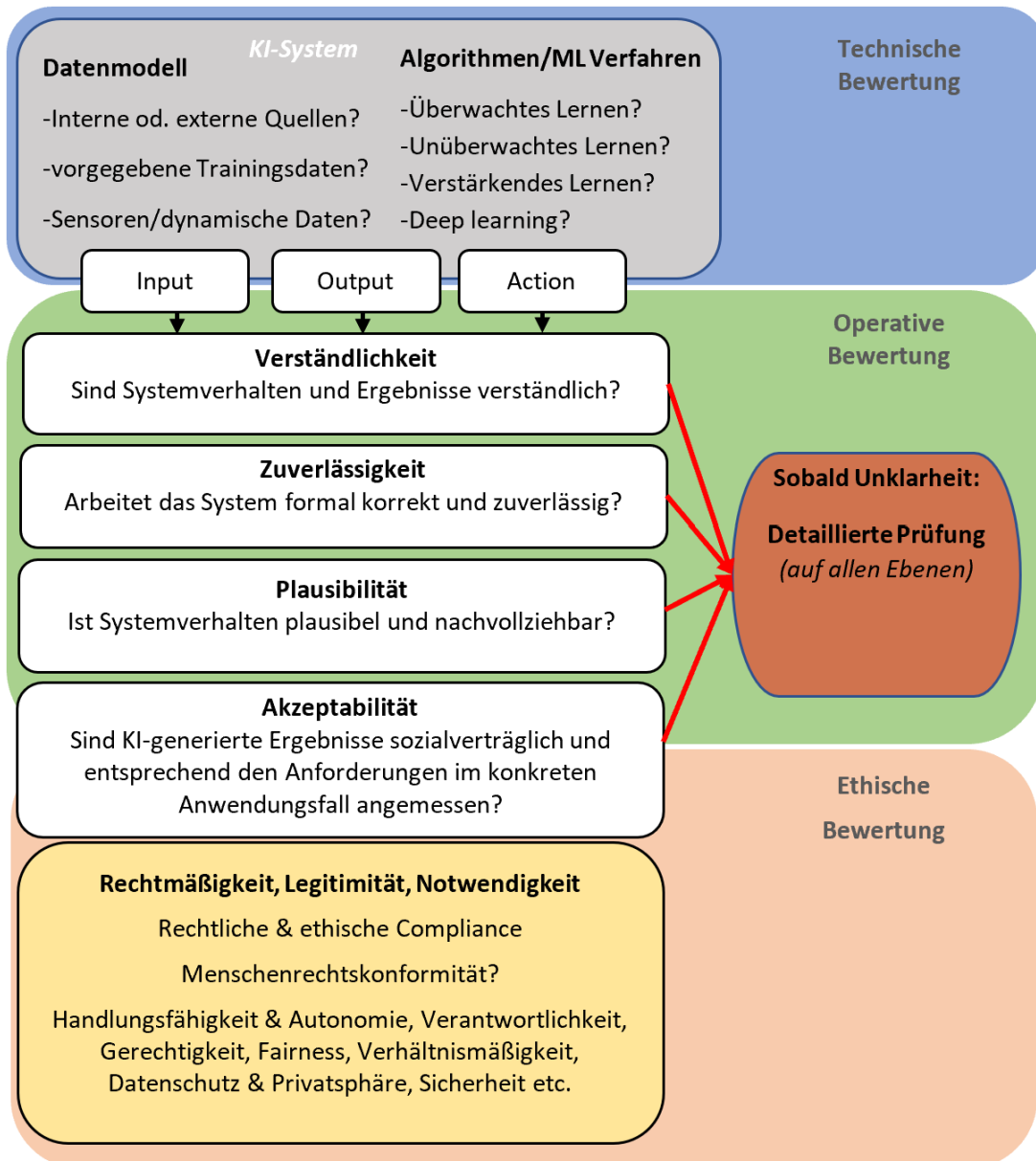


Abbildung 4: CAIL-Framework (Quelle: adaptiert von Strauß 2021a)

Der Ansatz basiert auf drei verschiedenen, miteinander verbundenen analytischen Ebenen, um zwischen technischen, operativen und ethischen Einflussfaktoren der Funktionalität eines KI-Systems differenzieren zu können. Die technische Ebene bezieht sich auf die Beschaffenheit des KI-Systems und seiner technischen Funktionsweise: D. h. die Art des Datenmodells und welche Daten, in welcher Form im KI-System verarbeitet werden. Relevante Aspekte sind hier z. B.: Basiert

*Technische
Grundfunktionalität
beeinflusst
Systemverhalten*

das Datenmodell auf vortrainierten Trainingsdaten? Woher stammen die Daten – aus internen oder externen Quellen? Wie geht das KI-System mit neuen Daten um? Gibt es externe Systeme, die zusätzliche Daten einspeisen? Z. B. über Sensoren andere Anwendungen, die zu einer erhöhten Dynamik des Datenbestands führen. Zudem beeinflusst die Art des Machine Learning-Verfahrens das Systemverhalten. Es macht einen wesentlichen Unterschied, ob das KI-System auf überwachtem, unüberwachtem, verstärkendem oder vertiefendem Lernen basiert (siehe dazu genauer Abschnitt 2). Die technische Grundfunktionalität beeinflusst auch den Eignungsgrad und das Verhalten des Systems im operativen Betrieb, also wie es in bestimmten Anwendungsfällen funktioniert.

Im Zentrum steht die Einschätzung des Systems im operativen Betrieb, also wie es sich im konkreten Anwendungsfall verhält. Um die Funktionsfähigkeit zu beurteilen, können vier Hauptindikatoren herangezogen werden:

1. *Erklärbarkeit/Verständlichkeit*: Ist das Systemverhalten verständlich?
2. *Gültigkeit/Zuverlässigkeit*: Arbeitet das System formal korrekt, stabil und zuverlässig?
3. *Plausibilität*: Sind die KI-generierten Ergebnisse plausibel und nachvollziehbar?
4. *Akzeptabilität*: Sind die KI-generierten Ergebnisse sozialverträglich und entsprechend den Anforderungen im konkreten Anwendungsfall angemessen?

Hauptindikatoren zur Bewertung des Systemverhaltens

Diese Indikatoren und die dazu gehörigen Leitfragen unterstützen bei der Einschätzung, ob eine KI-Anwendung ordnungsgemäß funktioniert und brauchbare Ergebnisse produziert für daran anknüpfende Handlungen. Sobald bei einem Indikator Unklarheiten festgestellt werden, kann die Anwendung noch detaillierter analysiert werden. Das kann dabei helfen, Diskrepanzen wie unzulängliches Systemverhalten in bestimmten Anwendungssituationen zu erkennen und genauer zu überprüfen. Es daher auch ein Tool zur Bewertung des DAB-Risikos. Sind alle Indikatoren eindeutig positiv, d. h. die Leitfragen mit ja beantwortbar, dann besteht kein erkennbares Risiko. Das heißt nicht, dass es grundsätzlich fehlerfrei funktioniert. Aber für den konkreten Fall zumindest funktionsfähig und hinreichend nutzbar. Sobald Unklarheiten im Systemverhalten entstehen, sollte das System unter Einbeziehung sämtlicher Faktoren auf allen drei Ebenen genauer analysiert werden, da diese eng miteinander verbunden sind. Fehlerhafte Daten können sich beispielsweise auf das gesamte System auswirken. Ist das Datenmodell in Ordnung, aber das Systemverhalten unzuverlässig, kann das etwa auch an zusätzlichen extern eingebundenen Anwendungen liegen oder auch daran, dass der dem System zugrunde liegende ML-Ansatz nicht für den Verwendungszweck geeignet ist etc. Sind alle Ergebnisse korrekt, aber ihr Zustandekommen nicht nachvollziehbar, kann das auch an mangelhafter Transparenz und Benutzerfreundlichkeit in der Anwendung liegen. So können etwa auch zusätzliche Anforderungen an die Benutzeroberfläche oder die Systembedienung identifiziert werden und dergleichen.

Der Faktor Akzeptabilität markiert den Übergang zwischen operativer und ethischer Bewertungsebene, weil inakzeptable KI-Ergebnisse mitunter auf größere, rechtliche und ethische Probleme hindeuten können, die dann einer tiefgehenden Analyse bedürfen. Zum Beispiel kann eine Anwendung einwandfreie Ergebnisse produzieren, aber trotzdem gegen Rechtsvorschriften verstoßen, Daten-

Zusammenspiel der Bewertungsebenen unterstützt differenzierte Analyse

schutz- und Sicherheitsprobleme verursachen oder ethische Probleme verstärken. Die dritte Ebene bezieht sich daher auf rechtliche und ethische Faktoren mit drei Hauptfaktoren Rechtmäßigkeit, Legitimität und Notwendigkeit des KI-Einsatzes.

Diese Unterscheidung zwischen den verschiedenen Ebenen dient auch dazu, eine differenzierte Analyse zu ermöglichen und nicht jeden Anwendungsfall nach breiten ethischen Kriterien bewerten zu müssen, die zwar wichtig sind, aber nicht immer hilfreich in der Arbeitspraxis für sämtliche Tätigkeiten Relevanz haben. (z. B. sind Fragen zur menschlichen Autonomie bei Beurteilung eines Bilderkennungssystems nicht vordergründig relevant). Es empfiehlt sich, die Bewertung nicht nur einmal, sondern für unterschiedliche Anwendungsfälle mehrfach durchzuführen, um ein genaueres Verständnis zur Funktionsweise eines KI-Systems zu erlangen.

