

KARL GF5 Leitfaden

Whitepaper zu ethischen, rechtlichen und sozialen Aspekten im
Kontext von Künstlicher Intelligenz für Arbeit und Lernen

Publikation: August 2023



Künstliche Intelligenz
für Arbeit und Lernen



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Autor:innen

Sascha Alpers¹, Bettina-Johanna Krings², Welf Schröter³, Christoph Becker¹,
Josef Brücklmayr¹, Ann-Katrin Dreher¹, Philipp Frey², Miriam Klöpper¹, Maria Rill¹,
Marius Take¹, Aline Vugrincic¹, Maria Weinreuter¹

¹ FZI Forschungszentrum Informatik, Haid- und Neu-Straße 10-14, 76131 Karlsruhe, {alpers, christoph.becker, bruecklmayr, dreher, kloeppe, m.rill, take, vugrincic, weinreuter}@fzi.de, Webseite: www.fzi.de

² Karlsruher Institut für Technologie (KIT), Institut für Technikfolgenabschätzung und Systemanalyse (ITAS), Karlstraße 11, 76133 Karlsruhe, {bettina-johanna.krings, philipp.frey}@kit.edu, Webseite: www.itas.kit.edu

³ Forum Soziale Technikgestaltung beim Deutschen Gewerkschaftsbund Bezirk Baden-Württemberg (FST), Leitung und Moderation Welf Schröter, Rietsweg 2, 72116 Mössingen-Talheim, schroeter@talheimer.de, Webseite: www.blog-zukunft-der-arbeit.de

Projektwebseite: www.kompetenzzentrum-karl.de

Hinweis

Der vorliegende Leitfaden ist eine Beta-Fassung (Stand August 2023). Er gibt zentrale Ergebnisse aus dem Gestaltungsfeld 5 des Verbundforschungsprojektes KARL (Künstliche Intelligenz für Arbeit und Lernen in der Region Karlsruhe) wieder. Entsprechend dem Projektplan läuft zum Zeitpunkt der Veröffentlichung die Evaluation in verschiedenen Anwendungsfällen.





Inhalt

1.	Einleitung	3
2.	Definitionen & Abgrenzungen	5
2.1.	Ethische, rechtliche und soziale Aspekte (ELSA)	5
2.2.	Künstliche Intelligenz	6
2.3.	Problem- oder Technikzentrierung?	10
3.	Integration von ELSA in Vorgehensmodelle	12
3.1.	Vorgehensmodell	12
3.2.	Integration von ELSA in Vorgehensmodelle des Software-Engineerings.....	15
4.	Handlungsempfehlungen für Akteur:innen in (neuen) KI-Vorhaben in Forschungsorganisationen und Unternehmen	16
4.1.	Analytische Ebene: Haftungsmodelle im Zusammenhang mit KI	17
4.2.	Analytische Ebene: Organisationsentwicklung	20
4.3.	Analytische Ebene: Persönliche Kompetenzen ausbilden	24
4.4.	Analytische Ebene: Einsatz von KI-Systemen.....	27
4.4.1.	Datenschutzrechtliche Anforderungen	27
4.4.2.	Betriebsverfassungsrechtliche Anforderungen.....	30
5.	Handlungsempfehlungen an die Sozialpartner:innen für erfolgreiche Mitbestimmungsschritte	33
6.	Forschungs- und Gestaltungsdesiderate	43
7.	Quellenverzeichnis.....	46

Die einzelnen (Teil-)Kapitel wurden von unterschiedlichen Autor:innen verfasst. Trotz des Versuchs, einen Konsens über alle Aspekte herzustellen, geben sie doch teils nicht die Meinung aller Autor:innen wieder. Auch bestimmte Begriffe wie etwa "lernende" Technik werden aufgrund des unterschiedlichen Begriffsverständnisses in den beteiligten Disziplinen nicht von allen Autor:innen geteilt.

Die einzelnen Kapitel haben mit ihren unterschiedlichen Perspektiven unterschiedliche Adressat:innen. Nur teilweise konnten Doppelungen zwischen den Kapiteln vermieden werden. Dies war jedoch auf Grund der Erhaltung des guten Leseflusses nicht immer möglich.

1. Einleitung

Die Anwendung von Künstlicher Intelligenz (KI) ist eingebettet in die seit vielen Jahren anhaltende Digitalisierungsdynamik - auch von Produktionsprozessen. Die aktuellen Entwicklungen zeigen, dass große Erwartungen an den künftigen Einsatz von KI in der Arbeitswelt gestellt werden. Diese werden weitreichende Auswirkungen auf die Arbeitsorganisation und Arbeitsprozesse von vielen Beschäftigten haben. Ziel des Kompetenzzentrums KARL (Künstliche Intelligenz für Arbeiten und Lernen in der Region Karlsruhe) ist es, einen Beitrag für die Gestaltung dieser neuen Arbeitsformen zu leisten. Vor diesem Hintergrund ist das Ziel, menschenzentrierte, transparente und lernförderliche KI-unterstützte Arbeits- und Lernsysteme zu konzipieren und in konkreten Praxisanwendungen zu demonstrieren (siehe Abbildung 1, Dach). Die in KARL entwickelten Werkzeuge, Methoden und Handreichungen, wie beispielsweise dieser Leitfaden, sollen Unternehmen in die Lage versetzen, attraktive KI-gestützte Arbeitsumgebungen, die auch auf die Interessen der Betroffenen (z.B. Arbeitnehmende) abgestimmt sind, umzusetzen. Im Rahmen des Verbundforschungsprojektes KARL arbeiten sieben Forschungseinrichtungen, zwei Transfereinrichtungen und elf regionale Unternehmen gefördert durch das Bundesministerium für Bildung und Forschung (BMBF) zusammen. Zusätzlich sind weitere Organisationen assoziiert.

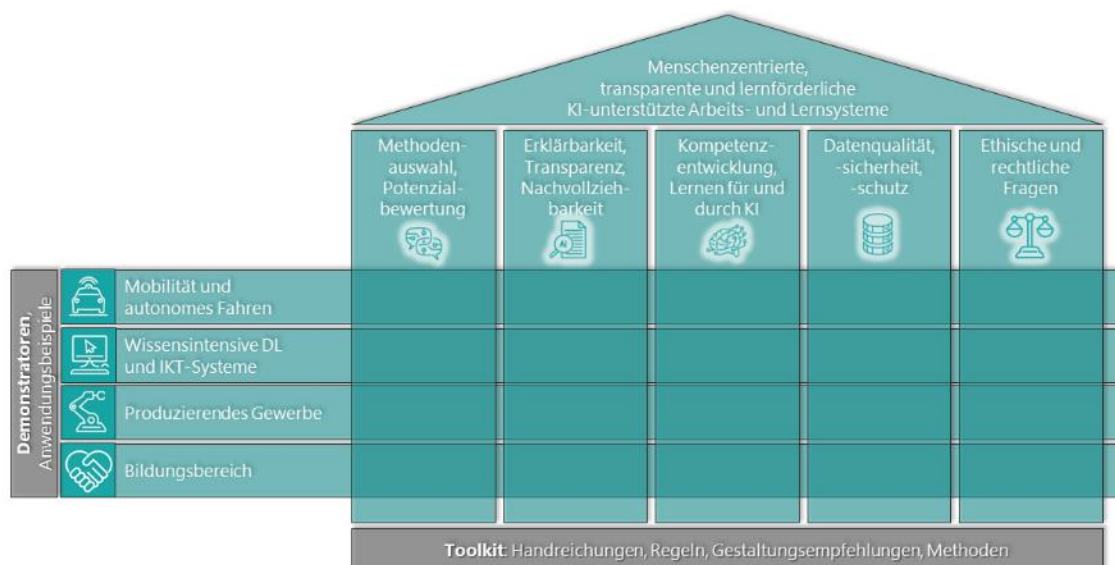


Abbildung 1: Übersichtsgrafik zu KARL. Zusammenwirken von Gestaltungsfeldern und Anwendungsbeispielen [Darstellung des Konsortialführers Hochschule Karlsruhe, Institut für Lernen und Innovation in Netzwerken].

Im Vorhaben KARL werden in fünf Gestaltungsfeldern (in Abbildung 1 als Säulen dargestellt) vorhandene und neue Erkenntnisse so aufbereitet, dass diese im Rahmen des neuen Kompetenzzentrums KARL in die unternehmerische Praxis überführt werden können. Die Handreichungen, Regeln, Gestaltungsempfehlungen und Methoden der fünf Gestaltungsfelder bilden zusammen das *Toolkit*, quasi den Werkzeugkoffer, von KARL.



Das **Gestaltungsfeld 5** hat sich in den letzten Jahren intensiv mit den **rechtlichen, sozialen und ethischen Dimensionen von KI-Anwendungen** in konkreten Arbeitsumgebungen beschäftigt. Diese Dimensionen werden im vorliegenden Leitfaden vorgestellt, exemplifiziert und entlang von Leitfragen für die Praxis vorbereitet. Hierbei werden Schwerpunkte gesetzt, die in der Natur der jeweiligen Dimensionen liegen. Beispielsweise sollen im Rahmen der Nutzung dieses Leitfadens auch Fragestellungen der digitalen Souveränität berührt werden (vgl. Beyerer, Müller-Quade und Reussner, 2018) und – sofern natürliche Personen betroffen sind – sich mit Fragen des Datenschutzes auseinandergesetzt werden. Zu Fragestellungen des Datenschutzrechtes verweisen wir auf die Arbeiten im KARL-Gestaltungsfeld 4 „Umgang mit Daten“ und deren Dokument „Datenschutz bei Künstlicher Intelligenz“ (Bao *et al.*, 2023).

So ist das Ziel dieses Leitfadens, auf Basis der bisherigen Erkenntnisse und Erfahrungen des KARL-Gestaltungsfeldes 5, für ethische, rechtliche und soziale Aspekte zu sensibilisieren und diese gemeinsam zu reflektieren. Trotz der unterschiedlichen Perspektiven und der unterschiedlichen institutionellen Einbindung der sozialen, rechtlichen und ethischen Themen in den Betrieben sollen diese zusammen gedacht und zusammengeführt werden. Die Anerkennung des personenbezogenen Datenschutzes beispielsweise ist in Deutschland auf der rechtlichen Ebene der Betriebsverfassung geschützt. Die Gratwanderungen, die jedoch mit dem Einsatz von KI-Systemen im Bereich des Datenschutzes entstehen können, erfordern frühzeitige soziale und ethische Reflexionen in der Gestaltung neuartiger Mensch-Maschine-Interaktionen in KI- gestützten Arbeitsumgebungen. Hierbei werden Themen berührt, die sich intensiv mit der Zukunft „guter“ Arbeit befassen. Die Autor:innen wünschen sich, dass der Leitfaden vor diesem Anspruch den Leser:innen wichtige Impulse zur Umsetzung „humanverträglicher“ KI-Systeme gibt.

Der vorliegende Leitfaden gliedert sich in mehrere Abschnitte. Der erste Abschnitt widmet sich den "Definitionen und Abgrenzungen" der KI in verschiedenen Ausprägungen. Dabei wird auf unterschiedliche Formen von KI eingegangen, um ein umfassendes Verständnis für die Bandbreite der Technologie zu vermitteln. Zudem wird ein Überblick über die sogenannten ELSA (Ethische, rechtliche und soziale Aspekte) gegeben. Im Anschluss daran wird ein Vorgehensmodell für die Softwareentwicklung vorgestellt, um ein besseres Verständnis von projekt- und praxisrelevanten Abläufen zu ermöglichen. Dabei können entsprechend angepasste Vorgehensmodelle sowohl im betrieblichen Umfeld als auch im wissenschaftlichen Umfeld genutzt werden.

Der Hauptteil des Leitfadens beinhaltet Handlungsempfehlungen, die sich an diverse Akteur:innen richten und entsprechend praxisorientierte Ratschläge für Unternehmen, Sozialpartner:innen, Entwickler:innen und Entscheidungsträger:innen bieten. Diese Empfehlungen basieren auf den zuvor behandelten Inhalten und sollen dazu beitragen, KI-Anwendungen verantwortungsbewusst zu gestalten und ausgewählte ELSA-Kriterien effektiv umzusetzen. Indem der Leitfaden Definitionen, Vorgehensmodelle und Handlungsempfehlungen kombiniert, bietet er eine wertvolle Ressource für die Entwicklung und den Einsatz von KI-basierten Systemen. Schließlich enthält der Leitfaden auch weiterreichende Handlungsempfehlungen für Forschende, Unternehmen und Beschäftigte sowie deren Vertreter:innen (Betriebsräte, Gewerkschaften) in Unternehmen.



2. Definitionen & Abgrenzungen

Dieses Kapitel widmet sich der Erörterung grundlegender Definitionen und Abgrenzungen von Systemen, deren Wirkweise durch Algorithmen getrieben ist oder auf Künstlicher Intelligenz (KI) basiert. So soll ein Fundament für das Verständnis der im Leitfaden behandelten Technologien geschaffen werden. Wir beleuchten die essenziellen Aspekte, die es ermöglichen, KI-Anwendungen einzuordnen und potentiell durch die während der Nutzung entstehenden Risiken navigieren zu können. Dabei wird neben den Eigenschaften der Systeme auch besonderes Augenmerk auf den Unterschied zwischen problemzentriertem und technikzentriertem Vorgehen gelegt.

2.1. Ethische, rechtliche und soziale Aspekte (ELSA)

Die ELSA-Kriterien (engl. *Ethical, Legal, and Social Aspects* – Ethische, rechtliche und soziale Aspekte) stellen ein inzwischen gut eingeführtes und angewandtes Rahmenwerk zur Bewertung wissenschaftlicher und technologischer Entwicklungen dar (Fiedeler und Nentwich, 2009). Neben den rein technischen Aspekten werden im Zuge technologischer Innovationen ethische, rechtliche und soziale Implikationen betont, die bei der Implementierung der Innovationen in soziale Kontexte berücksichtigt werden sollten. Diese Vorgehensweise beruht auf dem Anspruch, das Prinzip der Verantwortung verstärkt in die Technikgestaltung zu integrieren, d.h. potentielle Risiken zu antizipieren und zu vermeiden. Das Konzept der *Responsible Research and Innovation* (RRI) wurde erstmals 2014 als normative Perspektive in das Horizon 2020-Programm der Europäischen Kommission aufgenommen und in einer großen Anzahl von Projekten umgesetzt (Grunwald und Hillerbrand, 2021). Der Ursprungsgedanke von RRI fokussiert zum einen die Vermeidung potenzieller Risiken von technischen Innovationen. Zum anderen betont das Konzept vor allem den Implementierungsprozess, der inklusiv und partizipativ gestaltet werden sollte:

“Responsible Research and Innovation is a transparent, interactive process by which societal actors innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability, and societal desirability of the innovation process and its marketable products, in order to allow a proper embedding of scientific and technological advance in our society” (von Schomberg, 2013).