Dieser Bewertungsrahmen dient hier vor allem als praktikables Werkzeug in Betrieben, um sich rasch einen Überblick zu verschaffen, worauf zu achten ist, wenn ein KI-System eingesetzt wird oder werden soll. Er kann etwa als unterstützendes Instrument für Betriebsräte verwendet werden, um im Unternehmen zu einem besseren Grundverständnis über den KI-Einsatz und den damit verbundenen innerbetrieblichen Herausforderungen zu gelangen. Das kann eine einfache Grundlage für interne Schulungen sein oder auch als Entscheidungshilfe bei der Abwägung, ob und wie KI-Systeme eingesetzt werden sollen – idealerweise im gemeinsamen Austausch zwischen Arbeitgeber und Belegschaft. Insofern kann es einen Beitrag zur sozialverträglichen Gestaltung und Nutzung KI-basierter Technologien leisten.

*CAIL-Framework
als Instrument zur
Vermittlung kritischer
KI-Kompetenz*

6.2 HANDLUNGSEMPFEHLUNGEN FÜR BETRIEBE

Die folgenden Empfehlungen richten sich an Betriebe, die noch wenig Erfahrung mit KI-basierten Technologien haben und sich einen Überblick verschaffen wollen, welche Faktoren bei der Entscheidung, ob und wie ein KI-System eingeführt werden soll, zu beachten sind. Die Grafik unten soll dabei einen raschen Einstieg ermöglichen und eine Grundbedingung für konstruktiven KI-Einsatz verdeutlichen: Eine klare Differenzierung zwischen den verschiedenen Prozessen von der Planung, Einführung, Anwendung und Adaption. Denn KI-Systeme bringen andere Automatisierungsformen ins Unternehmen mit einer höheren Volatilität und Veränderungsdynamik, die sich nach Einführung dann auch auf die Arbeitsprozesse niederschlägt. Daher braucht es hier gekoppelt an die Anwendungen auch explizite Adaptionprozesse, um die Qualität und den praktischen Nutzen in Einklang mit der sozialen und organisationalen Arbeitspraxis zu gewährleisten.

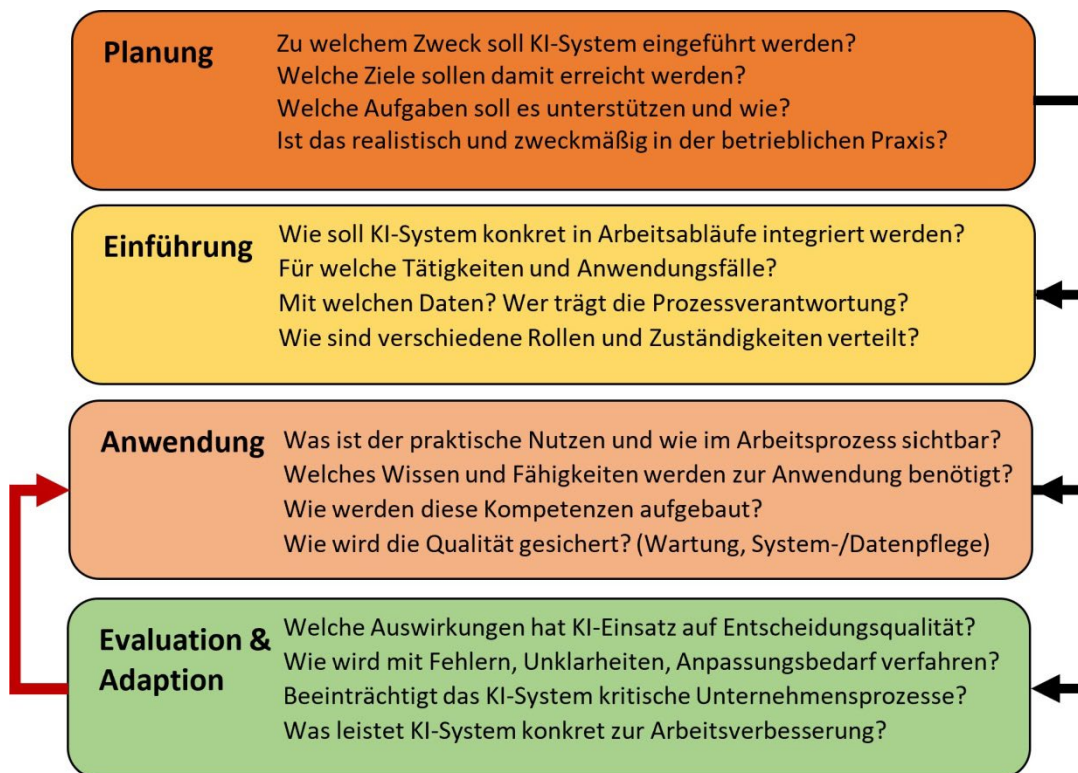


Abbildung 5: Entscheidungshilfe bei der KI-Einführung

MEHRAUFWAND DURCH KI-EINSATZ NICHT UNTERSCHÄTZEN

Die Art und Weise, wie KI-Systeme in betriebliche Abläufe integriert sind, beeinflusst ihren Mehrwert. Unternehmen, die KI-Einsatz anstreben, sollten sich zunächst bewusst machen, dass damit nicht weniger, sondern zunächst deutlich mehr Arbeitsaufwand verbunden ist. Das betrifft nicht nur die Planung, sondern die gesamte Prozesskette von der Planung, Einführung, bis zur Nutzung. Zudem kann aufgrund der Besonderheiten von KI-Systemen auch ein erhöhter Anpassungsbedarf entstehen, damit die Systeme sinnvoll und mit praktischem Nutzen in Arbeitsabläufe integriert werden können.

KI-EINSATZ REALISTISCH PLANEN UND BEDARF VORAB KLÄREN

In der Planungsphase ist vorab zu klären, warum, ob, wie und welche KI-Technologie eingeführt werden soll. Wichtig ist dabei auch, nicht nur die Erwartungen an den KI-Einsatz, sondern auch deren Realisierbarkeit und deren mittel- und längerfristige Folgen im Betrieb zu berücksichtigen. Das ist für jede Art von Technologie-Einsatz relevant und insbesondere bei IT-Projekten, zu denen auch die KI-Einführung zählt. Versäumnisse in der Planung erschweren es, einen Mehrwert zu erzielen.

EINSATZZWECK, KLARE ZIELE UND NICHT-ZIELE DES KI-EINSATZES DEFINIEREN

Es braucht daher Klarheit darüber, ob, zu welchem Zweck und für welche Tätigkeiten in welchem Ausmaß mit KI automatisiert unterstützt werden sollen und welche Ziele damit verfolgt werden. Eine zentrale Grundbedingung hierfür ist zudem, dass in jedem Fall menschliche Intervention und Entscheidungshoheit im Arbeitsablauf integriert ist. Das Ausmaß dieser Intervention ist wiederum abhängig davon, inwieweit mit KI automatisiert werden soll und welche Folgen das hat. Dafür braucht es auch Bewusstsein.

AUTOMATISIERUNGSGRADE BEWUSST MACHEN UND MITGESTALTEN

Die Beachtung dieser Aspekte – was und warum soll mit KI automatisiert werden und was nicht – ist bei KI-basierten Technologien umso wichtiger, weil ihr Einsatz mit Besonderheiten verbunden ist, die zu bedenken sind. Das betrifft zunächst den Umstand, dass KI-Einsatz keine „Effizienzmaschine“ ist, sondern neue Formen der Automatisierung ins Unternehmen bringt. Daraus allein entstehen weder mehr Effizienz noch Produktivitätsgewinne, noch sonstige der zahlreichen erhofften Vorteile (siehe Abschnitt 1). Die aktive Gestaltung des Automatisierungsgrades ist daher entscheidend.

KOMPLEXITÄT UND VERÄNDERUNGSDYNAMIK VON ARBEITSPROZESSEN BEDENKEN

KI-Einsatz bringt auch mehr Komplexität und Veränderungsdynamik in Arbeitsprozesse: 1) weil KI-generierte Ergebnisse nicht immer eindeutig sind, sondern mitunter schwer interpretierbar. 2) weil bei KI mit höherem Wartungsaufwand und tendenziell häufigeren Veränderungen im KI-System zu rechnen ist (z. B. durch aktualisierte Datenmodelle, Veränderungen/Erweiterungen der Trainingsdaten, kürzere Update-Zyklen etc.). Das kann die Nutzbarkeit des Systems und damit die Qualität der damit verbundenen Arbeitspraktiken beeinflussen und sich hemmend auf die Produktivität auswirken.

BEDARFSANALYSEN, UM ERWARTETE VORTEILE AUF IHRE REALISIERBARKEIT ZU PRÜFEN

Nach erfolgreicher Einführung können KI-Systeme helfen, Komplexität in Arbeitsprozessen im Idealfall besser zu bewältigen, indem sie zum Beispiel: Beschäftigte bei Routineaufgaben unterstützen und entlasten, entlang der Wertschöpfungskette Arbeitsprozesse optimieren helfen oder brauchbare Zusatzinformationen für Entscheidungsprozesse generieren. Gelingt das, sind u. a. Effizienz- und Produktivitätsgewinne möglich. Bevor ein KI-System eingeführt wird, sollte aber im

Rahmen von Bedarfsanalysen geklärt werden, inwieweit das auch tatsächlich realisierbar ist in konkreten Arbeitsprozessen.

INTEGRIERBARKEIT VON KI IN BESTEHENDE ARBEITSORGANISATION PLANEN

Darauf aufbauend braucht es Strategien und Abläufe, um zu klären, wie eine KI-Anwendung konkret in bestehende Arbeitsabläufe und Praktiken integriert werden kann und soll. Dabei sollte die Integration der Technologie so erfolgen, dass sie bestehende Arbeitsprozesse, Arbeitspraktiken und die Arbeitsorganisation sinnvoll ergänzt und nicht ersetzt. Dazu braucht es auch die Einbindung der Belegschaft, um Arbeitspraktiken durch den KI-Einsatz sinnvoll zu unterstützen, sodass ein Mehrwert geschaffen werden kann.

KLARE KONTEXTUALISIERUNG UND NUTZUNGS-TRANSPARENZ STÄRKEN

Neben der oft genannten Transparenz von KI-Systemen braucht es auch eine Transparenz der Nutzung. Das erfordert neben Klarheit über Nutzungszweck und Ziele von KI-Anwendungen auch eine klare Kontextualisierung, d. h. der KI-Einsatz sollte nicht beliebig und intransparent auf andere Arbeitsbereiche erweitert werden, sondern nur für bestimmte Anwendungsfälle und konkrete Tätigkeiten gelten, sofern ein praktischer Nutzen gegeben ist.

KENNZEICHNUNG VON KI-ANWENDUNGEN UND KI-GENERIERTEN ERGEBNISSEN

Damit verbunden trägt auch die Kennzeichnung von KI-basierten Anwendungen und KI-generierten Ergebnissen im Unternehmen für betriebliche Entscheidungen zur Stärkung der Nutzungs-Transparenz bei. Die Kennzeichnung unterstützt auch, mehr Bewusstsein für etwaige genauere Prüfungen von Qualität und praktischem Nutzen für Mitarbeiter:innen in den betroffenen Arbeitsprozesse zu schaffen.

ANWENDUNGS- UND PROZESSQUALITÄT GEWÄHRLEISTEN UND RESSOURCENBEDARF BEDENKEN

Beim Nutzungsprozess sind menschliche (und teils manuelle) Arbeitsschritte zur Datenaufbereitung und Datenpflege (z. B. Trainingsdaten, Testdaten, Eingangsdaten), Qualitätsmanagement, Fehlerdetektion und Fehlerkorrektur und Wartung von KI-Systemen, Ergebniskontrolle, und Handlungs- & Entscheidungsmanagement („verwenden oder verwerfen“) etc. explizit zu bedenken, um den tatsächlichen, KI-gestützten Arbeitsablauf und den damit verbundenen zusätzlichen Arbeits- und Ressourcenaufwand sichtbar und damit bewältigbar zu machen. Ein sehr unterschätzter Ressourcenfaktor ist hierbei die benötigte Zeit. Zudem beeinflussen auch etwaige Abhängigkeiten zu externen Technologien die Qualität der KI-gestützten Arbeitsprozesse. Diese Faktoren sind wichtig, um ein bestimmtes Niveau von Anwendungs- und Prozessqualität zu gewährleisten.

SCHULUNGS- UND QUALIFIKATIONSBEDARF BEDENKEN UND CAIL STÄRKEN

Es ist klar, dass mittelfristig der Bedarf nach Kompetenzen und Fachwissen im Umgang mit KI weiter zunimmt. Der Kompetenzaufbau sollte aber nicht nur auf technische Skills und Knowhow beschränkt sein. Vielmehr umfasst kritische KI-Kompetenz die Fähigkeit, KI-generierte Ergebnisse und ihre Aussagekraft in

situativen Kontexten zu erkennen, zu verstehen sowie ihre Brauchbarkeit und Tauglichkeit in konkreten Arbeitsprozessen zu beurteilen (siehe Abschnitt 6). Die Etablierung und Stärkung kritischer KI-Kompetenzen ist ein zentraler Erfolgsfaktor für den KI-Einsatz.

BETRIEBLICHE MITBESTIMMUNG ZULASSEN UND AKTIV FÖRDERN

Die aktive Einbindung von Betriebsräten und die Berücksichtigung von Mitarbeiterinteressen, Sorgen und Ängsten ist ein zentraler Aspekt zur Förderung von Akzeptanz, Akzeptabilität und Vertrauen bei der KI-Einführung und -Nutzung. Ein gemeinsam mit der Belegschaft erarbeitetes Grundverständnis von KI-Einsatz im Unternehmen unter Beachtung der damit verbundenen Potenziale, Möglichkeiten, Risiken und Grenzen kann erheblich zur Verbesserung der Arbeitsqualität, Mitarbeiterzufriedenheit und damit auch der Produktivität beitragen. Betriebsräte haben hierbei eine zentrale Funktion, die gerade im Zuge des KI-Einsatzes gestärkt werden sollte, um einen bewussten und konstruktiven Umgang mit diesen Technologien zu etablieren.

Die Berücksichtigung der genannten Faktoren ist wesentlich, um KI-Einsatz so zu gestalten, dass ein praktischer Nutzen entsteht. Andernfalls besteht die Gefahr, dass Komplexität und Aufwand unterschätzt werden. Das kann zu steigender Verunsicherung im Betrieb und zusätzlicher Arbeitsbelastung führen und damit auch wenig Akzeptanz für den Technologieeinsatz. In Summe kann sich das letztlich negativ auf das Vertrauen ins Unternehmen und die Produktivität auswirken. Zentrale Grundvoraussetzung für den konstruktiven Einsatz KI-basierter Technologien ist daher auch die sozialverträgliche Einbindung in die betriebliche Arbeitsorganisation.

7 ZUSAMMENFASSUNG UND AUSBLICK

Ein grundlegender Aspekt für den konstruktiven Umgang mit KI-basierten Technologien ist folgende Erkenntnis: KI kann keine Arbeit abnehmen, sie verändert sie nur. Diese triviale Aussage dient hier als Aufhänger um wesentliche Ergebnisse der Analyse zusammenzufassen. Wie sich in unterschiedlicher Ausprägung zeigt, führt KI-Einsatz bislang nicht dazu, dass ganze Arbeitsbereiche automatisiert werden. Vielmehr nimmt die automatisierte Unterstützung bestimmter Tätigkeiten zu und es werden Veränderungen in der Arbeitsorganisation sichtbar. Eine zentrale Grundfrage für die konstruktive und sozialverträgliche Nutzung KI-basierter Technologien ist daher, wie sich Arbeitsprozesse und betriebliche Praxis durch deren Einsatz verändern, und ob diese Veränderungen mit Unternehmenszielen, der Arbeitsorganisation und realen Arbeitspraktiken vereinbar sind. Die Beantwortung dieser Frage erfordert innerbetriebliches Wissen über die Funktionsweisen und Grenzen der Technologien in konkreten Anwendungskontexten.

KI-basierte Automatisierung ist komplexer, dynamischer und volatiler als klassische Automatisierung. Daher erfordert konstruktive KI-Nutzung auch mehr Bewusstsein für diese Besonderheiten und die damit verbundenen Risiken wie Deep Automation Bias (Abschnitt 4). Ein kritischer Faktor des KI-Einsatzes ist in jedem Anwendungsfall die Anwendungs- und Prozessqualität. Die Erwartung, KI würde per se Arbeit erleichtern, Produktivität steigern und Qualität verbessern, ist nicht haltbar. Im Gegenteil ist ein praktischer Mehrwert nur dann erzielbar, wenn Klarheit darüber herrscht, ob und wie KI-Anwendungen in Arbeitsprozesse integriert werden können, und wie und mit welchen Maßnahmen die Qualität und Stabilität dieser Prozesse gewährleistet wird. KI kann rasch Ergebnisse produzieren, aber die Interpretation und Überprüfung bleibt eine zentrale menschliche Aufgabe, die künftig noch wichtiger wird. Gleiches gilt für Tätigkeiten des Datenmanagements (Datenaufbereitung, Datenpflege, Bereinigung) sowie der Qualitätssicherung und Fehlerkorrektur. Es entstehen also neue Aufgaben, die auch mehr Arbeitsaufwand mit sich bringen. Das gilt insbesondere für das Qualitätsmanagement der KI-Anwendungen. Eine Prozess-Sicht auf die Einbettung von KI-Anwendungen in Arbeitsabläufe ist daher wichtig, um die Tätigkeiten so zu planen, dass die möglichen Entlastungen in einem Arbeitsbereich nicht zur Überlastung in anderen Arbeitsbereichen führen. Zum Beispiel kann KI die Produktion von multimedialen Inhalten erheblich beschleunigen. Zugleich entsteht aber auch mehr Arbeitsaufwand zur Qualitätskontrolle, Anpassung und Korrektur der Inhalte etc. Gleiches gilt für Prozessoptimierungen, die nur mit ausreichender Datenqualität möglich sind, die wiederum explizite Qualitätssicherungsaufgaben umfasst, die nicht einfach an KI ausgelagert werden können, sondern mehr innerbetrieblichen Arbeitsaufwand bedeuten. Um den praktischen Mehrwert von KI im Unternehmen einzuschätzen, braucht es auch einen Abgleich zwischen Erwartungen und real erzielbarem Nutzen (Abschnitt 5.1). Daher ist entscheidend, dass der Nutzungszweck und die Ziele des KI-Einsatzes im Vorfeld geklärt werden.