Dem Ansatz liegt die Vorstellung zugrunde, gesellschaftliche Werte sichtbar zu machen, die die aktuellen Produktions-, Arbeits- und Lebensprozesse prägen, wie beispielsweise das Prinzip der Produktivitätssteigerung auf Kosten (langanhaltender) sozialer oder ökologischer Folgewirkungen. Erst wenn diese Werte und Normen sichtbar, offen und transparent in wissenschaftlichen und öffentlichen Debatten diskutiert werden, können Entscheidungen darüber getroffen werden, *ob* und *welche Normen und Werte* zukünftige technologische Entwicklungen prägen und gestalten *sollen*. Hierbei wird im Rahmen des Konzeptes die Prämisse der „Verantwortung“ (Jonas, 2020) ausgearbeitet, die auf den unterschiedlichen institutionellen Ebenen der Technikgestaltung handlungsanleitend werden soll. Auf der Basis der ELSA-Kriterien werden diese Werte und Normen zum einen sichtbar gemacht, zum anderen in konkreten Handlungsfeldern operationalisiert. Langfristig sollen sie dazu



beitragen, verantwortungsbewusste und nachhaltige Innovationen zu fördern und den gesellschaftlichen Nutzen neuer Technologien (neu) zu definieren, bzw. in Gleichklang mit auszuhandelnden Prämissen des gesellschaftlichen Gemeinwohls zu bringen. So sollen sie einen Leitfaden für eine ausgewogene Abwägung zwischen Fortschritt und den Interessen von Individuen sowie der Gesellschaft als Ganzem bieten.

Tabelle 1: Auszug der strukturellen ELSA-Kriterien in KARL

ELSA-Kriterium	Aspekte und Fragen
Ethik	Gibt es ethische Bedenken hinsichtlich der Technologie? Was ist der Zweck und das Ziel der KI-Anwendung in diesem konkreten Kontext? Wer profitiert von der Anwendung? Wie verändert sich die Mensch-Maschine-Interaktion? Werden die Beschäftigten in diese Veränderungen einbezogen? Auf welchen Normen und Werten basiert die KI-Anwendung im konkreten Anwendungsfall? Wer entscheidet über die Prozesse?
Recht	Erfüllt die Technologie bestehende Gesetze und Vorschriften? Welche rechtlichen Folgen (Risiken) könnten sich aus der Nutzung in den konkreten Handlungsfeldern in KARL ergeben?
Soziale Aspekte	Wie könnte die Technologie bestehende soziale Strukturen, wie die Arbeitsorganisation, beeinflussen? Welche Gruppen könnten benachteiligt werden? Entstehen neue Qualifizierungsbedarfe? Wie können diese abgedeckt werden? Drohen negative Beschäftigungseffekte? Können diese durch eine Betriebs- oder Dienstvereinbarung abgewendet werden? Gibt es transparente Kommunikation über die Technologie und ihre Folgen? Wie können Interessenvertreter und Betroffene mit einbezogen werden?

Über die ELSA-Kriterien hinaus können die vierzig Kriterien des FST zur Gestaltung von künstlicher Intelligenz berücksichtigt werden (siehe: [Schröter, 2023a](#)).

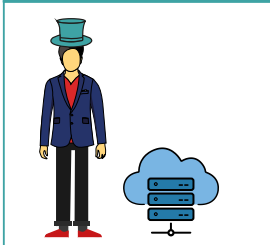
2.2. Künstliche Intelligenz

Auch wenn der Begriff der Künstlichen Intelligenz (KI) gegenwärtig weitverbreitet ist, fehlt eine allgemein akzeptierte Definition. Ursprünglich wurde KI mit dem Ziel formuliert, Aspekte intelligenten Verhaltens so zu beschreiben und zu formalisieren, dass sie durch ein Computersystem simulierbar sind (McCarthy et al., 2006). Problematisch dabei ist, dass auch die Definition von „Intelligenz“ nicht eindeutig ist. Die Frage, ob sich ein System, ein Programm und ein Mensch etc. „intelligent“ verhalten, ist subjektiv und lässt sich daher nicht allgemein abgrenzen (Turing, 1948). Entsprechend haben sich über die Zeit unterschiedliche Richtungen der KI-Forschung entwickelt, die jeweils unterschiedliche Ziele verfolgen (Russell und Norvig, 2016). Beispielsweise zielt eine Sparte darauf ab, Denkprozesse des Menschen zu erforschen und zu simulieren, während eine andere danach strebt, Computersysteme zur Lösung von Aufgaben zu befähigen, zu denen Menschen im Allgemeinen fähig sind, die aber (noch) nicht „klassisch“ in Computersoftware abgebildet werden können (z. B. Sprachverarbeitung, logisches Schließen, bildliche Wahrnehmung).

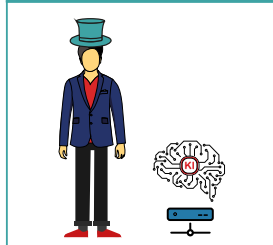
Assistenz- bzw. Unterstützungstechnik

Systeme, die Menschen unterstützen bzw. ihnen assistieren. Entscheidungen verbleiben bei Menschen.

regelbasierte Systeme



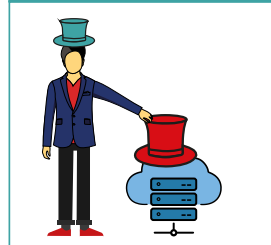
KI-Systeme



Delegationstechnik

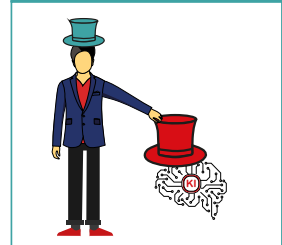
Systeme, die von Menschen bevollmächtigt werden, selbst Entscheidungen zu treffen.

regelbasierte Systeme



Im Falle von regelbasierten Systemen wurden – genau betrachtet – die Entscheidungen von den Regel-erstellern getroffen.

KI-Systeme



Zu unterscheiden sind KI-Systeme, die sich während ihres Einsatzes verändern („dazulernen“), von Systemen, die sich nach Beginn des Einsatzes nicht mehr verändern. Außerdem sind Systeme, deren Entscheidungen (durch sie selbst) erklärt werden können, von nicht erklärbaren bzw. selbsterklärenden KI-Systemen zu unterscheiden.

Abbildung 2: Unterscheidung von Unterstützungs- und Assistenztechnik sowie Delegationstechnik

Die neue Softwarewelt, welche die Arbeitsplätze immer mehr erreicht, lässt sich zunächst in zwei große Richtungen gliedern (vgl. Abbildung 2):

- Software, die den Menschen unterstützt und die Entscheidungen beim Menschen belässt (sogenannte Assistenz- und Unterstützungstechnik)
- Software, die selbst Entscheidungen zu treffen hat beziehungsweise trifft (sogenannte Delegationstechnik)

Geht es um Unterstützungstechnik, d. h. um „Assistenztechnik“, oder um Technik, auf die eine Entscheidungsvollmacht übertragen und delegiert werden kann, also um „Delegationstechnik“? In der Fähigkeit, zwischen „Assistenztechnik“ und „Delegationstechnik“ unterscheiden zu können, liegt ein wesentlicher zukünftiger Erfolgsfaktor im Dialog der Sozialpartner:innen. Dieser Aspekt wird zu einer Schlüsselfrage der Aushandlungsprozesse vor Ort. Nachgelagert ist die Frage, wie diese Systeme realisiert sind. Hier gibt es ebenfalls zwei Gruppen:

- regelbasierte Systeme (auch algorithmische Systeme genannt)
- KI-Systeme (die dann noch weiter unterschieden werden können, insbesondere bezüglich des „Lern“-Zeitpunkts und der Erklärbarkeit)

Zur Bedeutung dieser Unterscheidungsfähigkeit hat Welf Schröter, Leiter und Moderator des FST, für das KARL-Vorhaben zehn Videos als Kurzimpulse gestaltet (Schröter, 2023b).

Die im Leitfaden angesprochenen Formen algorithmischer und KI-basierter Systeme sowie potenziell durch den Umgang mit diesen Systemen entstehende Risiken werden im Folgenden in den beiden Dimensionen „Technologie“ und „Auswirkungen“ kurz skizziert.



Dimension 1: Technologie

Es gibt unterschiedliche Ausbaustufen der Systeme, die im Diskurs um KI Beachtung finden. Insgesamt repräsentieren diese Technologien unterschiedliche Ansätze im Bereich der künstlichen Intelligenz, die von statischen, regelbasierten Systemen bis hin zu dynamischen, lernenden Systemen reichen. Jede Art hat ihre eigenen Vor- und Nachteile und ist für verschiedene Anwendungsfälle geeignet.

Regelbasierte Systeme: Ein regelbasiertes System, auch „Expertensystem“, ist ein System, das auf einer Sammlung von vordefinierten Regeln basiert (Jiang, Y. *et al.*, 2022). Die Idee hinter solchen KI-Systemen ist es, Domänenwissen systematisch in Form von Fakten und "Wenn-Dann-Regeln" formalisiert abzubilden. Die Regeln werden von menschlichen Expert:innen erstellt und beschreiben, wie das System auf verschiedene Szenarien reagieren soll. Wenn das System eine Eingabe erhält, werden die Regeln fortlaufend oder parallel angewendet, um eine passende Antwort oder Aktion zu erzeugen (Russell und Norvig, 2016). Die Funktionsweisen der Systeme sind für Nutzer:innen leichter nachvollziehbar, da die Regeln von Menschen erstellt wurden. Zudem können die Regeln jederzeit nachjustiert werden. Dadurch entsteht ein höherer manueller Aufwand, der häufig Expert:innen-Wissen voraussetzt. Es kann darüber hinaus schwierig sein, vorab wichtige Szenarien und Regeln zu erfassen und zu definieren.

Lernende Systeme: Beim Ansatz des maschinellen Lernens (ML) werden keine Regeln manuell in das System eingepflegt. Stattdessen analysieren hier „Lernalgorithmen“ große Mengen an Daten und abstrahieren sie zu einem Modell (Russell und Norvig, 2016). Das Modell bildet Muster und Regelmäßigkeiten ab, die aus den Daten extrahiert wurden und kann vom KI-System zur Lösung neuer Probleme verwendet werden. Es gibt im Bereich des ML verschiedene Arten von Modellen, die jeweils unterschiedlich gut vom Menschen nachvollziehbar sind (Molnar, 2022). Beispielsweise gibt es Systeme, die aus den Daten formale Regeln extrahieren. Diese können, je nach Komplexität, auch vom Menschen nachvollzogen werden. Andere Modelle hingegen basieren darauf, die gelernten Muster in Form von komplexen mathematischen Berechnungen abzubilden - ein typisches Beispiel hierfür sind sogenannte "künstliche neuronale Netze". Diese Systeme haben sowohl Vor- als auch Nachteile. So verfügen sie über eine hohe Anpassungsfähigkeit und können auch komplexe Muster in großen Datenmengen erkennen. Allerdings werden sie auf Grund von möglicher Intransparenz und komplexer innerer Strukturen häufig als Black Box bezeichnet, die für Menschen nicht mehr nachvollziehbar ist (Loi, 2021). Zudem sind die Ergebnisse der Systeme stark von der Qualität der ursprünglichen Trainingsdaten abhängig (Tursunbayeva *et al.*, 2022). Insgesamt können beide Ansätze ihre Stärken und Schwächen haben und die Wahl zwischen einem regelbasierten und einem selbstlernenden System hängt von den Anforderungen der spezifischen Anwendung ab. Oft werden auch hybride Ansätze verwendet, bei denen regelbasierte und selbstlernende Elemente kombiniert werden, um die Vorteile beider Welten zu nutzen (Russell und Norvig, 2016).

Dabei können Systeme über sogenannte Maturity Level verfügen, mit denen die potentielle analytische Kapazität der Systeme bezeichnet wird (Giermindl *et al.*, 2022). Es wird unterschieden zwischen deskriptiven, prädiktiven und präskriptiven Systemen (Giermindel *et al.*, 2023; Klöpper, 2023). Deskriptive Systeme nutzen einfache statistische Methoden wie



Korrelationsanalyse oder Mittelwerte um Daten auszuwerten und einen Überblick über aktuelles oder vergangenes Geschehen zu liefern. Prädiktive Systeme nutzen fortgeschrittene statistische Methoden um Wahrscheinlichkeiten für zukünftig eintretende Ereignisse zu berechnen. Präskriptive Systeme gehen einen Schritt weiter und nutzen maschinelles Lernen, um Handlungsempfehlungen zu geben, die einen Einfluss auf die Wahrscheinlichkeit für das Eintreten von Ereignissen haben oder die Ergebnisse beeinflussen. Aktuell findet sich in der wissenschaftlichen Literatur auch vermehrt ein viertes Level, sogenannte autonome Systeme (Klöpfer, 2023). Dieses vierte Level beschreibt Systeme, die eigenmächtig in der Lage sind, die von ihnen vorgeschlagenen Handlungen auch umzusetzen (Giermindl *et al.*, 2022). Diese Form der Systeme in den Bereichen Arbeit und Lernen - generell im Umgang mit Personalentscheidungen oder personenbezogenen Daten - ist weitestgehend nicht mit der deutschen Rechtsprechung zu vereinbaren.

Dimension 2: Auswirkungen

Die Systeme finden zunehmenden Einsatz als Instrumente zur Unterstützung menschlichen Entscheidens. Dieses Paradigma spiegelt das Prinzip menschlicher Letztentscheidungen wider, bei dem KI-Systeme nicht darauf abzielen, menschliche Urteile und Entscheidungen zu ersetzen, sondern vielmehr dazu dienen, umfassendere Informationen und Datenanalysen bereitzustellen. Diese sollen die menschliche Expertise ergänzen und erweitern (Jarrahi *et al.*, 2021). In diesem Zusammenhang werden KI-Systeme als leistungsfähige Assistenz porträtiert, die in der Lage ist, Muster zu erkennen, komplexe Daten zu verarbeiten und relevante Einblicke zu liefern, die wiederum menschlichen Entscheidungsträger:innen helfen sollen, fundierte und informierte Schlussfolgerungen zu ziehen. Diese Schlussfolgerungen sollen darüber hinaus frei von menschlichen Vorurteilen (*Biases*) sein und somit dazu beitragen, Diskriminierung zu vermeiden (Adams-Prassl, 2022). Die grundlegende Idee besteht darin, die Synergie zwischen menschlicher Intelligenz und maschinellem Lernen zu nutzen, um bessere Entscheidungsgrundlagen zu schaffen und letztendlich effektivere Ergebnisse in verschiedenen Bereichen zu erzielen.