KI kann Arbeit nicht abnehmen, doch sie verändert die Arbeitsorganisation

Mehrwert mit KI braucht Klarheit über Nutzen und Integrierbarkeit in Arbeitspraktiken

Das Bewusstsein für diese Grundanforderungen ist je nach Anwendungsbereich unterschiedlich ausgeprägt. Wie in Abschnitt 5 gezeigt, besteht eine Art Kluft zwischen der Nutzung von KI als bewusstes Werkzeug in einem Expertensystem und als „Nebenbei“-Technologie ohne klaren Tätigkeitsbezug und klare Qualitätsanforderungen. In ersterem ist tendenziell ein relativ hohes Maß an Fachexpertise und Problembewusstsein gegeben. Das begünstigt die Einführung und Integration der Technologie in Einklang mit bestehenden Arbeitspraktiken zur Unterstützung konkreter Tätigkeiten. So lässt sich mit dem Einsatz von KI-Systemen auch ein praktischer Nutzen erzielen. Wird KI ohne bestimmte Expertise und konkrete Planung für die Integration in Arbeitsprozesse eingeführt, sind Probleme in der praktischen Nutzung deutlich wahrscheinlicher. Zum einen kann mangelhafte Einbeziehung der Arbeitnehmer:innen dazu führen, dass Anwendungsdefizite erst nach Technologie-Einführung sichtbar werden. Erfolgt der KI-Einsatz zu vorschnell, können neue technisch-organisatorische Abhängigkeiten entstehen, die sich eher belastend auf die Effizienz der Arbeitsabläufe und damit auch negativ auf die Akzeptanz und Akzeptabilität im Unternehmen auswirken (Abschnitt 5.3).

Ein wachsendes und unterschätztes Problemfeld ist die schleichende Einführung von KI über Standardsoftware wie z. B. weit verbreitete Betriebssysteme, die sukzessive KI-Funktionen in verschiedene Office-Anwendungen integrieren (siehe Abschnitt 5.4). Das kann erhebliche Folgen für die Arbeitsorganisation quer durch alle möglichen Branchen haben. Neben dem Einsatz zur Unterstützung von Arbeitsprozessen können KI-Systeme auch anderen Zwecken dienen, die nicht ohne weiteres ersichtlich sind. Anwendungen werden immer vernetzter, dadurch werden auch Datenflüsse immer schwerer kontrollierbar. Generell können verdeckte KI-Systeme zu erheblichen Problemen in Unternehmen führen. Neben Gefahren zunehmender Mitarbeiterüberwachung und Datenschutzproblemen (sei es proaktiv durch Arbeitgeber oder verdeckt über zusätzlich in Anwendungen erfasste Daten) können dadurch auch zusätzliche Probleme für die Sicherheit entstehen. Sind etwa zentrale Geschäftsprozesse abhängig von KI-Systemen, kann das die Vulnerabilität des Unternehmens erhöhen. Vor allem, wenn KI-Systeme zu einem Bestandteil der unternehmensinternen kritischen Infrastruktur werden. Aufgrund der starken Marktmacht einzelner großer Softwareanbieter besteht hier i. d. R. ein hohes Abhängigkeitsverhältnis. Die meisten Betriebe sind auf diese Systeme angewiesen und haben daher oftmals nur wenig Möglichkeiten zur Mitgestaltung, um KI-Anwendungen abgestimmt auf betriebliche Anforderungen nutzbringend einzusetzen. Das kann sich negativ auf die notwendige Kontextualisierung auswirken, also die gezielte Einbindung von KI-Anwendungen für bestimmte Nutzungskontexte entsprechend den Anforderungen der betroffenen Arbeitsabläufe. Die Rolle von Standardsoftware bei der KI-Einführung verdeutlicht auch den wachsenden Bedarf nach Open-Source KI-Systemen. Denn die Verbreitung von KI ist letztlich auch Bestandteil neuer Geschäftsmodelle großer Technologiebetreiber. Unternehmen sollten daher auch bedenken, welche innerbetrieblichen Daten von KI-Systemen verarbeitet werden. Im Zuge der Entscheidung ob KI eingeführt werden soll, empfiehlt sich daher auch, Open-Source Varianten für KI-Anwendungen in Erwägung zu ziehen, um einige der mit Standardanwendungen einher gehenden Probleme zu vermeiden.

Fachexpertise und Problembewusstsein begünstigen praktischen Nutzen

Unterschätztes Problemfeld: Schleichende KI-Einführung über Standardsoftware

Der erzielbare Mehrwert (Abschnitt 5.1) und die Auswirkungen auf die Arbeitsorganisation hängen wesentlich von der Art der Nutzung von KI-Anwendungen und ihrer Integration in konkrete Tätigkeiten ab. Abseits hoher Erwartungen und hoher Automationsgrade mit KI ist das konkrete Ausmaß, wie viel einer Tätigkeit im betrieblichen Arbeitskontext mit KI automatisiert werden soll, gestaltbar. Ein grundsätzlich hoher KI-Automationsgrad impliziert noch keinen praktischen Nutzen. D. h. es gilt zu unterscheiden zwischen dem, was eine KI-Anwendung für bestimmte Arbeitsschritte leisten kann, und inwieweit das auch in der betrieblichen Arbeitspraxis sinnvoll umsetzbar ist. Also inwieweit die KI-Anwendung in bestehende Arbeitsabläufe integrierbar ist. Entgegen gängiger Annahmen eignet sich KI weniger für komplexe Aufgaben, sondern vielmehr für einfachere, aber mitunter aufwändige Routinetätigkeiten. Hierbei ist auch die Form der Wissensarbeit und damit verbundene Tätigkeiten wesentlich. Wie in Abschnitt 3 ausgearbeitet, umfasst Wissensarbeit i.d.R. eine Kombination aus „manuellen“ bzw. repetitiven Routinetätigkeiten und komplexen Aufgaben. Letztere sind nicht ohne weiteres mit KI automatisierbar. Der konstruktive, sinnvolle Einsatz von KI in der Wissensarbeit ist möglich, wenn die Einbettung der Technologie derart gelingt, dass Routineaufgaben automatisiert werden, sodass Freiräume für komplexere Aufgaben entstehen. Problematisch wird es, wenn KI-Einsatz dazu führt, dass Wissensarbeit überformalisiert und monotonisiert wird. Das kann zu Demotivation führen und folglich auch zu geringerer Produktivität. Damit stünde dann auch die stabilisierende Funktion von Wissensarbeit in Unternehmen in Frage.

In Anbetracht des derzeitigen Entwicklungsstands steht ein größerer Wandel der Wissensarbeit mittels KI-Automatisierung entsprechender Tätigkeiten aber noch relativ am Anfang. Zwar gibt es unzählige Anwendungsfelder, aber hier zeigt sich, dass es zu einer Erweiterung bzw. Verlagerung menschlicher Tätigkeiten kommt. KI-Anwendungen können – sofern die Ergebnisse zuverlässig sind – Wissensarbeit und Wissensproduktion in vielerlei Hinsicht ergänzen. Jedoch nicht durch Generierung neuen Wissens, sondern nur durch Kombination und Aufbereitung bestehender Daten als Informationen. Ihre Interpretation und die Entscheidung, wie brauchbar oder unbrauchbar sie in bestimmten Kontexten sind, obliegt weiterhin dem Menschen und wird zu einer immer wesentlicheren Aufgabe. Das Verstehen und Interpretieren KI-generierter Information und ihre Einbettung in konkrete Anwendungskontexte wird daher Teil des Aufgabenspektrums von Wissensarbeit. D. h., Wissensarbeit benötigt zusätzlich zur „klassischen“ Problemlösungskompetenz nun auch die Kompetenz zu erkennen, ob und wie KI-Einsatz bei einer Tätigkeit sinnvoll ist oder nicht. Das ist mit einem wachsenden Bedarf an Wissen verknüpft, um im Sinne einer Critical AI Literacy zu einem klaren Grundverständnis darüber zu gelangen, was KI leisten kann und was nicht und wie KI-produzierte Resultate zu interpretieren sind.

Wie in Abschnitt 4.2 aufgezeigt, stehen den zentralen Grundanforderungen an KI-Systeme allem voran: Zuverlässigkeit, Nachvollziehbarkeit, Überprüfbarkeit und Interpretierbarkeit auch Anforderungen auf Anwender:innenseite gegenüber. Im Kern ist damit der Bedarf nach kritischer KI-Kompetenz angesprochen. Wie in Abschnitt 6 ausgeführt, ist die Etablierung einer Critical AI Literacy entscheidend, um die betriebliche und individuelle Handlungsfähigkeit im Umgang mit KI zu stärken. Als Hilfestellung für Betriebe und Organisationen wurde dazu mit dem CAIL-Framework ein einfacher Bewertungsrahmen ausgearbeitet. Dieses Instrument kann zur Vermittlung des Basiswissens dienen, das zur

Art der Nutzung und Automatisierungsgrad mitgestalten

KI ergänzt Wissensarbeit, wenn es bei Routineaufgaben entlastet und tatsächlich neue Freiräume schafft

Critical AI Literacy wird zur Aufgabe von Wissensarbeit

Grundanforderungen an KI-Systeme und deren Nutzung

Beurteilung des praktischen Nutzens und der Grenzen von KI-Anwendungen und KI-generierten Ergebnissen benötigt wird. Die Handlungsempfehlungen in Abschnitt 6.2 richten sich vor allem an Unternehmen und Betriebsräte, die noch wenig Erfahrung mit KI-basierten Technologien haben und einen Überblick zu relevanten Faktoren benötigen, was es bei der Entscheidung, ob und wie ein KI-System eingeführt werden soll, zu bedenken gilt. Diese Instrumente tragen idealerweise auch dazu bei, dass Betriebe gemeinsam mit Belegschaft und Interessensvertretungen zu einer konstruktiven, sozialverträglichen Gestaltung und Nutzung KI-basierter Technologien gelangen.

Es besteht wenig Zweifel, dass KI die Organisation von Arbeit in den nächsten Jahren deutlich verändern wird. Bei genauerer Betrachtung ist das Ausmaß der weitgehend angekündigten Revolution oder Disruption der Arbeitswelt jedoch in vielerlei Hinsicht offen. Zudem gibt es je nach Bereich unterschiedlich ausgeprägte Diskrepanzen zwischen hohen Erwartungen und teils relativ geringem Realisierungsgrad. Das zeigt sich auch in den Nutzungszahlen von KI-Technologien, die zwar langsam steigen, aber zugleich wachsen auch die Sorgen vor Verschlechterung der Arbeitsbedingungen. Es ist daher nicht ausgeschlossen, dass auf die hohe Erwartungshaltung mittelfristig eine gewisse Ernüchterung folgt. Denn trotz enormer Potenziale steht der KI-Einsatz in vielen Bereichen erst am Anfang und teils ist noch wenig praktischer Nutzen sichtbar.

Die Transformation der Wissensarbeit mit KI ist in vielerlei Hinsicht noch am Anfang und mit vielen offenen Fragen verbunden. Das derzeitige Entwicklungsstadium lässt sich als Teil eines größeren, institutionellen und gesellschaftlichen Lernprozesses darüber begreifen, ob und wie KI-basierte Technologien konstruktiv und praktikabel einsetzbar sind. Dabei muss zunächst auch in Unternehmen geklärt werden, was Technologieeinsatz leisten kann, ob das in der Arbeitspraxis zielführend ist und welche Folgen sich daraus für die jeweiligen Arbeitsbereiche ergeben. In vielen Unternehmen hat gerade erst eine Art Lern- und Experimentierphase mit KI begonnen. Diese trägt idealerweise dazu bei, über die Organisation von Arbeit, die Bedeutung sozialverträglicher Arbeitsbedingungen und Mitarbeiterzufriedenheit neu zu reflektieren. Wenn es gelingt, in den nächsten Jahren zu einem tieferen Grundverständnis von KI-basierten Technologien und der Etablierung einer kritischen Kompetenz im Umgang damit zu gelangen, ist ein wesentliches Lernziel für die konstruktive Nutzung dieser Technologien erreicht.

Diskrepanz zwischen hohen Erwartungen und unklarem Realisierungsgrad

Gesellschaftlicher Lernprozess über die konstruktive Nutzung von KI im Gange

8 LITERATUR

- Abrahamczyk, M. (2022), Systemausfall droht: Großer Tesla-Rückruf, 01.09., t-online, https://www.t-online.de/auto/elektromobilitaet/elektroauto/id_100046900/systemausfall-droht-grosser-tesla-rueckruf.html.
- Albrecht, T., Kellermann, C. (2020), Künstliche Intelligenz und die Zukunft der digitalen Arbeitsgesellschaft: Konturen einer ganzheitlichen Technikfolgenabschätzung. Working paper No.200, Hans-Böckler-Stiftung <https://www.econstor.eu/bitstream/10419/228963/1/1743497164.pdf>.
- Anlen, S., Llorente, R.V. (2024), Spotting the deepfakes in this year of elections: how AI detection tools work and where they fail. April 15, <https://reutersinstitute.politics.ox.ac.uk/news/spotting-deepfakes-year-elections-how-ai-detection-tools-work-and-where-they-fail>.
- APA – Austria Presse Agentur (2023), Leitlinie zum Umgang mit künstlicher Intelligenz. Stand: Juli 2023. <https://apa.at/wp-content/uploads/2023/07/Leitlinie-zum-Umgang-mit-kuenstlicher-Intelligent-2023-2.pdf>.
- Bager, J., Wiegand, D., Wischner, S. (2024), Copilot-Alternativen: Welche KI bei Office-Aufgaben besser helfen kann. Heise 05. April, <https://www.heise.de/ratgeber/Copilot-Alternativen-Welche-KI-bei-Office-Aufgaben-besser-helfen-kann-9659097.html>.
- Bailer, W., Thallinger, G., Krawarik, V., Schell, K. und Ertelthaler, V. (2022), AI for the media industry: application potential and automation levels, International Conference on Multimedia.
- Barenkamp, M., Rebstadt, J., Thomas, O. (2020), Applications of AI in classical software engineering. *AI Perspect* (2)1. <https://doi.org/10.1186/s42467-020-00005-4>.
- Barr, A.; Feigenbaum, E. (1981), *The Handbook of Artificial Intelligence*; HeurisTech Press: Stanford, CA, USA, Volume 1.
- Bastian, M. (2024), Sprachmodelle können große Dokumente laut neuem Benchmark nicht zuverlässig verarbeiten. 18. Juli, the decoder, <https://the-decoder.de/sprachmodelle-koennen-grosse-dokumente-laut-neuem-benchmark-nicht-zuverlaessig-verarbeiten/>.
- BBC (2024), BBC News Labs, <https://www.bbc.co.uk/rdnewslabs/>.
- Beer, K. (2024), Deutscher Lehrerverband: Lehrer brauchen datenschutzkonformen Zugang zu KI. Heise 04.01., <https://www.heise.de/news/Deutscher-Lehrerverband-Lehrkraefte-brauchen-datenschutzkonformen-Zugang-zu-KI-9586621.html>.
- Beiersmann, S. (2018), Niederlande: Sammlung von Microsoft-Office-Telemetriedaten verstößt gegen DSGVO. ZDNet 15.11., <https://www.zdnet.de/88347263/niederlande-sammlung-von-microsoft-office-telemetriedaten-verstoest-gegen-dsgvo/>.
- Bertelsmann (2022), Was Deutschland über Algorithmen und Künstliche Intelligenz weiß und denkt. Ergebnisse einer repräsentativen Bevölkerungsumfrage: Update 2022. Bertelsmann-Stiftung, <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/was-deutschland-ueber-algorithmen-und-kuenstliche-intelligenz-weiss-und-denkt-all>.
- Bhagwat, N. et al., (2018), Modeling and prediction of clinical symptom trajectories in Alzheimer’s disease using longitudinal data, *PLOS Computational Biology* 14(9), e1006376 doi.org/10.1371/journal.pcbi.1006376.
- BMBWF – Bundesministerium für Bildung, Wissenschaft und Forschung (2023a), Die Verwendung KI-basierter Tools beim Erstellen abschließender Arbeiten – Potenziale, Risiken und beurteilungsrelevante Aspekte. Informationen für Betreuer/innen abschließender Arbeiten an AHS und BMHS https://www.bmbwf.gv.at/dam/jcr:3bc6eb26-f4b1-499c-a601-675e7fd6fa0f/ki_abarb.pdf.
- BMBWF – Bundesministerium für Bildung, Wissenschaft und Forschung (2023b), Auseinandersetzung mit Künstlicher Intelligenz im Bildungssystem. Stand: 29.08.2023 https://www.bmbwf.gv.at/dam/jcr:b77eacd7-3926-460e-955a-0754e419e577/ki_bildungssystem.pdf.