Neben allen positiven Aspekten, die eine Nutzung der Systeme verspricht, birgt diese jedoch auch eine Vielzahl an Risiken. Ein deutliches und in der Forschung mittlerweile viel diskutiertes Risiko betrifft die potenzielle Verstärkung von Vorurteilen und Diskriminierung durch KI-Systeme (Marjanovic, Cecez-Kecmanovic und Vidgen, 2021). Da diese Systeme auf historischen Daten trainiert werden, können sie unbewusst vorhandene Vorurteile aufgreifen und reproduzieren, was zu ungerechten oder diskriminierenden Entscheidungen führen kann. Diese zu belegen ist für Betroffene häufig schwierig, ungerechte Entscheidungen können daher möglicherweise nicht angefochten werden. Auch die Frage der Verantwortlichkeit für mögliche Diskriminierung gestaltet sich schwierig, wenn algorithmische Systeme involviert sind (Tursunbayeva *et al.*, 2022). Darüber hinaus birgt die Anwendung von KI das Risiko von Pfadabhängigkeiten, bei denen einmal eingeschlagene Richtungen und Entscheidungen zu langfristigen Auswirkungen führen können, ohne die Möglichkeit einer einfachen Umkehr. Self-Fulfilling Prophecies, bei denen KI-Systeme aufgrund ihrer eigenen Prognosen zukünftige Ereignisse beeinflussen können, sind ebenfalls eine Sorge (Giermindl *et al.*, 2022). Ein Beispiel dafür ist die (vermeintliche) Identifizierung von besonders produktiven Beschäftigten mittels



KI. Werden die vom System als herausragend eingestuft Personen besonders gefördert und erzielen dementsprechend herausragende Leistung, ist letztendlich nicht mehr erkennbar, ob die Leistung der Beschäftigten tatsächlich immer schon herausragend war, oder ob die Leistung erst durch die gezielte Aufmerksamkeit und Förderung nach der Diagnose des Systems herausragend geworden ist. Ein weiteres Risiko ist der Kompetenzverlust, insbesondere bei Führungskräften, da übermäßige Abhängigkeit von KI dazu führen könnte, dass menschliche Fähigkeiten verkümmern und in bestimmten Bereichen die menschliche Kontrolle abnimmt (Klöpper und Köhne, 2023). Dies kann insbesondere dann eintreten, wenn menschliche Entscheider:innen die Wirkweisen der Systeme nicht nachvollziehen können und daher an die Unfehlbarkeit einer KI glauben (Tursunbayeva *et al.*, 2022). Die Reflexion dieser Herausforderungen erfordert eine sorgfältige Abwägung und eine umfassende ethische Betrachtung, um sicherzustellen, dass die möglichen Vorteile der KI-Technologie ohne unerwünschte Risiken realisiert werden können. Letzterer Aspekt leitet zum nächsten Abschnitt über, in dem in aller Kürze dargelegt wird, dass es am Ende wichtige Abwägungsprozesse bedarf, um sich für neue technologische „Pfade“ zu entscheiden. Diese Abwägung wiegt umso schwerer, als dass technologische Innovationen wie die KI als wichtige Lösungsstrategien für jegliche gesellschaftlichen Probleme herangezogen werden.

2.3. Problem- oder Technikzentrierung?

Dieser Leitfaden richtet sich, wie oben beschrieben, an Akteur:innen, die sich mit der Einführung von KI-Anwendungen befassen. Insofern wird in ihm vorausgesetzt, dass die Entscheidung zur Implementierung eines KI-Systems bereits getroffen worden ist. Im Folgenden sollen dennoch einige Aspekte dieser Entscheidungsfindung reflektiert werden:

Neben der Überprüfung der fachlichen, technischen und organisationalen Voraussetzungen im Betrieb („KI-Readiness“) ist vordringlich zu klären, welche Ziele durch den Einsatz von KI verfolgt werden sollen — und inwieweit KI-basierte Lösungsansätze diese Ziele (am besten) erreichen. Die Autor:innen wünschen sich ein problem- und lösungsorientiertes Vorgehen, das ohne zu frühe Festlegung auf eine Technologie nach der „besten“ Lösung sucht.

Zielsetzungen können beispielsweise die Stärkung der Konkurrenzfähigkeit, die Verbesserung von Arbeitsbedingungen sowie die Ressourceneinsparung sein. Dieser Abwägungsprozess findet dabei faktisch nicht im luftleeren Raum, sondern vor dem Hintergrund diskursiver Dynamiken („Hypes“) statt – einer weitverbreiteten Präferenz für vermeintliche technische Lösungen für soziale sowie organisatorische Herausforderungen („Solutionismus“) und förderpolitische Vorfestlegung, die in den frühen Phasen von Technologieentwicklung funktional sein können (Nachtwey und Seidl, 2017). Angesichts des Energiebedarfs und der ökonomischen Kosten, die mit der Einführung von KI-Systemen verbunden sind, sollte diesem Abwägungsprozess jedoch schon aus Gründen der ökologischen Suffizienz sowie der ökonomischen Rationalität Platz eingeräumt werden, um kosten- und ressourcenintensives „Overengineering“ zu vermeiden (Rohde *et al.*, 2021).

Die Entscheidung zur Einführung eines KI-Systems sollte entsprechend am Ende eines Abwägungsprozesses unter Einbindung der späteren Nutzer:innen sowie weiterer Betroffener stehen, in dem verschiedene soziotechnische Alternativen zur Problemlösung diskutiert



werden – von rein organisationalen Veränderungen des Arbeits- und Lernprozesses bis hin zu stärker technisch geprägten Lösungsansätzen, wie dem Einsatz von KI – und nicht vorausgesetzt werden (siehe Empfehlungsimpuls 8 in Kapitel 5). So wäre beispielsweise danach zu fragen, ob eine verbesserte Arbeitsorganisation tatsächlich am zielführendsten durch die Einführung eines KI-Systems zu realisieren ist oder alternative Ansätze, wie eine relativ ressourcenschonende Weiterbildung von Führungskräften im Bereich der Konfliktbewältigung und der Mitarbeitendenführung oder die Erweiterung von Autonomiespielräumen der einzelnen Beschäftigten, nicht schneller zu Verbesserungen in der Arbeitsorganisation führen können. Gerade die Erschließung des Erfahrungswissens der Beschäftigten kann dabei ein großes Innovationspotenzial erschließen (Gerlsbeck und Herzog, 2019). Dazu kann sich auch partizipativer Methoden, wie des Design Thinkings (Curedale, 2019), bedient werden, um Verbesserungspotenziale inklusiv zu erheben und Lösungsansätze unter Einbindung des Wissens der Betroffenen zu identifizieren.



3. Integration von ELSA in Vorgehensmodelle

Ethische, rechtliche und soziale Aspekte (ELSA) lassen sich durch unterschiedliche Mittel in Projekten zu KI-Systemen integrieren bzw. achten. Allen gemeinsam ist, dass sie einen Raum eröffnen, über diese Aspekte (in strukturierter und ggf. moderierter Weise) im jeweiligen konkreten Vorhaben nachzudenken. Dieses Nachdenken als Aktivität innerhalb der einzelnen Vorhaben lässt sich beispielsweise durch vorgelagerte Organisationsentwicklung und Ausbildung persönlicher Kompetenzen sowie allgemeine Forschung unterstützen (vgl. Abbildung 4 zu Beginn des nächsten Kapitels) – aber nicht ersetzen.

Ein Mittel ist, ELSA als „normale“ Anforderungen über das Requirements Engineering jeweils in einem Projekt zu verankern¹. Ganz so einfach ist es jedoch nicht, weil bereits die Erfassung von ELSA als Anforderungen in frühen Phasen einer Projektidee, insbesondere bei innovativ-disruptiven Ideen, für alle Beteiligte schwierig ist. Auch die weitere Beachtung ist herausfordernd, da eine Vielzahl nicht funktionaler qualitativer Anforderungen zu erwarten ist und typische Entwicklungsteams diese in größerer Menge nicht gut in ihre Entwicklungsarbeit einbeziehen können. In der Praxis der KARL-Use-Case-Projekte hat sich gezeigt, dass dieses Mittel nicht für eine effiziente Einbeziehung von ELSA ausreicht.

Eine andere Möglichkeit sind spezielle Vorgehensmodelle „nur“ für ELSA, wie das „IEEE Model Process for Addressing Ethical Concerns during System Design“ (IEEE 7000-2021). Diese wären dann zusätzlich und nebenläufig zu bereits eingesetzten Vorgehensmodellen des Software-Engineerings anzuwenden. Ein Nachteil besteht darin, dass die Komplexität wegen der nebenläufigen Anwendung mehrerer Vorgehensmodelle und der Aufwand durch Überschneidungen der einzelnen Vorgehensmodelle steigen.

Existierende Vorgehensmodelle des Software-Engineerings sowie solche zur Einführung von Standardsystemen berücksichtigen die Frage der Wirkungs- und Wertbeitragsproblematik von Informationssystemen nicht hinreichend (Schütte et al., 2019).

Daraus ergibt sich folgende These: „Wir brauchen mehr strukturelle Verankerungen von ELSA in Standardvorgehensmodellen“ (Alpers, 2022). Im Sinne dieser These wird im Folgenden gezeigt, wie eine strukturelle Verankerung von ELSA in Vorgehensmodellen adressiert werden kann. Hierfür wird zunächst der Begriff „Vorgehensmodell“ näher erläutert, bevor im Kapitel 3.2 eine kurze Erörterung der These (zusammengefasst aus Alpers, 2022 sowie Alpers und Krings, 2023) dargestellt wird.

3.1. Vorgehensmodell

Ein Modell ist eine vereinfachte Abbildung eines Modelloriginals. Im Kontext von Vorgehensmodellen ist das Modelloriginal immer ein Vorgehen. Ein Vorgehensmodell definiert, wie der Weg von einem Ausgangspunkt zu einem bestimmten Ziel ausgeführt werden soll. Durch das Vorgehensmodell wird das Vorgehen idealisiert dargestellt, indem es

¹ Hierfür werden teils spezifische Methoden entwickelt bzw. vorhandene Methoden angepasst, beispielsweise (Weinreuter et al., 2022).



vorschreibt, wie dieser Weg effektiv und effizient auszuführen ist (Fischer et al., 1998; Schütte et al., 2019). Ein Vorgehensmodell bietet einen Überblick über die wesentlichen Aspekte in einem Vorgehen, wodurch es Klarheit schafft und Komplexität reduziert. Es abstrahiert so weit von einem konkreten Vorgehen, dass die Beschreibung auf eine Vielzahl von Vorhaben angewandt werden kann (Fischer et al., 1998). Bei der Anwendung eines Vorgehensmodells in einem konkreten Vorhaben bietet das Vorgehensmodell so eine Hilfestellung in Bezug darauf, welche Aspekte zu beachten sind, lässt aber für die konkrete Ausgestaltung (Adaption) großen Freiraum (Kuhrmann und Linssen, 2014).²

Das Ordnungsschema eines Vorgehensmodells ist der Abbildung 3 zu entnehmen. Diese stellt es im Kern dar und ergänzt es hier konkret um den gesellschaftlichen Rahmen. Jener ist neben weiteren Rahmenbedingungen zu beachten. Dabei kann es Rahmenbedingungen geben, die vordringlich sind, und solche, die nachrangig gegenüber gesellschaftlichen Rahmenbedingungen zu adressieren sind.

Ein Vorgehensmodell besteht aus einer Kombination von Methoden (Kuhrmann und Linssen, 2014; Alpers, 2022), welche im Rahmen festgelegter Prinzipien ausgewählt bzw. adoptiert wurden (Schütte et al., 2019).

Die Inhalte einer Methode sind in weiten Teilen dem Begriffsverständnis des „Method-Engineerings“ (Brinkkemper, 1996) entnommen. Dabei besteht jede Methode aus Methodenfragmenten bezüglich des Vorgehens und Methodenfragmenten hinsichtlich der Artefakte.

Die Kernkomponente des Vorgehens ist ein Prozess, typischerweise beschrieben durch ein Prozessmodell (*teils auch als Vorgehensmodell im engeren Sinne bezeichnet*) in einer Modellierungssprache, wie BPMN 2.0³ oder Petri-Netze (Freund und Rücker, 2019; Reisig, 2010; Adamo et al., 2021).

Das Prozessmodell legt fest, unter welchen Voraussetzungen

- welche Aktivitäten (elementare Arbeitsschritte innerhalb der Prozessdefinition)
- von welchen Rollen (und ggf. von welchen anderen Ressourcen, wie IT-Systemen)

in welcher logischen Ordnung ausgeführt werden (Schütte et al., 2019; Adamo et al., 2021; Rosenthal et al., 2021; Alpers, 2022; Johannsen et al., 2023;). Im Rahmen der einzelnen Aktivitäten werden Artefakte erarbeitet und (von späteren Aktivitäten) (wieder-)eingesetzt (Fischer et al., 1998).

Die Art und Weise, wie Artefakte (d. h. Ergebnisse) erarbeitet und beschrieben werden, werden durch die Methodenfragmente bezüglich der Artefakte unterstützt und teils auch weiter spezifiziert. Dabei legt die Sprache fest, wie ein Artefakt zu dokumentieren ist (Kuhrmann und Linssen, 2014; Adamo et al., 2021; Alpar et al., 2023). Die Technik gibt durch Vorschriften vor, wie ein Artefakt zu erzeugen ist (Fischer et al., 1998, Kuhrmann und Linssen,

² Die konkrete Ausgestaltung ist nicht nur möglich, sondern auch notwendig; und die richtige Nutzung der Freiheitsgrade ist für den Erfolg in einem konkreten Kontext mitentscheidend.

³ Business Process Model and Notation Version 2.0, Standard der Object Management Group, abrufbar unter <https://www.omg.org/spec/BPMN/2.0.2/PDF>.

2014;). Zur Unterstützung einer Technik können Werkzeuge eingesetzt werden. Diese haben sich in der Praxis als vorteilhaft erwiesen, sind in der Regel (teil-)automatisiert und sollen den Einsatz von Techniken erleichtern (Brinkkemper, 1996; Johannsen et al., 2023).

Im Kontext von Softwareentwicklungsprojekten gibt es etwa Sprachfestlegungen für Modelle (z. B. zur Beschreibung von Datenstrukturen) und für Quellcode. Als Technik würden sich etwa Modellierungstechniken, aber auch objektorientiertes, prozedurales oder deklaratives Programmieren nennen lassen. Beispiele für Werkzeuge wären Modellierungsumgebungen, wie Horus Business Modeler, oder Entwicklungsumgebungen, wie das SDK von Eclipse.

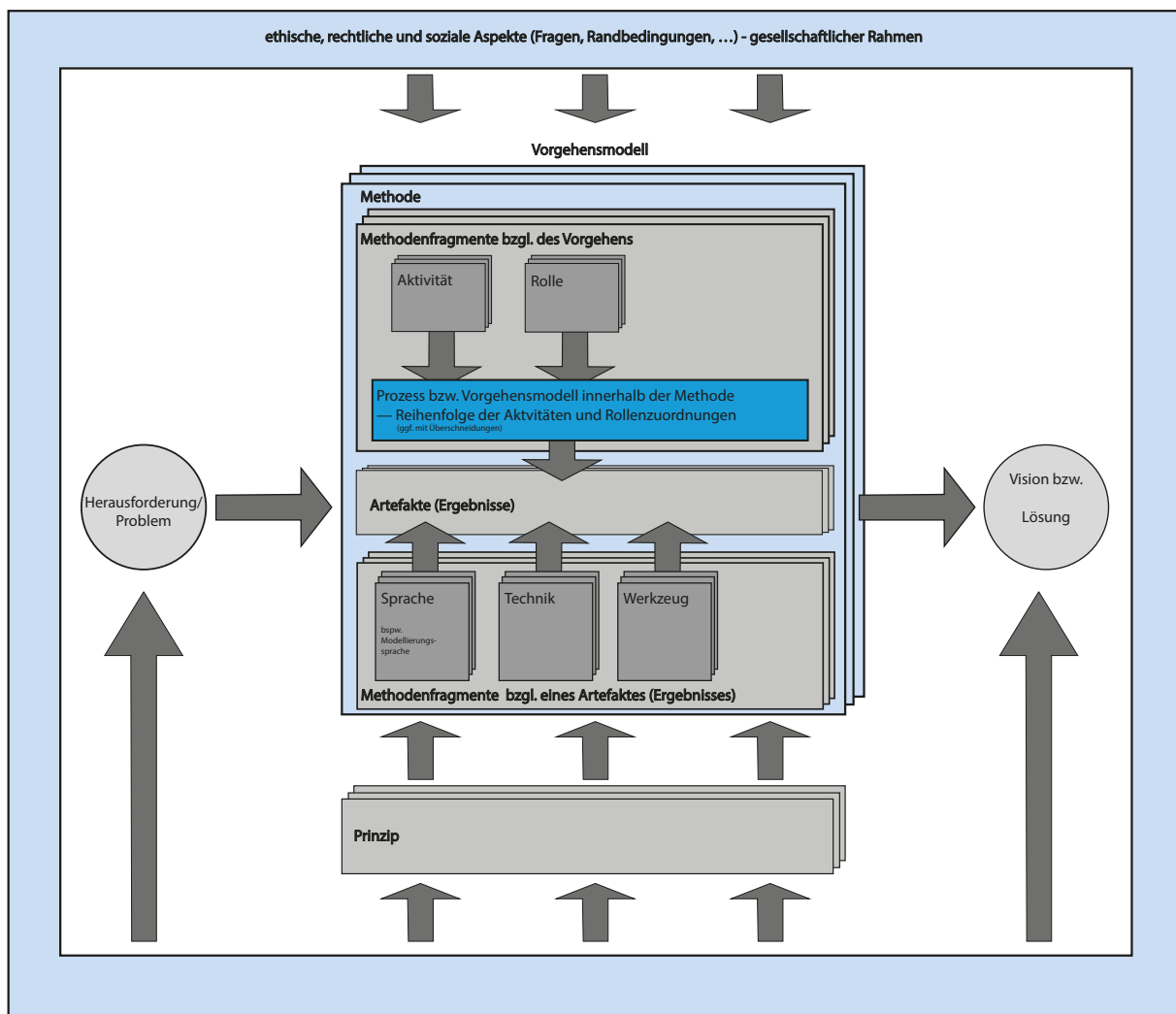


Abbildung 3: ELSA und Vorgehensmodelle für Softwareentwicklungs- und Softwareeinführungsprojekte (Alpers, 2022)



3.2. Integration von ELSA in Vorgehensmodelle des Software-Engineerings

In Abbildung 3 sind ELSA als Teil des gesellschaftlichen Rahmens für Vorgehensmodelle dargestellt. Freilich gibt es auch weitere Rahmenbedingungen – und diese stehen in einem Verhältnis zueinander. Aus Sicht der Autor:innen müssen sich beispielsweise betriebliche Rahmenbedingungen den gesellschaftlichen unterordnen. Auch Epping definiert das Vorgehensmodell daher ausgehend vom Begriff „Wert“: „Ein Vorgehensmodell ist eine Zusammenstellung von Elementen, die auf einem oder mehreren Werten basieren und von denen jedes durch eine oder mehrere Techniken umgesetzt wird“ (Epping, 2011).

Die strukturelle Verankerung von ELSA in Vorgehensmodelle kann (vgl. Abbildung 3) beispielsweise durch spezifische Aktivitäten (z. B. Quality Gates mit Ethikcheck, Einbeziehung von „Betroffenen“, wie der Vertretung von Arbeitnehmer:innen), spezifische Rollen (etwa Ethikverantwortliche), konkrete Verankerung der Aktivitäten und Rollen in einem Vorgehensmodell sowie spezifische Artefakte (z. B. Werteregister) erfolgen. Hierdurch wird eine systematische Betrachtung von ELSA gefördert. Die Integration ist dann besonders vielversprechend, wenn es gelingt, etablierte Standardvorgehensmodelle anzupassen. Davon betroffen sind sowohl etablierte Vorgehensmodelle innerhalb von Unternehmen als auch Standardvorgehensmodelle für Projekte innerhalb einer Branche.

Ein Beispiel hierzu ist das im Bereich der Softwareentwicklung verbreitete Vorgehensmodell SCRUM (Hanser, 2010). Ein konkreter Vorschlag zur Integration von ELSA besteht darin, diese als Annahmekriterien von Anforderungen (z. B. innerhalb der Definition of Ready, wie sie auch mehrere andere agile Methoden kennen) oder Abnahmekriterien (etwa innerhalb der Definition of Done) zu formulieren (Silva et al., 2017; Dalton, 2019).

Für die weitere Erörterung der These wird auf (Alpers, 2022) sowie auf (Alpers und Krings, 2023) verwiesen.

4. Handlungsempfehlungen für Akteur:innen in (neuen) KI-Vorhaben in Forschungsorganisationen und Unternehmen

Nachfolgende Handlungsempfehlungen sind in vier analytische Ebenen gegliedert, die sich als zentrale Handlungsfelder im Lauf der Projektarbeit herauskristallisieren werden. Diese können in ein idealisiertes Phasenmodell (vgl. Abbildung 4) eingebettet oder davon unabhängig betrachtet werden. Ihre konkrete Ausgestaltung ist abhängig vom jeweiligen Kontext, in dem KI gestaltet werden soll; auch sind die analytischen Ebenen zum Teil eng miteinander verwoben. Abbildung 4 zeigt einen idealisierten Ablauf eines KI-Gestaltungsprozesses, in dem die einzelnen Kontexte, die von der Einführung der KI-Systeme beeinflusst werden, analytisch getrennt dargestellt sind. Dies stellt sich in der Praxis oft anders dar. Dort sind Fragen der Organisationsentwicklung beispielsweise selten sauber von der Frage der Kompetenzentwicklung der einzelnen Beschäftigten zu differenzieren. Darüber hinaus unterscheiden sich die Gestaltungsprozesse zwischen einer Forschungseinrichtung und einem Unternehmen. Die Vorgehensweisen folgen divergierenden Methodensets und die Zielstellungen sind sehr verschieden angelegt. Der Leitfaden wird dieser Differenzierung gerecht, indem im Folgenden zwei soziale Gruppen, die der Forschenden sowie die der Betriebsräte / Beschäftigten, mehrheitlich getrennt betrachtet und angesprochen werden.

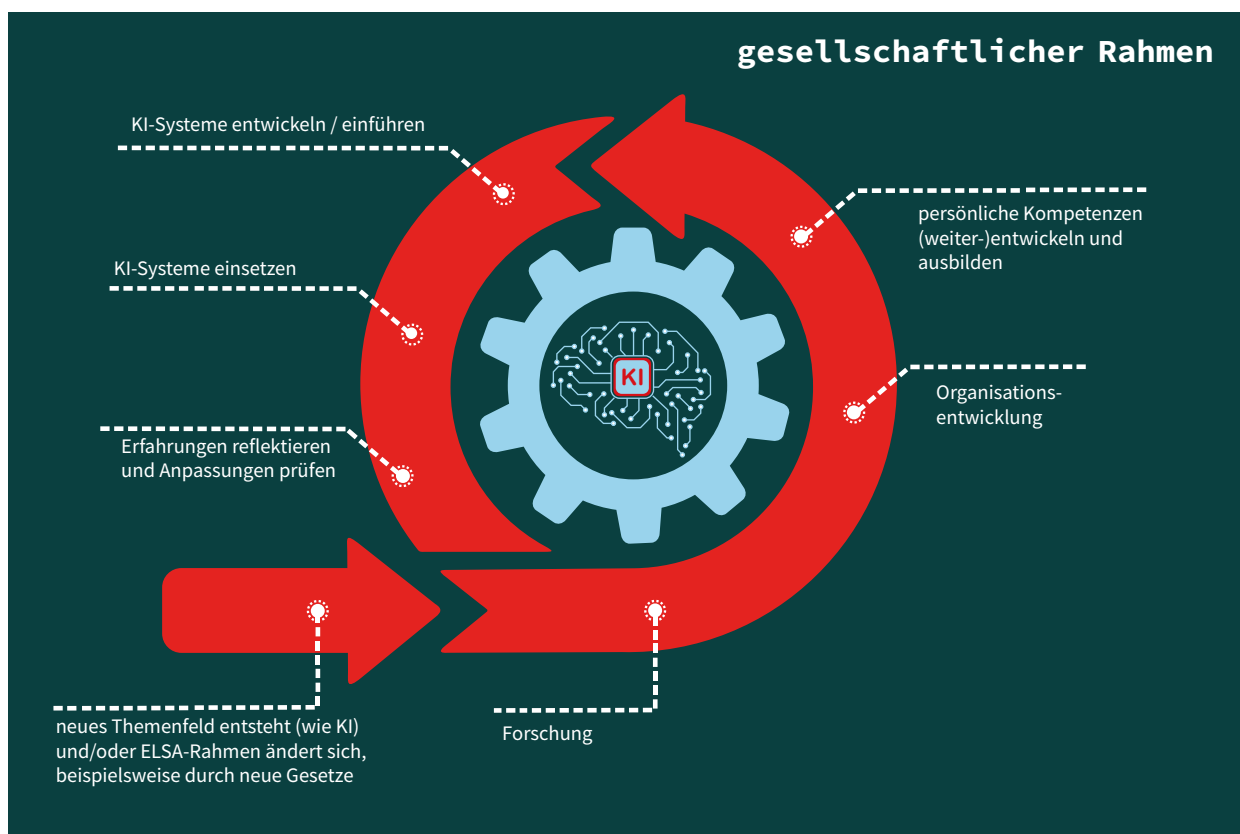


Abbildung 4: Anordnung der analytischen Ebenen in einem idealisierten Phasenmodell



Die Ausgestaltung dieser idealisierten analytischen Ebenen muss regelmäßig überprüft und angepasst werden. Hierzu sind Erfahrungen in konkreten Vorhaben zu analysieren und auszuwerten. Auf Basis der so gewonnenen Erkenntnisse erfolgen dann Anpassungsvorschläge auf den genannten analytischen Ebenen. Diese können etwa den Zuschnitt der einzelnen Ebenen oder ihren Stellenwert umfassen. Beispielsweise können zusätzliche Qualifikationsbedarfe identifiziert werden, die die Frage der Kompetenzentwicklung in ihrer Bedeutung aufwerten, oder Probleme bei der Einführung eines algorithmischen Systems machen die Notwendigkeit organisatorischer Weiterentwicklungen, etwa in der Zusammenarbeit von Sozialpartner:innen und Entwickler:innen, deutlich. Die Anpassungsvorschläge sind dann in möglichst inklusiver Weise zu erörtern und ggf. aufzunehmen – sowie bei Bedarf erneut zu prüfen. So wird die Ausgestaltung dieser analytischen Ebenen iterativ weiterentwickelt.

4.1. Analytische Ebene: Haftungsmodelle im Zusammenhang mit KI

Die Haftung im Zusammenhang mit dem Einsatz von KI kann sehr vielschichtig sein, da KI-Technologien vergleichsweise noch neu sind und sich ständig weiterentwickeln. Mit der schnellen Entwicklung und Verbreitung von KI innerhalb kürzester Zeit hat sich nicht nur unser technologisches Terrain verändert, sondern auch eine komplexe Vielfalt an rechtlichen Problemfeldern und Fragestellungen aufgeworfen. Von grundlegendem Interesse ist hierbei die Frage der Haftung und den verschiedenen Haftungsmodellen beim Einsatz von KI. Zum jetzigen Zeitpunkt gibt es keine spezifischen Gesetze, die sich ausschließlich mit der Haftung von KI befassen. Daher müssen bestehende rechtliche Rahmenbedingungen und Präzedenzfälle herangezogen werden, um die Haftungsfragen zu klären.

Vorliegend soll ein besonderes Augenmerk auf die rechtlichen Rahmenbedingungen verschiedener Haftungsmodelle beim Einsatz von KI im Arbeits- und Lernkontext liegen. In Deutschland existieren grundsätzlich verschiedene Haftungsmodelle, die sowohl auf vertraglicher als auch auf gesetzlicher Basis beruhen, sich daneben aber auch abhängig von der Geschäftsbeziehung oder bei arbeitsrechtlichem Kontext unterscheiden können. Eine umfassende Betrachtung dieser Haftungsmodelle ist unerlässlich, um die rechtlichen Implikationen von KI-Anwendungen zu verstehen. Im Folgenden werden die Haftungsmodelle kurz erläutert und die relevanten gesetzlichen Normen benannt.

Die vertragliche Haftung ist grundlegend im Zusammenhang mit dem Einsatz von KI-Systemen. Vertragliche Vereinbarungen zwischen den beteiligten Parteien können Haftungsregelungen enthalten, darunter Haftungsausschlüsse, Haftungsbeschränkungen und Schadensersatzregelungen und dementsprechend die jeweiligen Rechte und Pflichten, sowie die Verantwortlichkeiten der Vertragsparteien durch den Vertrag festlegen. Hierbei kommt dem Bürgerlichen Gesetzbuch (BGB) eine zentrale Bedeutung zu, insbesondere den §§ 311 und 280 bis 280e BGB, die sich mit Vertragsverletzungen und Schadensersatz auseinandersetzen (Ebers et al., 2020). Ein Beispiel für die vertragliche Haftung einer KI könnte ein Entwicklungsvertrag für eine KI-Anwendung sein, in dem bestimmte Garantien oder Gewährleistungen festgelegt werden. Falls die KI die vereinbarten Anforderungen nicht erfüllt,



kann der Entwickler verpflichtet sein, den entstandenen Schaden zu kompensieren (Hoeren, Sieber, Holznagel, 2021).