- Buchanan, B.G. (2005), A (Very) Brief History of Artificial Intelligence. *AI Mag.*, 26, 53–60.
- Chibanguza K., Kuß, C., Steege, H. (2022), *Künstliche Intelligenz: Recht und Praxis automatisierter und autonomer Systeme*. 1. Auflage. Nomos: Baden-Baden.
- Cunnigham, A. (2023), Microsoft 365's AI-powered Copilot is like an omniscient version of Clippy. *Ars Technica* 16.03., <https://arstechnica.com/information-technology/2023/03/microsoft-365s-ai-powered-copilot-is-like-an-omniscient-version-of-clippy/>.
- Dabbous, A., Aoun Barakat, K. and Merhej Sayegh, M. (2022), Enabling organizational use of artificial intelligence: an employee perspective, *Journal of Asia Business Studies*, 16(2), 245-266. <https://doi.org/10.1108/JABS-09-2020-0372>.
- Davenport, T., Kalakota, R. (2019), The potential for artificial intelligence in healthcare, *Future healthcare journal* 6(2), 94.
- Davenport, T.H., De Long, D. W., Beers, M. C. (1998): Successful knowledge management projects. *Sloan Management Review* 2(2):43-57.
- Deloitte (2023), Generative AI's fast and furious entry into Switzerland. Usage and attitudes of the Swiss workforce towards Generative AI. <https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/technology/deloitte-ch-en-generative-ai.pdf>.
- Deloitte (2024), 2024 Global Human Capital Trends: Thriving beyond boundaries: Human performance in a boundless world. Deloitte Insights. https://www2.deloitte.com/content/dam/insights/articles/glob176836_global-human-capital-trends-2024/DI_Global-Human-Capital-Trends-2024.pdf.
- Die Presse (2023), Wie der ORF bereits mit Künstlicher Intelligenz arbeitet. *Die Presse Online*, 11.10.2023, <https://www.diepresse.com/17731209/wie-der-orf-bereits-mit-kuenstlicher-intelligenz-arbeitet>.
- Dinges, S. (2020), Microsoft Office 365: Überwachung des Verhaltens Angestellter soll beschränkt werden. *Netzpolitik* 02.12., <https://netzpolitik.org/2020/microsoft-office-365-ueberwachung-des-verhaltens-angestellter-soll-beschraenkt-werden/>.
- Domingos, P. (2012), A Few Useful Things to Know about Machine Learning. *Commun. ACM*, 55, 78–87.
- Dörr, K.N. (2016), Mapping the field of Algorithmic Journalism. *Digital Journalism*, 4(6):700-722. DOI: <https://doi.org/10.1080/21670811.2015.1096748>.
- Dreyfus, H.L. (1992), *What Computers Still Can't Do: A Critique of Artificial Reason*; MIT Press: Cambridge, MA, USA.
- Drucker, P.F. (1999), Knowledge-Worker Productivity: The Biggest Challenge. *California Management Review* (41):2, 79-94.
- Drucker, P.F. (2001), *Management Challenges for the 21st Century*. New York: Harper Business.
- DSK – Datenschutzkonferenz (2022), Festlegung der Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder, Stand: 24.11.2022, AG DSK „Microsoft-Onlinedienste“. https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DSK/DSKBeschluessePositionspapiere/104DSK-Festlegung-Microsoft-Onlinedienste.pdf?__blob=publicationFile&v=1.
- Dupuis, I. (2016), Data journalism. *Studienbrief*, Deutsches Journalistenkolleg. <https://www.journalistenkolleg.de/documents/10157/161315/Data+Journalism.pdf/87bf3dd7-32b4-4f0a-bdea-59598f002979>.
- Edwards, J.; Perrone, A.; Doyle, P.R. (2020), Transparency in Language Generation: Levels of Automation. In *Proceedings of the 2nd Conference on Conversational User Interfaces CUI'20*, Bilbao, Spain, 9–10 July.
- Emmanouilidis, C., Pistofidis, P., Bertonecelj, L., Katsouros, V., Fournaris, A., Koulamas, C., Ruiz-Carcel, C. (2019), Enabling the human in the loop: Linked data and knowledge in industrial cyber-physical systems, *Annual Reviews in Control*, 47, pp. 249-265.
- Floridi, L. et al. (2018), AI4People – an ethical framework for a good AI society. Opportunities, risks, principles, and recommendations. In: *Minds & Machines* 28, pp. 689–707.
- Frank, N. (2024), Washington Post, Virginia Tech collaborate on AI news search tool. *Virginia Tech* Sept. 9, <https://news.vt.edu/articles/2024/09/washington-post-virginia-tech-ai-news-climate-answers.html>.

- Frohm, J.; Lindström, V.; Winroth, M.; Stahre, J. (2008), Levels of Automation in Manufacturing. *Ergon. Int. J. Ergon. Hum. Factors*, 30, 181–207.
- Funer, F. (2022), Accuracy and Interpretability: Struggling with the Epistemic Foundations of Machine Learning-Generated Medical Information and Their Practical Implications for the Doctor-Patient Relationship. *Philosophy & Technology* (2022) 35: 5, <https://doi.org/10.1007/s13347-022-00505-7>.
- Gerlach, A., Eisele, J. (2022), Künstliche Intelligenz für akustische Sensorsysteme. Tagungsband DAGA – Jahrestagung für Akustik 2022, Stuttgart, S. 62-65.
https://pub.dega-akustik.de/DAGA_2022/data/articles/000012.pdf.
- Giering, O. (2021), Künstliche Intelligenz und Arbeit: Betrachtungen zwischen Prognose und betrieblicher Realität. *Zeitschrift für Arbeitswissenschaft* 76: 50-64.
- Goddard, K.; Roudsari, A.; Wyatt, J.C. (2012), Automation bias: A systematic review of frequency, effect mediators, and mitigators. *J. Am. Med. Inform. Assoc.*, 19, 121–127.
- Goddard, K.; Roudsari, A.; Wyatt, J.C. (2014), Automation bias: Empirical results assessing influencing factors. *Int. J. Med Inform.*, 83, 368–375.
- Goertzel, B. (2015), Are There Deep Reasons Underlying the Pathologies of Today’s Deep Learning Algorithms?. In: Bieger, J., Goertzel, B., Potapov, A. (eds) *Artificial General Intelligence. AGI 2015. Lecture Notes in Computer Science()*, vol 9205. Springer, Cham.
https://doi.org/10.1007/978-3-319-21365-1_8.
- Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning (Adaptive Computation and Machine Learning)*; MIT Press: Cambridge, MA, USA, 2016; ISBN 978-0262035613.
- Große, P. (2024), Prompt-Tipp: KI am Limit: Analyse vieler Dokumente. *AI-Journalist*, 20.06.,
<https://www.aijournalist.de/prompt-tipp-ki-am-limit-analyse-vieler-dokumente>.
- Grunwald, A. (2008), Akzeptanz und Akzeptabilität technikbedingter Risiken. In: Grunwald, A.: *Technik und Politikberatung*. Frankfurt a. M.:Surkamp, S. 339–367.
- Hallensleben, Sebastian et al. (2020), From principles to practice. An interdisciplinary framework to operationalise AI ethics. Gütersloh: Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WKIO_2020_final.pdf.
- Hamet, P. und Tremblay, J., 2017, Artificial intelligence in medicine, *Metabolism* 69, S36-S40.
- Haoyuan, M. (2024), Marketing Strategy of Open AI. *Advances in Economics, Management and Political Sciences*. 73 (1), <https://doi.org/10.54254/2754-1169/73/20231500>.
- Harding, W., Kloster, M. (2024), Coding on Copilot, 2023 Data Shows Downward Pressure on Code Quality. January 16, <https://gitclear-public.s3.us-west-2.amazonaws.com/Coding-on-Copilot-2024-Developer-Research.pdf>.
- Hasija, Abhinav & Esper, Terry. (2022), In artificial intelligence (AI) we trust: A qualitative investigation of AI technology acceptance. *Journal of Business Logistics*, 43. DOI:10.1111/jbl.12301.
- Hiltawsky, K. et al. (2024), KI für bessere Abläufe in Medizin und Pflege. Anwendungen und Potenziale in organisatorischen Prozessen. Whitepaper aus der Plattform Lernende Systeme, München. DOI: https://doi.org/10.48669/pls_2024-3.
- HLEG—High-Level Expert Group on Artificial Intelligence (2019), *Ethics Guidelines for Trustworthy AI*. European Commission.
<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
- Hofstadter, D.R. (1996), *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*; Basic Books: New York, NY, USA.
- Hohenwalde, C.E., Wahl, M., Lehmkuhl, L. (2024), Kann ChatGPT komplexe wissenschaftsjournalistische Texte verständlich machen? Fehler bei der automatischen Übersetzung mit ChatGPT in einfache Sprache. Karlsruhe Institut für Technologie. DOI: 10.5445/IR/1000174282
<https://publikationen.bibliothek.kit.edu/1000174282>.
- Hosbach, W. (2024), Schlechte Code-Qualität durch die KI-Assistenten GitHub Copilot und ChatGPT Heise online, 26.01., <https://www.heise.de/news/Schlechte-Code-Qualitaet-durch-die-KI-Assistenten-GitHub-Copilot-und-ChatGPT-9609271.html>.