Daneben kann auch die Produkthaftung bei der Haftung für Schäden, die durch fehlerhafte Produkte verursacht werden, eine bedeutende Rolle spielen. Das Produkthaftungsgesetz (ProdHaftG) regelt die Haftung des Herstellers für solche Schäden. Gemäß §§ 1-17 ProdHaftG können Herstellende für Schäden haftbar gemacht werden, die durch ein fehlerhaftes Produkt entstehen. Darunter können auch KI-Systeme fallen, die als Produkte betrachtet werden. Ein Beispiel für die Produkthaftung einer KI wäre ein autonomes Robotersystem, das in einer Fabrik eingesetzt wird und aufgrund eines Fehlers im System einen Schaden verursacht. In einem solchen Fall könnten die Herstellenden des Robotersystems gemäß ProdHaftG haftbar gemacht werden (Beierle, 2021).

Eine Partei kann auch dann für Schäden haftbar gemacht werden, wenn sie ihre Sorgfaltspflichten verletzt oder gegen gesetzliche Vorschriften verstößt. Hierbei sind die Bestimmungen des § 823 BGB entscheidend, wonach eine Haftung für Schäden entsteht, die durch eine Rechtsverletzung verursacht werden (Münchener Kommentar zum Bürgerlichen Gesetzbuch Bd. 7, 2022). Im Zusammenhang mit KI könnte dies beispielsweise der Fall sein, wenn ein Unternehmen in der Verarbeitung personenbezogener Daten gegen die Bestimmungen der Datenschutz-Grundverordnung (DSGVO) verstößt und dadurch Schäden verursacht. In einem solchen Fall könnte das Unternehmen für die Datenschutzverletzung haftbar gemacht werden.

Die Haftung im Zusammenhang mit KI beschränkt sich jedoch nicht ausschließlich nur auf das Vertragsrecht und den Verbraucherschutz, sowie die allgemeinen gesetzlichen Regelungen im BGB und ProdHaftG. Daneben sind auch Datenschutz und Datensicherheit von entscheidender Bedeutung, insbesondere wenn personenbezogene Daten durch KI-Systeme erhoben oder verbreitet werden. Unternehmen müssen sicherstellen, dass sie die geltenden Datenschutzbestimmungen einhalten, sie demnach personenbezogene Daten rechtmäßig verarbeiten und darüber hinaus angemessene Sicherheitsmaßnahmen zum Schutz dieser Daten implementieren, um mögliche Risiken im Zusammenhang mit der Datensicherheit und dem Datenschutz angemessen bewerten und behandeln zu können (Gola und Klug, 2023). Bei Verstößen gegen den Datenschutz können Haftungsansprüche gemäß der DSGVO geltend gemacht werden. Insbesondere die Artikel 5, 32 und 82 der DSGVO sind hierbei relevant (Wolff et. al. 2023). Ein Beispiel für die Haftung im Zusammenhang mit Datenschutz und Datensicherheit wäre ein Unternehmen, das eine KI-Anwendung zur Datenanalyse nutzt und dabei versehentlich personenbezogene Daten offenlegt. Dadurch kann es zu einem Verlust der Privatsphäre und zu finanziellen Schäden für die betroffenen Personen kommen. In einem solchen Fall könnte das Unternehmen für die Datenschutzverletzung haftbar gemacht werden (Auer-Reinsdorff und Conrad, 2019).

Die Haftungsfragen im Zusammenhang mit künstlicher Intelligenz bemessen sich darüber hinaus an weiteren verschiedenen rechtlichen Rahmenbedingungen, je nachdem welche Beziehung zwischen den Vertragsparteien besteht. Eine grundlegende Unterscheidung liegt hierbei in dem Verhältnis von Verbraucher:innen zu Unternehmer:innen (Business-to-Consumer - B2C) oder Geschäftsbeziehungen zwischen Unternehmen (Business-to-Business - B2B).



Im B2C-Bereich gelten die bestehenden Verbraucherschutzgesetze (Münchener Kommentar zum Bürgerlichen Gesetzbuch Bd. 3, 2022). Wenn Unternehmen KI-Technologien einsetzen, um Produkte oder Dienstleistungen an Verbraucher:innen zu verkaufen, müssen sie sicherstellen, dass diese den gesetzlichen Anforderungen entsprechen. Unternehmen können für Schäden haftbar gemacht werden, die durch fehlerhafte KI-Anwendungen oder unsachgemäßen Einsatz entstehen, insbesondere wenn dadurch die Sicherheit oder die Rechte der Verbraucher:innen beeinträchtigt werden. In einigen Fällen können die Herstellenden der KI-Technologie selbst haftbar sein, während in anderen Fällen das Unternehmen, das die KI einsetzt, die Verantwortung trägt (Schrader, 2022).

Verbraucher:innen sind in diesem Bezug diejenigen Vertragsparteien, die KI-basierte Dienstleistungen oder Produkte nutzen. Das deutsche Verbraucherschutzrecht enthält Regelungen, die Verbraucher:innen vor unlauteren Geschäftspraktiken und Produkten schützen sollen, weshalb Verbraucher:innen dann Ansprüche geltend machen können, wenn KI-basierte Produkte oder Dienstleistungen nicht den erwarteten Standards entsprechen oder zu Schäden führen (Schrader, 2022).

Im B2C-Bereich gelten in Deutschland umfangreichere Schutzmaßnahmen zum Wohl der Verbraucher:innen, als dies bei einem Unternehmerverhältnis (B2B) der Fall ist. Beim Verkauf von Waren oder Dienstleistungen durch ein Unternehmen an Verbraucher:innen, müssen Unternehmende gesetzliche Vorschriften einhalten. Verbraucher:innenverträge sind in Deutschland stark durch das Bürgerliche Gesetzbuch (BGB) und spezifische Verbraucher:innenschutzgesetze geregelt (Auer-Reinsdorff, 2023).

Unternehmen, die an Verbraucher:innen verkaufen, haften für Mängel der verkauften Produkte oder Dienstleistungen. Verbraucher:innen haben in solchen Fällen das Recht auf Gewährleistung, Rücktritt vom Vertrag, Minderung des Kaufpreises oder Schadensersatz. Das Gesetz geht davon aus, dass Verbraucher:innen weniger Verhandlungsmacht und Fachkenntnisse haben und daher zusätzlichen Schutz benötigen.

Darüber hinaus gibt es bestimmte gesetzliche Regelungen für den Fernabsatz (z.B. Online-Käufe), bei denen Verbraucher:innen ein Widerrufsrecht haben und ohne Angabe von Gründen vom Vertrag zurücktreten können.

Auch kann in diesem Kontext eine Haftung für fehlerhafte Produkte in Betracht kommen, deren Haftung im Produkthaftungsgesetz (ProdHaftG) festgelegt ist. Wenn ein KI-Produkt Schäden verursacht, könnten die Hersteller:innen haftbar gemacht werden, sofern das Produkt nicht den erforderlichen Sicherheitsstandards entspricht. (s.o.)

Die Haftung einer Geschäftsbeziehung zwischen Unternehmen (B2B) im Zusammenhang mit KI ist dem gegenüber haftungsrechtlich anders zu betrachten. Sie betrifft regelmäßig die Nutzung eines KI-Systems im Unternehmen zur Prozessverbesserung in verschiedensten Formen und kann besonders komplex sein, da sie oft von vertraglichen Vereinbarungen abhängt und dementsprechend ebenfalls eine grundlegende und bedeutende Rolle spielt. Hier gibt es weniger gesetzliche Schutzmechanismen im Vergleich zum B2C-Bereich, da man davon ausgeht, dass Unternehmen über eine größere Verhandlungsmacht und Fachkenntnisse verfügen oder sich diese durch die Hinzuziehung von Spezialisten verschaffen



(können). Im B2B-Kontext können die Vertragsparteien regelmäßig ihre eigenen Vertragsbedingungen aushandeln und festlegen, welche Haftung für Vertragsverletzungen oder Schäden bestehen sollen. Dabei sind die Risiken und potenziellen Schäden zu berücksichtigen, die aus dem Einsatz von KI resultieren könnten, wie beispielsweise fehlerhafte Entscheidungen aufgrund von Algorithmen oder Datenschutzverletzungen. Dabei kann eine Schadenskompensation beispielsweise durch Schadensersatz oder Vertragsstrafen geregelt werden, welche durch Service Level Agreements (SLAs) oder Gewährleistungsklauseln geregelt werden können (Auer-Reinsdorff, 2023).

Wenn Unternehmen KI-Technologien einsetzen, ist es besonders wichtig in den zugrundeliegenden Verträgen klar festzulegen, welche Leistungen von dem KI-System erwartet werden (können) und welche Haftung im Falle von Fehlfunktionen oder Nichterfüllung dieser Erwartungen besteht. Die Vertragsparteien sollten hier auch die Verantwortlichkeiten klar regeln und Haftungsausschlüsse oder -beschränkungen vereinbaren, um potenzielle Risiken zu begrenzen (Auer-Reinsdorff und Conrad, 2019).

Die genannten Haftungsmodelle und Beispiele sind vereinfacht dargestellt, weil die konkrete Haftung von verschiedenen Faktoren abhängt. Dazu gehören beispielsweise der Grad der Kontrolle über das System, die Art des KI-Systems selbst, das Verschulden oder das Vorhandensein von Haftungsausschlüssen in Verträgen. Da die Entwicklung von KI und die damit verbundenen vielfältigen rechtlichen Aspekte weiterhin rasch fortschreiten, ist es ratsam, aktuelle Gesetze und Gerichtsentscheidungen stets im Auge zu behalten und bei Bedarf rechtlichen Rat von Expert:innen einzuholen, um die spezifischen Haftungsfragen im Zusammenhang mit dem Einsatz von KI zu klären.

4.2. Analytische Ebene: Organisationsentwicklung

Das Themenfeld „Organisationsentwicklung von Arbeit“ ist ein komplexes und vielschichtiges Feld, das sehr viele Facetten beinhaltet und aus unterschiedlichen disziplinären Perspektiven betrachtet und bearbeitet wird. Das heutige Verständnis speist sich in hohem Maße aus der U.S.-amerikanisch geprägten Entwicklung des Human Resource Managements, das über gezielte Maßnahmen Verhaltens- und Einstellungsänderungen der Beschäftigten herbeizuführen versucht, um die organisationale Effektivität zu steigern. Diese Effektivität lässt sich in der Regel daran messen, wie Arbeits- und Prozessabläufe optimiert, die Flexibilität der Abläufe erhöht und die organisationale Innovationsfähigkeit gesteigert werden kann. Hierbei haben sich die Motive im Hinblick auf und Strategien der Organisationsentwicklung in den letzten Jahrzehnten signifikant verändert. Beispielsweise haben sich zentrale Organisationsstrukturen vielfach zu dezentralen Organisationsstrukturen entwickelt, Hierarchien wurden (teilweise) abgebaut, die Eigenverantwortung der Beschäftigten innerhalb der Organisationen wurde über Instrumente wie projektformig organisierte Arbeitsabläufe und Delegation von Verantwortung gestärkt. Externe Faktoren wie die Einführung von digitalen Technologien und ihren vielseitigen Möglichkeiten, neue nationale und globale Marktkonstellationen, neue Erwartungen an die Ausgestaltung von Arbeitsplätzen von Seiten der Beschäftigten haben seit den 1990er Jahren zu weitläufigen Veränderungen der Organisationsentwicklung geführt. Als Stichwort seien hier neue Formen einer



Humanisierung der Arbeitswelt genannt, die in allen Sektoren und beruflichen Bereichen zu mehr Flexibilität geführt haben sowie den Anspruch erheben, vermehrt Partizipation und Mitgestaltung auf Seiten der Beschäftigten durchzuführen.

Neben vielen Disziplinen wie Arbeitspsychologie, Arbeitsforschung, Betriebsingenieurwesen, die sich mit der Gestaltung von Organisationsentwicklung beschäftigen, sei die Perspektive der Arbeits- und Industriesoziologie in aller Kürze angeführt. Diese nimmt die Geschichte der Arbeitsteilung in den Blick, aus der sich nach Moldaschl (2010) zwei zentrale Prämissen für unterschiedliche Formen der Organisationsentwicklung entwickelt haben: (a) Der Grad der Arbeitsteilung wird wesentlich bestimmt von den verfügbaren Technologien; (b) zwischen dem Grad der Arbeitsteilung und den Qualifikationsanforderungen (bzw. den mit ihnen verbundenen Handlungsspielräumen der Arbeitenden) besteht eine negative Kausalbeziehung. Je schärfer die Arbeitsteilung, insbesondere die vertikale zwischen Planung und Ausführung, desto geringer die Qualifikationsanforderungen im Hinblick auf die Beschäftigten. Im Laufe der Geschichte der Organisationsentwicklung wurden diese Spannungsfelder kontinuierlich zwischen den involvierten gesellschaftlichen Gruppen, aber auch zwischen den Tarifparteien identifiziert und (teilweise kontrovers) austariert. Im Jahre 1984 wurde sogar das „Ende der Arbeitsteilung“ in der industriellen Arbeit verkündet (Piore und Sabel, 1984). Die Autoren konstatierten hier auf einer empirischen Grundlage eine Requalifizierung der Industriearbeit in europäischen Ländern, gleichzeitig beschrieben sie jedoch die Entwicklung von Kern- und Randbelegschaften, bzw. die Arbeitsteilung zwischen Ländern, wohin niedrig qualifizierte Arbeit in Billiglohnregionen weltweit ausgelagert wurde (Piore und Sabel, 1984). Diese Spannungsfelder prägen noch heute die Debatten um Organisationsveränderungen im Hinblick auf Qualifikation, Effizienzgewinne, Innovationsfähigkeit etc. Wichtig scheint angesichts der Einführung avancierter Technologien, das Ganze der Organisationsentwicklung in den Blick zu nehmen, also über die Grenzen der Organisationen hinaus, die weitreichenden Auswirkungen arbeitsteiliger Effekte zu bewerten. Gleichzeitig scheinen Analysen der Technikfolgen auf die Arbeitsumgebungen, die Analyse der Verteilung des (Erfahrungs-)Wissens in den Prozessabläufen sowie die Gestaltung von Aushandlungsprozessen im Rahmen der Modellierung von Organisation wichtige Voraussetzungen dafür zu sein, humanverträgliche Arbeitsplätze zu generieren.

Die Frage nach organisationalen Veränderungen im Hinblick auf die Einführung von KI-Systemen spielt in KARL zunächst keine explizite Rolle. Im Vordergrund des Projektes steht die spezifische Gestaltung von KI-Systemen in unterschiedlichen sektoralen Feldern sowie die Frage nach notwendigen Lernprozessen im Hinblick auf die Einführung dieser Systeme. Die Ideen sowie die Umsetzung dieser Projektziele finden in interdisziplinär und transdisziplinär organisierten Use Cases statt, die sich spezifischen Themenfeldern widmen. Die übergeordneten Ziele der Use Cases in KARL sind die potenzielle Anwendung von KI-Systemen und deren reflexive Überprüfung im Hinblick auf rechtliche, soziale und betriebliche Risiken. Die spezifischen und zu entwickelnden Anwendungsgebiete der KI sollen insgesamt eine Steigerung der Adaptivität und Flexibilität der sozio-technischen Systeme in den Unternehmen erreichen. So wird das weitreichende Ziel formuliert, die Erschließung neuer Automatisierungspotentiale im Rahmen der Planungs-, Organisations- und Produktionsprozesse voranzutreiben, die zu weitreichenden (a) Entlastungen der



Beschäftigten und zu (b) einer qualitativen Ergänzung und Erweiterung ihres Aufgaben- und Arbeitsportfolios führen sollen. Hierbei können durchaus neue Modelle der Arbeitsteilungen zwischen Menschen und Maschinen entstehen mit teilweise weitreichenden Folgen für die Organisationsentwicklung in den Betrieben. Hierbei gehen mit diesen Modellen „entsprechende technische, organisatorische und personalpolitische Gestaltungsbedarfe einher“ (Müller, 2023), die in KARL und vor allem in Gestaltungsfeld 5 identifiziert werden.

Der normative Ausgangspunkt und der Referenzrahmen in KARL basiert auf der Prämisse, dass der Mensch bei diesen Veränderungen im Mittelpunkt stehen soll, d.h. es sollen menschenzentrierte Arbeitsprozesse entstehen. Der Begriff der „Menschenzentrierung“ knüpft an die bestehenden Debatten im Diskurs der Industrie 4.0 an mit dem Ziel, dass es nicht nur darum gehen soll, menschliche Arbeitskraft durch KI-Systeme zu ersetzen, um höhere Effizienzgewinne für die Unternehmen zu erreichen (Frey und Osborne, 2017). Es geht zunehmend mehr darum, Technisierungsprozesse und deren Auswirkungen auf Organisationen als gesamtgesellschaftliche Aufgabe im Hinblick auf die Zukunft der Arbeit und zukunftsgerichtete Aufgabe zu erkennen und zu erschließen. Auf gewerkschaftlicher und politischer Ebene werden seit einigen Jahren „unter dem Label ‚Arbeit 4.0‘ Gestaltungspotentiale und neue Regulationserfordernisse digitaler Arbeit kritisch diskutiert“ (Hirsch-Kreinsen, Hartmut, 2023, S. 171). In KARL wird an diesen Diskurs auf der Basis rechtlicher, sozialer und ethischer Kriterien angeknüpft. Vor dem Hintergrund der aktiv geführten Debatte um die Gefahren von KI im Rahmen der Arbeit, aber auch in lebensweltlichen Tätigkeitsfeldern, spielt die Überprüfung rechtlicher, sozialer und ethischer Fragen bei der Veränderung organisatorischer Einheiten eine wichtige Rolle. Wie oben dargestellt, spielen die drei Themenfelder: Qualifikation der Beschäftigten, Effizienzgewinne im Rahmen der Produktions- und Arbeitsprozesse sowie die Innovationsfähigkeit eines Unternehmens eine große Rolle in der Gestaltung von Organisationseinheiten.

In der Literatur wird in diesem Kontext häufig die Frage gestellt, welche neuen Tätigkeitsfelder in den jeweiligen Arbeitszusammenhängen der Organisation entstehen. Hierbei wird der Fokus in der Regel auf das Themenfeld „berufliche Qualifikation“ sowie auf zukünftige „Kompetenzen“ der Beschäftigten im Umgang mit KI gelegt und weniger auf die aktive Gestaltung von Organisationseinheiten, bzw. der gesamten Organisation. Dies scheint auf den ersten Blick verwunderlich, da sich in den „vernetzten, intelligenten und teilweise interaktiven Systemen“ (Hirsch-Kreinsen, Hartmut, 2023, S. 47) ein großes Potential für organisatorische Veränderungen versteckt. So entstehen auf der Basis dieser Systeme Arbeitsprozesse, die neue Funktionalitäten in den Arbeitszusammenhängen und neue Schnittstellen zu benachbarten Tätigkeitsfeldern einführen und somit zu signifikanten organisatorischen Veränderungen führen können. Mehr noch: in den Visionen einer Industrie 4.0 beispielsweise wird das Potential einer technologisch autonomen Planung sowie die autonome Organisation von Produktionsprozessen formuliert und die Steuerung von Organisationseinheiten ins Auge gefasst (Pfeiffer, 2015). KI-basierte Systeme bilden hier wichtige Voraussetzungen, um diese Ziele zu erreichen. Je nachdem, wie und auf welche Weise die Eingriffstiefen dieser KI-Systeme ihre Wirkung entfalten. Wie oben dargestellt, unterscheidet die Gestaltungsgruppe 5 zwischen KI-Systemen als Assistenz- und KI-Systemen als Delegationstechnologien. Gleichwohl diese Unterscheidung in Arbeitsumgebungen wie in der Industrie beispielsweise fließend sind, bzw.



häufig eine Verschiebung der technischen Funktionen mit der Implementierung erfolgt, ist diese Unterscheidung für die Bewertung und für die Gestaltung der KI notwendig. Die Erfahrungen von Technikprozessen in Arbeitsumgebungen zeigen, dass sich digitale Technologien sehr unterschiedlich auf die Arbeitsprozesse auswirken. In der Regel jedoch perpetuieren sie organisatorische Prozesse, so dass diese kaum mehr rückgängig zu machen sind. Ein Beispiel ist etwa die Auslagerung der Kund:innenberatung von Unternehmen, die in Call-Center ausgelagert wurde und noch immer wird (Flecker, 2018). Die kontinuierliche Technisierung der Beratung durch KI-Systeme beinhaltet eine hochstandardisierte Qualität der MMI im Hinblick auf die Kund:innen sowie die Möglichkeit der Bewertung der Beratungsqualität auf Seiten der Kund:innen, angefragt durch die Unternehmen. Diese Veränderungen betreffen Prozesse auf allen Ebenen wie die Shopfloor-Ebene sowie indirekte Bereiche wie Planung, Steuerung und Verwaltungsaufgaben. Vor allem bei Themenfeldern wie „Steuerung“ und „Planung“ scheint es wichtig, detailliert nachzuvollziehen, wie sich auf der Basis von KI-Systemen Steuerungssysteme verändern und welche Auswirkungen diese Veränderungen auf Qualifikationsprofile der Beschäftigten, Team- oder Gruppenarbeit und/oder Kontrollinstrumente in den Unternehmen haben. Diese Aspekte weisen insgesamt auf ein hohes Potential organisatorischer Veränderungen, die umfassend reflektiert werden sollten. Veränderungen organisationaler Bereiche nachzuvollziehen, zu analysieren und im Sinne einer ‚guten‘ Organisationsentwicklung zu gestalten, wird als eine bedeutsame Herausforderung und Zukunftsaufgabe in diesem Leitfaden erkannt. Die Überprüfung dessen, was als eine „gute“ Organisationsentwicklung verstanden wird, orientiert sich an den rechtlichen Rahmenbedingungen (Betriebsverfassungsgesetz) sowie an einer Unternehmenskultur, die sich zu (ethischen) Leitlinien einer guten Unternehmenskultur bekennt. Folgende Aspekte werden in der Literatur als positiv für eine verbindliche Unternehmenskultur genannt (Franken, 2019).

- Integration von Mitarbeitenden im Unternehmen durch verbindliche Sinnorientierungen und Wertvorstellungen. Das Unternehmen wird dadurch zu einer Einheit – einzelne Mitarbeitender, Teams und Abteilungen begreifen sich als Teile eines Ganzen;
- Identifikation und Motivation: Identitätsbildung der Mitarbeitenden und des Unternehmens durch erfolgreiche Zusammenarbeit und Erfolgserlebnisse trägt zur Entwicklung eines Zusammengehörigkeitsgefühls bei, was eine stärkere Motivation der Mitarbeitenden zur Folge hat;
- Koordinierungs- und Orientierungsfunktion: Die Unternehmenskultur gibt Orientierung für individuelle und kollektive Entscheidungen aufgrund von gemeinsamen Erfahrungen und ergänzt formale Regeln und Abstimmungen durch ungeschriebene, selbstverständliche Abläufe und Beziehungen;
- Effiziente Kommunikation: Die Abstimmungsprozesse gestalten sich durch einheitliche Orientierung wesentlich einfacher, Entscheidungen werden beschleunigt, der Informationsfluss verläuft zum großen Teil informell, Konflikte können früher erkannt und gelöst werden;
- Steigerung der Effizienz und Dynamik des Unternehmens durch Lernprozesse und Wissensaustausch, erhöhte Innovationsbereitschaft, bessere Kommunikation, Kooperation und Konfliktlösung;

- 
- Ein positives Bild des Unternehmens in der Öffentlichkeit, besseres Ansehen bei Kund:innen, Lieferant:innen, Bewerber:innen.

Diese Aspekte entsprechen den Prinzipien der Durchführung des unten beschriebenen moderierten Spezifikationsdialog, in dem die Gestaltungsprozesse der KI-Systeme in den Unternehmen entwickelt werden sollen.



Handlungsempfehlungen Organisationsentwicklung:

- Identifizieren & analysieren Sie, wo und wie durch die Einführung eines KI-Systems Veränderungen im organisationalen Ablauf Ihrer Arbeitsprozesse entstehen!
- Berücksichtigen Sie hierbei folgende Fragen:
 - Wie verändern sich die qualifikatorischen Erwartungen an die Beschäftigten?
 - Welche Ebenen sind betroffen? Beispielsweise
 - Mensch-Maschine-Interaktionen
 - Ebene der Organisation (Teamzusammensetzung, Kontrollmechanismen, neue Formen der Arbeitsteilung etc.)
 - strategische Ausrichtung sowie Ziele des Unternehmens
- Diskutieren Sie im Team / Ihrer Arbeitsgruppe, welche Risiken Sie durch die Einführung von KI-Systemen sehen. Wie kann man mit diesen Risiken umgehen? Welches Wissen (z.B. technisches und/oder rechtliches Knowhow) brauche ich?
- Diskutieren Sie im Team/ Ihrer Arbeitsgruppe die Frage, ob diese Veränderungen den ethischen Grundsätzen der Unternehmenskultur Ihrer Organisation entsprechen?

4.3. Analytische Ebene: Persönliche Kompetenzen ausbilden

Die Ausbildung von persönlichen Kompetenzen im Berufsleben ist eng an die Organisationsentwicklung von Unternehmen sowie an politisch-institutionelle Rahmenbedingungen der jeweiligen Länder gebunden. Wie oben dargestellt, spielt die Betonung auf die Herausbildung beruflicher Kompetenzen in den Debatten um Organisationsformen seit vielen Jahrzehnten eine besondere Rolle, da hier im Hinblick auf Grade der Arbeitsteilung und/oder oder im Hinblick auf Grade der Mensch-Maschine-Interaktion (MMI) über die Ausgestaltung und Qualität der Arbeitsumgebungen, bzw. über vielseitige Möglichkeit der beruflichen Kompetenzerweiterung entschieden wird. Vor allem unter dem Label der „Humanisierung der Arbeit“ (HdA) entstanden in den 1960er und 1970er



Jahren in Deutschland vielversprechende Ansätze, die eine Antwort auf die Frage geben sollten, wie im Rahmen der monotonen und belastenden Fließbandarbeit in den Fabriken die beruflichen Anforderungen verbessert werden können. „Den verstärkten Motivations-, Beteiligungs- und Entfaltungsbedürfnissen der Beschäftigten entsprachen steigende berufliche und kooperative Anforderungen in einer flexiblen und qualitätsorientierten Fertigung“ (Oehlke 2017, S. 167). Diese Forderungen verschmolzen Ende der 1960er Jahre mit der Phase gesellschaftlicher Umbrüche („Studentenbewegung“), in der das Offenlegen politisch-ökonomische Widersprüche des Produktions- und Wirtschaftssystems mit dem Ruf nach „Arbeiterkontrolle, Fabrikräten und Selbstverwaltung“ (Oehlke 2017, S. 167) verschmolzen. Mitte der 1970er Jahre entstand eine Reihe von Initiativen in Politik, Forschung und auf Ebene der Gewerkschaften wie beispielsweise das Aktionsprogramm „Forschung zur Humanisierung des Arbeitslebens“ (1974-1989), das in der Folge zentrale reformpolitische Prämissen auf den Weg brachte. Eine Prämisse bezieht sich hierbei auf die „paritätische Beteiligungsregeln bei den Gestaltungsprojekten der Unternehmen“ (Oehlke 2017, S. 167), in der eine Vielzahl von rechtlichen und organisationalen Maßnahmen entwickelt wurden, um technische und wirtschaftliche Interessen mit humanen Aspekten in der Ausgestaltung der Arbeitsplätze zu verbinden. Im Hinblick auf die Förderung von Kompetenzen im Berufsbildungsbereich formuliert in Deutschland das deutsche Berufsbildungsgesetz „Mindeststandards für die duale Berufsausbildung, also den Erwerb beruflicher Qualifikation mit den Lernorten Betrieb und Berufsschule“ (Bosch 2017, S. 83). Im Rahmen dieser formalisierten Ausbildungswege werden die zu erlernenden Kompetenzen im Rahmen der Berufe definiert und festgelegt. Das Erlernen der Praxis sowie des Berufsalltags erfolgt in den Unternehmen und Betrieben.

Seit Jahrzehnten spielt im Zuge der fortschreitenden Digitalisierung der Berufswelt im nationalen und internationalen Kontext die Diskussion um neue Kompetenzen und Fähigkeiten der Beschäftigten in allen Sektoren eine herausragende Rolle. Auf internationaler Ebene haben sich die UNESCO sowie die ILO auf die Strategie geeinigt, „the study of technologies and related sciences, and the acquisition of practical skills, attitudes and understanding and knowledge relating to occupations in various sectors“ (Unevoc 2006 in: Bosch 2017, S. 82) in besonderem Maße zu fördern. Die Anfang der 2000er Jahre politisch und wirtschaftlich forcierte Debatte um „Wissensgesellschaften“ (Stehr, 2001) als Selbstbeschreibung hochindustrialisierter Gesellschaften führte darüber hinaus zu einem Bedeutungszuwachs von Wissen, Kompetenzen und den Zusammenhang von lernenden, bzw. ‚intelligenten‘ Systemen in beruflichen Umgebungen. Dieser Zusammenhang bezog auch qualifizierte und hochqualifizierte Berufsbilder mit ein und betont in besonderem Maße die aktive Rolle des Lernens auf Seiten der Beschäftigten genauso wie die Vorstellung des lebenslangen Lernens in beruflichen Kontexten (Krings, 2011).