- Hube, G. (2005), Beitrag zur Beschreibung und Analyse von Wissensarbeit. Heimsheim: Jost Jetter.
- IAB – Institut für Arbeitsmarkt- und Berufsforschung (2021). Job Futuromat <https://job-futuromat.iab.de/>;
Zuletzt abgerufen: 16.03.2022.
- ifb (2024), Die unsichtbaren Helden der Künstlichen Intelligenz: Wie Klickarbeiter die Grundlage für KI legen und dabei oft vergessen werden. 06.02., <https://www.betriebsrat.de/news/kuenstliche-intelligenz/die-unsichtbaren-helden-der-kuenstlichen-intelligenz-3073557>.
- IFES – Institut für empirische Sozialforschung (2023), Digitalisierung: 74 % befürchten mehr Überwachung am Arbeitsplatz https://www.arbeiterkammer.at/interessenvertretung/wirtschaft/konsument/Presseunterlage_20230202.pdf.
- IOZ (2024), (Update August 2024) M365 Copilot: Erfahrungsbericht nach 6 Wochen Praxiseinsatz. 23.09., <https://www.ioz.ch/blog/m365-copilot-erfahrungsbericht-nach-6-wochen-praxiseinsatz-productivity-news-vom-01-03-2024/>.
- ISO – International Standards Organization (2022) What is artificial intelligence (AI)? ISO/IEC 22989:2022 Information technology – Artificial intelligence – Artificial intelligence concepts and terminology <https://www.iso.org/artificial-intelligence/what-is-ai>.
- Janssen, J.-K. (2024), Microsofts Copilot für Office ist eine Frechheit. Heise 16.02., <https://www.heise.de/news/Microsofts-Copilot-fuer-Office-ist-eine-Frechheit-c-t-3003-9631015.html>.
- Joos, T. (2024), Microsoft Recall schon jetzt geknackt. Security Insider 28.06., <https://www.security-insider.de/microsoft-recall-datenschutz-dilemma-um-ki-funktion-windows-a-42aea86c03ae979b0204554d1e55b8c4/>.
- Kelter, J., Rief, S., Bauer, W., Haner, U.-E. (2009), Office 21 Studie Information Work: Über die Potenziale von Informations- und Kommunikationstechnologien bei Büro- und Wissensarbeit. Dieter Spath (Hrsg.), Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO. https://office21.de/wp-content/uploads/2017/10/Fraunhofer-IAO-Studie_Information_Work2009.pdf.
- Kemper, J. (2024), KI-Modelle verstehen zwar Fotos, können aber die Uhr nicht lesen. 28. Juli, the decoder, https://the-decoder.de/ki-modelle-verstehen-zwar-fotos-koennen-aber-die-uhr-nicht-lesen/?utm_source=pocket-newtab-de-de.
- Kornwachs, K. (2023), KI und die Disruption der Arbeit. Carl Hanser Verlag: München.
- Krause-Pilatus, A., Rinne, U., Schneider, H. (2019), Arbeitszufriedenheit in der modernen Arbeitswelt. IZA Standpunkte Nr. 94. IZA – Institute of Labor Economics, <https://docs.iza.org/sp94.pdf>.
- Krawarik, V., Schell, K., Ertelthaler, V., Thallinger, G. und Bailer, W., (2021), *AI and the Austrian Media Sector: Mapping the Landscape, Setting a Course*. Wien: Bundesministerium für Klimaschutz, U., Energie, Mobilität, Innovation und Technologie.
- Krempel, S. (2023), Schadenersatz droht: Datenschützer mahnt Aus für Microsoft 365 an Schulen an. 11.02., <https://www.heise.de/news/Schadenersatz-droht-Datenschuetzer-mahnt-Aus-fuer-Microsoft-365-an-Schulen-an-7493247.html>.
- Kretschmer, C. (2023), Wie Klickarbeiter in Kenia ausgebeutet werden. Tagesschau online 24.08., <https://www.tagesschau.de/wissen/technologie/ki-klickarbeiter-trainingsdaten-100.html>.
- Lane, M., Williams, M., Broecke, S. (2023), The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers. OECD Social, Employment and Migration Working Papers No. 288 <https://dx.doi.org/10.1787/ea0a0fe1-en>.
- LeCun, Y.; Benigo, Y.; Hinton, G. (2015), Deep Learning. Nature, 521, 436–444.
- Leible, S., Gücük, G.-L., Simic, D., von Brackel-Schmidt, C., Lewandowski, T. (2024), Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft. HMD Praxis der Wirtschaftsinformatik (2024) 61:344–370. <https://doi.org/10.1365/s40702-024-01050-x>.
- Li, M., Zhang, S., Liu, Y., Chen, K. (2024), NeedleBench: Can LLMs Do Retrieval and Reasoning in 1 Million Context Window? <https://arxiv.org/pdf/2407.11963>.

- Liu, N.F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., Liang, P. (2023), Lost in the Middle: How Language Models Use Long Contexts, <https://arxiv.org/pdf/2307.03172> <https://the-decoder.de/grosse-sprachmodelle-und-das-phaenomen-der-vergessenen-mittel/>.
- Lohmann, S. (2023), Photoshop-KI ausprobiert: Einfach unglaublich, aber nicht perfekt! Macwelt, 30.06., <https://www.macwelt.de/article/1969634/photoshop-ki-test.html>.
- Long, D., Magerko, B. (2020), What is AI literacy? Competencies and design considerations Proceedings of the 2020 CHI conference on human factors in computing systems, pp. 1-16.
- Looi, J., & Kahlor, L. A. (2024). Artificial Intelligence in Influencer Marketing: A Mixed-Method Comparison of Human and Virtual Influencers on Instagram. *Journal of Interactive Advertising*, 24(2), 107–126. <https://doi.org/10.1080/15252019.2024.2313721>.
- Lyell, D.; Coiera, E. (2016), Automation bias and verification complexity: A systematic review. *J. Am. Med. Inform. Assoc.*, 24, 424–431.
- Maier, F., Hoffmann, C. (2024), Microsoft Copilot Pro im Test. CIO Magazin 06.08., <https://www.cio.de/a/microsoft-copilot-pro-im-test,3698431>.
- Malik, P., Pathania, M. und Rathaur, V. K., 2019, Overview of artificial intelligence in medicine, *Journal of family medicine and primary care* 8(7), 2328.
- McCarthy, J. (2007), Basic Questions, What Is Artificial Intelligence? Stanford University: Stanford, CA, USA, <http://www-formal.stanford.edu/jmc/whatisai/>.
- McCarthy, J., Minsky, M., Rochester, N. and Shannon, C. (1955), A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, <https://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.
- McDermott, D. (1981), Artificial Intelligence meets natural stupidity. In *Mind Design. Philosophy, Psychology, Artificial Intelligence*; Haugheand, J., Ed.; MIT Press: Cambridge, MA, USA; London, UK.
- Mey, S. (2024), Blamage um Chatbot des AMS wirft weitere Fragen auf. *Der Standard online*, 19.01, <https://www.derstandard.at/story/3000000203652/chatbot-ams>.
- Mirbabaie, M., Brünker, F., Möllmann Frick, N.R.J. et al. (2022), The rise of artificial intelligence – understanding the AI identity threat at the workplace. *Electron Markets* 32, 73–99 <https://doi.org/10.1007/s12525-021-00496-x>.
- MIT – Massachusetts Institute of Technology (2024), Detect DeepFakes: How to counteract misinformation created by AI. <https://www.media.mit.edu/projects/detect-fakes/overview/>.
- Mitchell, (2006), T.M. *The Discipline of Machine Learning*. <http://www.cs.cmu.edu/~tom/pubs/MachineLearningTR.pdf>.
- Newman, N. (2024), Journalism, media, and technology trends and predictions 2024 <https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2024>.
- Ng, D., T. K., Leung, J. K. L., Chu, S., K., W., Qiao, M., S. (2021), Conceptualizing AI literacy: An exploratory review, *Computers and Education: Artificial Intelligence*, Volume 2, 100041, <https://doi.org/10.1016/j.caeai.2021.100041>.
- Nickel, O. (2024a), Was Microsoft Copilot in Unternehmen kann – und was nicht. *Golem* 17. April, <https://www.golem.de/news/microsofts-ki-assistent-was-microsoft-copilot-im-unternehmen-kann-und-was-nicht-2404-184229.html>.
- Nickel, O. (2024b), Adobe startet Testphase für Videoclip-KI. *Golem*, 15. Oktober, <https://www.golem.de/news/firefly-adobe-startet-testphase-fuer-videoclip-ki-2410-189844.html>.
- Noyb (2024), Microsoft verletzt Datenschutz von Kindern – und schiebt Schuld auf Schulen, 04.06, <https://noyb.eu/de/microsoft-violates-childrens-privacy-blames-your-local-school>.
- NYT – New York Times (2024), Introducing the AI Initiatives Team, May 9, <https://www.nytc.com/press/introducing-the-a-i-initiatives-team/>.
- OECD (2018). Automation, skills, use and training. OECD social, employment and migration WP No. 202.
- OECD (2023), The Impact of AI on the Workplace: Evidence from OECD Case Studies of AI Implementation. OECD social, employment and migration working papers No. 288. JT03533166 <https://one.oecd.org/document/DELSA/ELSA/WD/SEM%282023%297/en/pdf>.