Der Bezug zwischen der Digitalisierung von Produktions- und Organisationskonzepten und neuen beruflichen Kompetenzen und „skills“ markiert seit Jahren die Debatte um die Gestaltungsmöglichkeiten der KI. Hierbei sollen die Chancen akzentuiert werden, die sich angesichts neuer technischer Innovationsschübe gesamtgesellschaftlich ergeben. Hierbei spielen die Themen Automatisierung, die Evolution der Berufsprofile, aber auch die ökonomisch angelegte Standortdebatte eine wichtige Rolle (Krings *et al.*, 2021). . Sehr



ausgeprägt gilt dies für das Konzept der Industrie 4.0 sowie den aktuellen Debatten um die Einführung von KI in der Arbeitswelt. Die Erwartung dieser Debatten, die Optionen der KI-Systeme zu betonen, spiegelt sich im Projekt KARL, wo neue berufliche Kompetenzen in Arbeits- und Lernwelten der KI explizit identifiziert und gefördert werden sollen. Hierbei basieren die neuen Aspekte zukünftiger Kompetenzen einerseits auf eher klassischen Automatisierungsfunktionen, die die KI im Rahmen von Arbeitsprozessen generieren. Andererseits sollen die KI-Systeme in den Anwendungsfeldern neue Funktionen für die Arbeitsprozesse bereitstellen, die neue berufliche Kompetenzen erforderlich machen. Hierbei werden die Systeme in die Lage versetzt, „Informationsvorgaben je nach Arbeitssituation zu optimieren und zu ergänzen sowie autonome Planungs- und Steuerungsentscheidungen über Arbeitsabläufe zu treffen“ (Hirsch-Kreinsen 2017, S. 172). Gemäß den rechtlichen Rahmenbedingungen (vgl. Kapitel 4.1) sollen diese Entscheidungen jedoch an die Beschäftigten delegiert werden, da die letztendlichen Entscheidungen im Prozessverlauf nicht automatisiert werden dürfen. Gleichzeitig werden im Rahmen der KI-Systeme neue MMI eingeführt, die eine große Bandbreite von neuen Kompetenzen auf Seiten der Beschäftigten nach sich ziehen. Diese umfassen etwa das Erlernen von neuen Programmiertechniken, das Erlernen von kollaborativ angelegten Systemen für die Steuerung, die Kontrolle, die Wartung und die Instandhaltung der Produktionsprozesse, der Umgang mit neuen Verarbeitungssystemen von optischen, haptischen und/oder akustischen Signalen und der Einsatz von „teaching methods“, also der Umgang mit Sprechakten oder Sprachbefehlen (Fischer *et al.*, 2017). Diese Aspekte betreffen den Produktions- genauso wie den Dienstleistungsbereich der Use Cases in KARL.

Das Nachdenken sowie die Gestaltung über neue Kompetenzprofile kann – wie oben beschrieben- nur eingebettet in die funktionale Arbeitsteilung zwischen Mensch und Maschine in neuen KI-Arbeitskontexten betrachtet werden. Hierbei zeigt die Erfahrung, je „intelligenter“ die sozio-technischen Systeme, desto komplexer und vielschichtiger werden die Anpassungsprobleme für die Arbeitsprofile, die sich daraus ergeben. So stellt, auch im Projekt KARL, die Darstellung und die Gestaltung der MMI die eigentliche Herausforderung der aufgabenorientierten Perspektive dar. Noch umfangreicher stellt sich die Herausforderung dar, wenn die Bewältigung unvorhergesehener Ereignisse integriert werden soll, da dann die Entscheidungsfähigkeit im Rahmen des Systems gewährleistet werden muss. Hier bei muss beachtet werden, dass ein KI-System „keinen inhaltlichen Zusammenhang zwischen Wissen/Erfahrung und der aktuellen Problemsituation herstellen, sondern nur einen formalen Zusammenhang“ (Fischer *et al.*, 2017, S. 10). Dies wird in KARL ebenfalls berücksichtigt, wobei die Frage virulent wird, wie dieses menschliche Wissen die Arbeitsprozesse bereichern und befruchten kann.



Handlungsempfehlungen Kompetenzentwicklung:

- Identifizieren & analysieren Sie, wo und wie durch die Einführung eines KI-Systems Veränderungen im organisationalen Ablauf Ihrer Arbeitsprozesse entstehen. **Welches Wissen der Beschäftigten wird von den Veränderungen berührt? Welche Wissensanteile werden an die KI-Systeme delegiert?**
- Wie verändert sich die Mensch-Maschine-Interaktion in Ihrer Arbeitsumgebung? **Welche neuen Qualitäten entstehen hierbei?** Diskutieren Sie diese im Team, auch unter der Frage nach Akzeptanz, Befürwortung und/oder Unbehagen.
- Welches **neue Wissen** (z.B. technisches und/oder rechtliches Knowhow) brauchen Sie angesichts der neuen Arbeitsumgebung mit KI-Systemen?
- Welche Fortbildungen brauchen Sie, um den neuen Kompetenzen gerecht zu werden? **Sind Sie an Weiter- und/oder Fortbildungen an Ihrem Arbeitsplatz interessiert?** Ergeben sich hier für Sie interessante Entwicklungsmöglichkeiten?

4.4. Analytische Ebene: Einsatz von KI-Systemen

Der Einsatz von KI im Arbeitsleben ist in vielen Bereichen schon längst Realität geworden. Dabei trifft man auch regelmäßig auf rechtliche Herausforderungen, die es zu bewältigen gilt. Im Folgenden wird insbesondere auf die datenschutzrechtlichen Anforderungen überblicksartig eingegangen. Der Fokus wird hierbei auf den möglichen Legitimationsgrundlagen, die den Einsatz von KI im Beschäftigtenkontext erlauben und den einzuhaltenden Datenschutzgrundsätzen mit Bezug zu Einsatzmöglichkeiten von KI liegen. Des Weiteren werden einige datenschutzrechtliche Grenzen des Einsatzes von KI aufgezeigt. Ein weiterer Schwerpunkt dieses Abschnitts werden die betriebsverfassungsrechtlichen Anforderungen, die für Unternehmen gelten in denen Betriebsräte existieren, darstellen. Hier wird auf die Mitbestimmungsrechte der Betriebsräte beim Einsatz von KI eingegangen.

4.4.1. Datenschutzrechtliche Anforderungen

Bei der Analyse von Leistung oder Verhalten der Arbeitnehmenden, werden meist personenbezogene Daten gemäß Art. 4 Nr. 1 DSGVO verarbeitet (Wolff et. al, 2023).

Da bei dem Einsatz von KI auch personenbezogene Daten der Beschäftigten verarbeitet werden sind die Grundsätze des Datenschutzrechtes aus Art. 5 DSGVO zu beachten. Fundamentale Grundsätze des Datenschutzrechts sind nach Art. 5 Abs. 1 lit a DSGVO die Transparenz sowie die Rechtmäßigkeit der Verarbeitung. Die Rechtmäßigkeit der



Verarbeitung kann nur dann vorliegen, wenn eine sogenannte Legitimationsgrundlage (Erlaubnis) für die Verarbeitung der personenbezogenen Daten vorliegt. Mögliche Erlaubnisgrundlagen können laut Art. 6 Abs.1 a-f DSGVO beispielsweise rechtliche Verpflichtungen oder gesetzliche Erlaubnisregelungen sein. Für die Verarbeitung im Beschäftigtenkontext bilden Art. 88 DSGVO bzw. § 26 Abs. 1 BDSG eine Erlaubnis für die Verarbeitung von Beschäftigtendaten. Personenbezogene Daten der Beschäftigten dürfen demgemäß dann verarbeitet werden, wenn dies für die Begründung, Durchführung oder Beendigung eines Beschäftigungsverhältnisses erforderlich ist. Die Datenverarbeitung darf dabei nicht unverhältnismäßig in das Persönlichkeitsrecht der Arbeitsnehmenden eingreifen (Düwell und Brink, 2017) und kann nach Art. 26 Abs. 2 BDSG auch aufgrund einer Einwilligung der Betroffenen erfolgen. Nach aktueller Rechtsprechung des EuGH vom 30.03.2023 könnte die Legitimation von Datenverarbeitungen auf Basis des § 26 Abs.1 S.1 BDSG rechtswidrig sein. Hier ist abzuwarten wie der deutsche Gesetzgeber auf das Urteil reagiert.⁴ Die Einwilligung der Beschäftigten wäre eine weitere mögliche Erlaubnisgrundlage, diese muss allerdings freiwillig erfolgen können. Aufgrund des Abhängigkeitsverhältnisses des Beschäftigten zum Arbeitgebenden ist diese Erlaubnisgrundlage besonders zu prüfen⁵ und beim breiteren Einsatz von Leistungs- und Analysefunktionen kaum praktikabel unter anderem auch da die Einwilligungen mit Wirkung für die Zukunft zurückgezogen werden können, wodurch ein weiterer Einsatz der KI-Anwendung nicht möglich wäre.

Die Transparenz soll den Betroffenen im Allgemeinen ermöglichen Informationen darüber zu erhalten wer, warum, wann, welche Daten über sie/ihn verarbeitet. Der Transparenzgrundsatz gilt auch im Beschäftigtenkontext, sodass den Arbeitnehmer:innen entsprechende Informationen auch beim Einsatz von KI vorliegen müssen (Weichert, 2023). Wird der Grundsatz insbesondere hinsichtlich der in Art. 7, Art 12. und Art.13 DSGVO normierten Anforderungen nicht beachtet, so kann dies zu einer unwirksamen Verarbeitung, beispielsweise weil eine erteilte Einwilligung nicht in informierter Weise getroffen werden konnte, führen. Ein weiter wichtiger Grundsatz bildet die Zweckbindung nach Art. 5 Abs. 1 lit. b DSGVO wodurch konkrete Verarbeitungszweck im Vorfeld der Verarbeitung zu identifizieren, zu benennen und zu beachten (Wolff et. al., 2023) sind.⁶

Zu beachten ist außerdem das Gebot der Datenminimierung aus Art. 5 Abs. 1 lit. c DSGVO, sodass die Datenverarbeitung auf das notwendige Minimum der Zweckerreichung zu beschränken ist. Im Arbeitsalltag könnten Arbeitnehmende beispielsweise durch ein Pseudonym in den Datensätzen repräsentiert werden, ohne einen unmittelbaren Personenbezug. Einen Schritt weiter geht die Anonymisierung der Beschäftigtendaten. Diese schützt laut Huff und Götz die Rechte der Mitarbeiter:innen am besten (Huff und Götz, 2019). Im Einklang mit den Datenschutzgrundsätzen ist die Anonymisierung von Daten grundsätzlich zu empfehlen sofern es der Zweck der Verarbeitung zulässt.

⁴ EuGH, 30.03.2023 - C-612/21

⁵ Entweder verfolgen der Arbeitgeber und die Mitarbeiter:in gleichgelagerte Interessen oder der Mitarbeiter:in entstehen rechtliche oder wirtschaftliche Vorteile, § 26 Abs. 2 BDSG. Siehe auch DSK Kurzpapier Nr. 14.

⁶ Zum Vorgehen bei Zweckvereinbarkeitsprüfungen und Zweckänderungen wird an dieser Stelle nicht eingegangen.



Die Komplexität des Zusammenspiels der Datenschutzgrundsätze lässt sich an folgendem Beispiel darstellen: Durch den Einsatz von KI kann die Leistung der Mitarbeitenden sehr leicht und ohne Probleme von dem Arbeitgeber:innen überwacht werden. Die Analyse dieser Daten kann im schlimmsten Fall zu Bewertungen, Abmahnungen oder auch Kündigungen führen. Gleichzeitig ist aber auch eine Verbesserung des Arbeitsklimas, der einzelnen Prozesse und der Mitarbeiter:innenzufriedenheit möglich. Als Verarbeitungszweck, das heißt als Grund für die Datenverarbeitung, könnte das Beispiel in Betracht kommen, wenn Arbeitgeber:innen ein Interesse an einer Verbesserung der Arbeitsergebnisse, der Effizienz und der Mitarbeiter:innenzufriedenheit haben. Werden dabei jedoch Eigenschaften, zum Beispiel Charakter- oder Persönlichkeits- oder Charaktereigenschaften, einer Mitarbeiter:in untersucht, die nichts mit seiner Tätigkeit zu tun haben, so genannte Persönlichkeitsanalyse, so ist diese Verarbeitung unzulässig, wenn sie nicht mehr von dem Zweck gedeckt ist (Joos, 2020) und hier das Persönlichkeitsrecht der Arbeitnehmenden überwiegt. Bezüglich der Intensität des Eingriffs in Persönlichkeitsrechte muss betrachtet werden, wie lange die Daten gespeichert werden, ob Verwechslungsmöglichkeiten möglichst ausgeschlossen sind, wer Zugriff auf die Daten hat etc. Bezüglich der Arbeitgebenden ist zu werten, ob durch die Erhebung und Verarbeitung der Daten beispielsweise Betriebsgeheimnisse geschützt werden sollen (Bongers et al., 2023).

Eine Einsatzmöglichkeit von KI liegt im Bereich des Human Resources Management und wird durch so genannte People Analytics Software realisiert. Dabei werden große Datenmengen aus unterschiedlichen Quellen analysiert, um daraus beispielsweise Informationen zur Leistung oder der Arbeitszufriedenheit ableiten zu können (Holthausen, 2021).

Werden beim Einsatz von KI im Arbeitskontext Entscheidungen mit rechtlicher Relevanz wie beispielsweise die Einstellung oder Kündigung einer Person (Abel, 2018), getroffen, so ist im Rahmen des Art. 22 DSGVO zu beachten, dass diese Entscheidungen nicht ausschließlich aufgrund der automatisierten Verarbeitung von personenbezogenen Daten getroffen werden dürfen, sondern stets ein Mensch die letztliche Entscheidung treffen muss. Art. 22 DSGVO verbietet grundsätzlich das so genannte Profiling, also die Verarbeitung von personenbezogenen Daten in Kombination mit einer Bewertung dieser Daten und die automatische Entscheidung aufgrund dieses Profilings (Rudkowski, 2019). Zur Entscheidungsunterstützung kann gemäß Art. 22 Abs. 2 DSGVO das KI-System jedoch herangezogen werden.

Mit der anfänglichen Verbreitung von KI-Systemen wurde argumentiert, dass ein computergesteuertes System objektiv und vorurteilsfrei bezüglich der Entscheidungsfindung angesehen werden können⁷, allerdings trifft auch ein KI-System durch die Analyse bestimmter Eigenschaften Entscheidungen oder Empfehlungen, die diskriminierend sein können (Bongers et al., 2023), was in einer Vielzahl an Studien und Forschungsarbeiten gezeigt wurde und zu beachten ist.⁸ Dabei ist entscheidend, welche Trainingsdaten dem System zur Verfügung gestellt werden und wie das System diese Daten dann jeweils verknüpft und bewertet, da dies letztlich die getroffenen Entscheidungen des Systems beeinflusst.