- ORF (2024), ORF gewinnt mit „AiDitor“ europäischen Technologie- und Innovations-Award. ORF Public Value 07.09., https://zukunft.orf.at/show_content.php?sid=151&blog_id=363&blog_mode=single.
- Osoba, O.; Welser, (2017), I.V.W. An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence; RAND Corporation.
- Ozkaya, I. (2023), The Next Frontier in Software Development: AI-Augmented Software Development Processes. *IEEE Software* (40)4, pp. 4-9,, doi: 10.1109/MS.2023.3278056.
- Peer, M. (2023), Wie eine Million Clickworker Googles KI für Niedriglöhne trainieren. Handelsblatt online, 30.06., <https://www.handelsblatt.com/technik/it-internet/appen-wie-eine-million-clickworker-googles-ki-fuer-niedrigloehne-trainieren/29225530.html>.
- Pfannstiel, M.A. (2022), Künstliche Intelligenz im Gesundheitswesen: Entwicklungen, Beispiele und Perspektiven. Springer Gabler: Wiesbaden.
- Prem, E., Ruhland, S. (2019), Artificial Intelligence Potenzial Österreich: Zahlen, Daten, Fakten. Eine Annäherung auf Basis wirtschaftsstatistischer Analysen. Bundesministerium für Verkehr, Innovation und Technologie. https://www.kmuforschung.ac.at/wp-content/uploads/2019/03/ai_potenzial_oesterreich.pdf.
- Proshofsky, A. (2024), Vorurteile und zweifelhafte Umsetzung: AMS-KI-Chatbot trifft auf Spott und Hohn. Der Standard online, 04.01., <https://www.derstandard.at/story/3000000201774/vorurteile-und-zweifelhafte-umsetzung-der-ams-ki-chatbot-trifft-auf-spott-und-hohn>.
- Raisch, S. Krakowski, S. (2021), Artificial Intelligence and Management: The Automation-Augmentation Paradox. *Academy of Management Review* 46(1), 192–210. <https://doi.org/10.5465/amr.2018.0072>.
- Reuters (2024), <https://www.thomsonreuters.com/en/artificial-intelligence/research.html>.
- Rosbach, O. (2020), Office 365 in der Schule: Grobe Verletzungen datenschutzrechtlicher Vorschriften. Netzpolitik 03.11. <https://netzpolitik.org/2020/office-365-in-der-schule-grobe-verletzungen-datenschutzrechtlicher-vorschriften/>.
- Russell, S., Norvig, P. (2010), *Artificial Intelligence: A Modern Approach*. 3rd Edition. Pearson: New Jersey.
- Schell, K. (2022), *Journalistische Textautomatisierung: Status, Potenziale, Limitationen*. Wien: Austria Presse Agentur.
- Schell, K. (2023), ChatGPT: Die KI, die schreiben kann. Aber kann sie auch Journalismus? ChatGPT: Die KI, die schreiben kann. Aber kann sie auch Journalismus? APA Blog, <https://apa.at/blog/chatgpt-im-journalismus/>
- Schörpf, P., Schönauer, A., Flecker, J. (2018), Entwicklungstrends digitaler Arbeit. Studie im Auftrag der Kammer für Arbeiter und Angestellte Wien.
- Shah, P. (2023), *AI and the future of education: Teaching in the age of artificial intelligence*. Jossey-Bass: New Jersey.
- Springer (2023), Axel Springer and OpenAI partner to deepen beneficial use of AI in journalism. <https://www.axelspringer.com/en/ax-press-release/axel-springer-and-openai-partner-to-deepen-beneficial-use-of-ai-in-journalism>.
- Standard (2024a), Microsoft stellt seinen KI-Copilot nun auch in der breiten Masse zur Verfügung. Der Standard online, 16.01., <https://www.derstandard.at/story/3000000203298/microsoft-stellt-seinen-ki-copilot-nun-auch-in-der-breiten-masse-zur-verfuegung>.
- Standard (2024b), Microsoft stellt Copilot Plus PC vor und will sämtliche Inhalte für die KI speichern. Der Standard online 21.05., <https://www.derstandard.at/story/3000000220842/microsoft-stellt-copilot-plus-pc-vor-und-will-saemtliche-inhalte-fuer-die-ki-speichern>.
- Statistik Austria (2023), 11 % der österreichischen Unternehmen nutzen künstliche Intelligenz <https://www.statistik.at/fileadmin/announcement/2023/10/20231017IKTU2023.pdf>.
- Statistik Austria (2024), Nutzung von künstlicher Intelligenz in Unternehmen innerhalb eines Jahres fast verdoppelt <https://www.statistik.at/fileadmin/announcement/2024/10/20241016IKTU2024.pdf>.
- Stowasser, S. (2023), *Künstliche Intelligenz (KI) und Arbeit: Leitfaden zur soziotechnischen Gestaltung von KI-Systemen*. Ifaa-Edition, Institut für angewandte Arbeitswissenschaft, Springer Vieweg.

- Strauß, S. (2018), From big data to deep learning. A leap towards strong AI or 'intelligentia obscura'? In: Big Data and Cognitive Computing 2 (3), 1–19. <https://www.mdpi.com/2504-2289/2/3/16>.
- Strauß, S. (2021a), „Don't let me be misunderstood“: Critical AI literacy for the constructive use of AI technology. TATuP – Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis 30(3), 44-49. <https://doi.org/10.14512/tatup.30.3.44>.
- Strauß, S. (2021b), Deep automation bias. How to tackle a wicked problem of AI? In: Big Data and Cognitive Computing 5(2), 1–14. <https://doi.org/10.3390/bdcc5020018>.
- Strauß, S., Bettin, S. (2023), Digitalisierung, Vulnerabilität und (kritische) gesellschaftliche Infrastrukturen: Entwicklungsstand, Trends und zentrale Herausforderungen. ÖAW/ITA: Wien <https://epub.oew.ac.at/ita/ita-projektberichte/ITA-pb-2023-01.pdf>.
- Strauß, S., Udrea, T. (2024), Menschzentrierte KI-Automatisierung: Zwischen Automation Bias und Critical AI Literacy. Publikation zur Wissenschaftskonferenz der Arbeiterkammer Vorarlberg im November 2023. Technikfolgenabschätzung aus Arbeitnehmer:innenperspektive, S. 226-249.
- Sutton, R.S., Barto, A.G. (2015), Reinforcement Learning: An Introduction. Second Edition, MIT Press: Cambridge, Massachusetts. <https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf>.
- Taeibi, B. (2017). Bridging the gap between social acceptance and ethical acceptability. Risk Analysis 37(10):1817-1827.
- Tschandl, P., Rinner, C., Apalla, Z. et al. (2020), Human-computer collaboration for skin cancer recognition. Nature Medicine 26(8): 1229-1234.
- Tschoepe, C., Duckhorn, F., Woff, M. (2017), Akustische Mustererkennung: Qualitätskontrolle – Vorausschauende Instandhaltung – Zustandsüberwachung. Fraunhofer IKTS, https://www.researchgate.net/publication/318946441_Akustische_Mustererkennung_Qualitätskontrolle_-_Vorausschauende_Instandhaltung_-_Zustandsüberwachung.
- Turing, A. (1950), Computing Machinery and Intelligence. Mind, LIX, 433–466.
- UNESCO (2022), Recommendation on the Ethics of Artificial Intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>.
- UNESCO (2024), UNESCO survey: Less than 10 % of schools and universities have formal guidance on AI. Sept. 18, <https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai>.
- VDE (2024), Zulassung von KI-basierten Medizinprodukten in Europa. VDE, 23.07., <https://www.vde.com/topics-de/health/beratung/zulassung-von-ki-basierten-medizinprodukten-in-europa>.
- Verdegem, P. (2021), Introduction: Why We Need Critical Perspectives on AI. In: Verdegem, P. (ed.) AI for Everyone? Critical Perspectives. Pp. 1–18. London: University of Westminster Press. DOI: <https://doi.org/10.16997/book55.a>.
- von Garrel, J., Mayer, J., Mühlfeld, M. (2023), Künstliche Intelligenz im Studium – Eine quantitative Befragung von Studierenden zur Nutzung von ChatGPT & Co. Online unter: [doi: 10.48444/h_docs-pub-395](https://doi.org/10.48444/h_docs-pub-395).
- von Richthofen, G., Köhne, S. Send, H. (2023), KI in der Wissensarbeit. Handlungsfelder und Ansätze für eine beschäftigtenorientierte Gestaltung. *HIIG Impact Publication Series* <https://doi.org/10.5281/zenodo.7541307>.
- Warren, T. (2024), Microsoft says its Recall uninstall option in Windows 11 is just a bug. The Verge 02.09., <https://www.theverge.com/2024/9/2/24233992/microsoft-recall-windows-11-uninstall-feature-bug>.
- Weber, R.O., Johs, A.J., Goel, P., Silva, J.M. (2024), XAI is in trouble. AI Magazine:1-17, DOI: 10.1002/aaai.12184.
- Weiß, E.M. (2024), Österreichische Arbeitsagentur veröffentlicht fragwürdigen KI-Chatbot. Heise online, 05.01., <https://www.heise.de/news/Oesterreichische-Arbeitsagentur-veroeffentlicht-fragwuerdigen-KI-Chatbot-9588098.html>.
- Weizenbaum, J. (1976), Computer Power and Human Reason: From Judgement to Calculation; W. H. Freeman: San Francisco, CA, USA, 1976.

- Wilhelm, K. (2024), KI und Film – was hat sich seit dem Hollywood-Streik getan? NDR online, 11.06., <https://www.ndr.de/kultur/film/KI-und-Film-was-hat-sich-seit-dem-Hollywood-Streik-getan,kiundfilm100.html>.
- Wilkens, A. (2023), Bild-Zeitung: KI soll das Layout machen. Heise 20.06. <https://www.heise.de/news/KI-Bild-Zeitung-setzt-auf-Kuenstliche-Intelligenz-9192608.html>.
- Wilkens, U. (2020), Artificial intelligence in the workplace – A double-edged sword. *International Journal of Information and Learning Technology*, 37(5): 253-265.
- Willke, H. (1998), Organisierte Wissensarbeit. *Zeitschrift für Soziologie*, Jg. 27, Heft 3, Juni 1998, S. 161-177. <https://ams-forschungsnetzwerk.at/downloadpub/organisierte%20wissensarbeit.pdf>.
- Wu, T-H., Biamby, G., Quenum, J., Gupta, R., Gonzalez, J.E., Darrell, T., Chan, D.M. (2024), Visual Haystacks: Answering Harder Questions About Sets of Images. <https://arxiv.org/abs/2407.13766>.
- Yi; Y. (2021), Establishing the concept of AI literacy: Focusing on competence and purpose. *European Journal of Bioethics* 12(2):353-368. DOI: [10.21860/j.12.2.8](https://doi.org/10.21860/j.12.2.8).
- Zhang et al. (2024), Multimodal Self-Instruct: Synthetic Abstract Image and Visual Reasoning Instruction Using Language Model. <https://arxiv.org/abs/2407.07053>.

Alle URLs zuletzt aufgerufen am 30.10.2024.



ÖAW



WWW.OEAW.AC.AT