⁷ Der SCOT-Ansatz von (Bijker und Pinsch, 1987) argumentiert diesbezüglich, dass es keine Objektivität von Technik gibt.

⁸ Siehe beispielsweise in Fleck, Rounding, Özgül, „Künstliche Intelligenz in der Personalauswahl, 2022“



Ein weiteres Einsatzgebiet von KI kann die Analyse von Videoaufnahmen sein. Hierbei können biometrische Daten nach Art. 4 Nr. 14 DSGVO verarbeitet werden. Biometrische Daten werden mittels spezieller technischer Verfahren gewonnen und liefern Auskunft zu physischen, physiologischen oder verhaltenstypischen Merkmalen (Byers, Winkler und Stelter, 2023). Hierfür bedarf es entweder einer Einwilligung der Mitarbeitenden oder des Vorliegens der Voraussetzungen von § 26 BDSG (Bongers et al., 2023). Diese Videoaufnahmen könnten beispielsweise Arbeitsprozesse im Bereich der Montage aufnehmen, wobei aufgrund der konkreten Bedienung einer Maschine oder bestimmter Bewegungen oder Bewegungsabläufe Rückschlüsse auf Personen gezogen werden könnten.

4.4.2. Betriebsverfassungsrechtliche Anforderungen

Durch Inkrafttreten des Betriebsrätemodernisierungsgesetzes (BReModG) am 18.06.2021⁹ wurden Beteiligungsrechte des Betriebsrates hinsichtlich der Einführung und des Einsatzes von KI erweitert beziehungsweise manifestiert. Dieser ist in den Prozess der Einführung und des Einsatzes von KI am Arbeitsplatz zu integrieren.

Wie bereits erwähnt kann die Verarbeitung von Beschäftigtendaten in bestimmten Fällen auf Basis einer Einwilligung erfolgen. Sofern diese jedoch nicht praktikabel ist und andere gesetzliche Erlaubnisregelungen für eine gewünschte Datenverarbeitung nicht vorliegen kann es sich anbieten, eine Kollektivvereinbarung nach Art. 88 Abs. 2 DSGVO iVm § 26 Abs. 4 BDSG zu schließen.

Die Kollektivvereinbarung richtet sich nach § 26 Abs. 4 S. 2 BDSG iVm Art. 88 Abs. 2 DSGVO, § 75 Abs. 2 BetrVG. Als Kollektivvereinbarungen können laut Erwägungsgrund 155 der DSGVO auch Betriebsvereinbarungen angesehen werden. Auch bei Betriebsvereinbarungen sind die Grundrechte und vor allem das allgemeine Persönlichkeitsrecht der Beteiligten zu schützen (Maschmann, 2018), sodass stets eine Abwägung der Interessen von Arbeitgebenden und Arbeitnehmenden erforderlich ist.

Die Betriebsvereinbarung muss Angaben zum jeweiligen Zweck der Verarbeitung, der Zugriffsberechtigungen und einem Löschkonzept enthalten (Kühling et al., 2020). Im Kontext von KI-Anwendungen ist auch das KI-System selbst so genau wie möglich zu beschreiben, um Transparenz und Nachvollziehbarkeit schaffen zu können (Holthausen, 2021). Eine Umgehung der datenschutzrechtlichen Anforderungen mittels der Vereinbarung ist unzulässig (Kort, 2018).

Der Betriebsrat hat gemäß § 80 Abs. 1 Nr. 1 BetrVG zu überwachen, dass die Vorschriften der DSGVO und des BDSG eingehalten werden, woraus sich für den Betriebsrat aus § 80 Abs. 2 BetrVG ein allgemeiner Informationsanspruch ergibt. Darauf ist bei einer Einführung oder Änderung einer KI und bei der Verarbeitung personenbezogener Daten von Mitarbeiter:innen zu achten (Bongers et al., 2023).

§ 80 Abs. 3 S. 2 BetrVG berechtigt den Betriebsrat dazu, eine:n Sachverständige:n hinzuzuziehen, wenn, „der Betriebsrat zur Durchführung seiner Aufgaben die Einführung und

⁹ BGBl. I 2021 S. 1762.



Anwendung von Künstlicher Intelligenz beurteilen“ muss. Es muss dabei für die Erfüllung der Aufgaben des Betriebsrates erforderlich sein, dass man die sachverständige Person hinzuzieht, um fehlendes Wissen des Betriebsrates in fachlicher und/oder rechtliche Art ausgleichen zu können. Bei der Einführung von KI ist diesbezüglich stets von der Erforderlichkeit der Hinzuziehung einer sachverständigen Person auszugehen (Witteler, 2023).

Da eine Datenauswertung basierend auf KI eine technische Einrichtung sein kann, welche die Leistung und/oder das Verhalten von Mitarbeitenden überwachen kann, bestimmt der Betriebsrat gemäß § 87 Abs. 1 Nr. 6 BetrVG bei der Implementierung einer KI-Software mit (Kort, 2018).

Gemäß § 90 Abs. 1 Nr. 3 BetrVG ist es verpflichtend, den Betriebsrat rechtzeitig darüber zu unterrichten, wenn geplant ist, KI einzusetzen, welche sich auf das Arbeitsverfahren oder Arbeitsabläufe auswirkt. Hat der Einsatz von KI einen Effekt auf die Arbeit der Mitarbeitenden, muss der Betriebsrat frühzeitig eingebunden werden (Grambow, 2021). Gemeint ist damit, dass sobald man ernsthaft in Erwägung zieht, Maßnahmen durch- oder einzuführen, der Betriebsrat umfassend und mit entsprechenden relevanten Unterlagen informiert werden muss (Müller-Glöge et al., 2023).

Weiter besteht ein Mitbestimmungsrecht des Betriebsrates gemäß § 94 BetrVG bei Personalfragebögen und Beurteilungsgrundsätzen in Verbindung mit einem KI-System. „Personalfragebögen sind dabei alle formalisierten und standardisierten Informationserhebungen des Arbeitgebers“ (Raif und Swidersky, S. 352, 2017), Beurteilungsgrundsätze sind Regelungen, um Leistungen und das Verhalten von Mitarbeitenden zu bewerten, wobei sie nicht individuell gelten, sondern für eine bestimmte Gruppe von Mitarbeitenden.

§ 95 Abs. 2a BetrVG stellt klar, dass der Betriebsrat einzubeziehen ist, wenn im Rahmen der Auswahlrichtlinien gemäß § 95 Abs. 1 und Abs. 2 BetrVG KI zum Einsatz kommt. Die Auswahlrichtlinien beziehen sich gemäß § 95 Abs. 1 S. 1 BetrVG auf Personalfragen zu „Einstellungen, Versetzungen, Umgruppierungen und Kündigungen“ und umfassen Sachverhalte, „wenn eine KI-Anwendung eigenständig oder innerhalb eines von einem Dritten vorgegebenen Rahmens Auswahlrichtlinien trifft“ (BR-Drs. 19/28899, 22).

Ausblick:

Für die Zukunft sind die Regelungen der kommenden KI-Verordnung (KI-VO)¹⁰ zu beachten. Durch die KI-Verordnung, welche aktuell (Stand: August 2023) noch als Entwurf vorliegt, werden KI-Systeme in vier Risikogruppen unterteilt werden, wobei jede Risikogruppen entsprechende Pflichten für die in Art. 2 Abs. 1 des Entwurfes genannten Personengruppen (Anbieter:innen und Nutzer:innen) mit sich bringen wird. In Art. 3 Nr. 2 – 8 des Verordnungsentwurfes werden weitere Rollen genannt, die für die Arbeitswelt relevant werden können, je nachdem, wie man KI am eigenen Arbeitsplatz einsetzt oder plant, einzusetzen. Insbesondere die so genannten Hochrisiko-KI-Systeme, die in Art. 6 in

¹⁰ Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union v. 21.4.2021, COM(2021) 206 final.



Verbindung mit Anhang III des Verordnungsentwurfes definiert werden, stehen im Fokus der KI-Verordnung. Hochrisiko-KI-Systeme sind solche, die aufgrund ihres Einsatzes oder ihrer Eigenschaften wahrscheinlich ein hohes Risiko für die Gesundheit, Sicherheit oder die Grundrechte von natürlichen Personen darstellen. Dies ist dann jeweils im konkreten Einzelfall zu entscheiden. Laut Erwägungsgrund 36 des Entwurfs werden auch solche KI-Systeme, die „für Entscheidungen über Beförderung und Kündigung sowie für die Zuweisung, Überwachung oder Bewertung von Personen in Arbeitsvertragsverhältnissen“ eingesetzt werden, als hochriskant eingestuft. Die KI-Verordnung verlangt unter Anderem, dass Trainingsdaten entsprechende Qualitätsanforderungen erfüllen müssen (Art. 10 KI-VO), oder dass die KI-Systeme unter menschlicher Aufsicht stehen müssen (Art. 14 Abs. 1, Abs. 2 KI-VO), sodass in die Vorgänge des KI-Systems bei Bedarf eingreifen kann (Sesing und Tschech, 2022). Nr. 4 des Anhangs III der KI-Verordnung spricht explizit den Arbeitskontext an, sodass die kommende KI-Verordnung auch Einfluss auf die Arbeitswelt haben wird. Dies wird sowohl die Entwickelnden von KI-Systemen betreffen, als auch die Nutzenden, die sich an die Vorgaben der Verordnung halten müssen.



Handlungsempfehlungen

- Beachten Sie die Grundsätze des Datenschutzes aus Art. 5 DSGVO, wie beispielsweise die Zweckbindung oder die Datenminimierung und insbesondere Die Rechtmäßigkeit und Transparenz, über den kompletten Prozess hinweg. Prüfen Sie hinsichtlich der Zweckbindung regelmäßig, ob der ursprüngliche Zweck noch verfolgt wird oder sich geändert haben könnte.
- Anonymisieren Sie die personenbezogenen Daten wenn möglich.¹¹
- Machen Sie sich bewusst, dass ein KI-System aufgrund der Trainingsdaten Entscheidungen treffen wird, sodass es aufgrund dessen zu Diskriminierungen durch die KI kommen kann. Die Trainingsdaten sollten daher so objektiv und/oder divers wie möglich sein.
- Informieren Sie und beziehen Sie den Betriebsrat im Zusammenhang von der Einführung und Nutzung von KI am Arbeitsplatz frühzeitig und umfassend ein. Konkrete Handlungsempfehlungen sind im Folgenden Kapitel dargelegt.
- Verfolgen Sie die gerichtlichen Entscheidungen und Gesetze, die sich mit KI und deren Auswirkungen auf die Arbeitswelt beschäftigen, um bei Bedarf rechtzeitig Anpassungen vornehmen zu können.

¹¹ Eine Hilfestellung entsteht unter anderem im Kompetenzcluster ANYMOS. Siehe (Kneis et al., 2023).



5. Handlungsempfehlungen an die Sozialpartner:innen für erfolgreiche Mitbestimmungsschritte

In diesem Teil des Leitfadens wechselt die Perspektive: In den Vordergrund rücken Handlungsempfehlungen an die Sozialpartner:innen, an unternehmerisch denkende Geschäftsleitungen und sozial handelnde betriebliche Beschäftigtenvertretungen.

Das Vorgehensmodell für IT-Teams zur Kreierung eines neuen Software-Werkzeuges ist nicht identisch mit dem Vorgehensmodell zur Einführung neuer komplexer Techniken. Es ist ebenfalls nicht identisch mit dem Vorgehensmodell der Mitbestimmungspraxis. Wir haben es mit insgesamt drei Vorgehensmodellen zu tun.

Die nachfolgenden neun **Empfehlungsimpulse entlang des Vorgehensmodells der Mitbestimmungspraxis** zielen auf die Schaffung einer guten Unternehmens- und Aushandlungskultur im Rahmen der Mitbestimmung (siehe dazu auch Kapitel 2.3 und Kapitel 3 sowie Kapitel 4.2, Kapitel 4.3 und Kapitel 4.4 dieses Leitfadens). Die Impulse richten sich an Akteur:innen in großen und kleinen Firmen, in Produktion, Dienstleistung und Verwaltung.

Bei der Einführung von komplexen Softwaresystemen rückt zumeist der Vorgang der Technikimplementierung in den Vordergrund. Einführungsprozesse konzentrieren sich häufig auf das Ausrollen neuer Softwarelösungen. Das hier zu nennende „Vorgehensmodell“ beschreibt dagegen die Abläufe der Interessensaushandlung zwischen den Sozialpartner:innen in Produktion, Dienstleistung und Verwaltung. Das Vorgehensmodell stellt die Fragen der Mitbestimmung, der ethischen sowie sozial nachhaltigen Faktoren, der wechselseitigen Information und der Vertrauensbildung ins Zentrum (Schröter, 2019). Ein technik- und sozialinnovatives Vorgehensmodell setzt die Aushandlungen vor die Phase der Beschaffung, vor das Experimentieren und Erproben und vor den Technikstart. Dies gilt insbesondere für die geplante Anwendung sich selbstverändernder „Delegationstechnik“.

Im Mittelpunkt steht dabei der Umgang mit jener technischen Entwicklung, die es seit fast siebzig Jahren gibt. Diese Technik entfaltet sich weiter und hat heute eine Fülle von Bezeichnungen erhalten. Man spricht unter anderem von „Künstlicher Intelligenz“ (KI), von „Artificial Intelligence“ (AI), von „sogenannter Künstlicher Intelligenz“, von „Algorithmischen Steuerungs- und Entscheidungssystemen“, von „Lernenden Systemen“ oder „Denkenden Systemen“, von „Mustererkennungstechnik“ oder „Maschinellem Lernen“, von „Expertensystemen“ und „Chatbots“, von „Neuronalen Netzen“ oder „Deep Learning“ beziehungsweise „Intelligenten Systemen“.

Für das Thema „KI“ liegen Dutzende, wenn nicht gar über einhundert Definitionen aus den Kreisen von Forschungsgruppen, Interessenverbänden und politischen Vereinigungen vor. Jedoch gibt es keine vereinbarte gemeinsame interdisziplinäre Definition aus den Bereichen Informatik, Mathematik, Technik und Gesellschaftswissenschaften.

Unabhängig von den Lücken in der Definition wollen nachfolgende Empfehlungen dazu beitragen, dass in den Arbeitswelten die Einführung neuer technischer Lösungen auf der Basis gut ausgehandelter ethischer, rechtlicher und sozialer Standards (vgl. oben ELSA-Kriterien) gelingt.



Empfehlungsimpuls 1:

Was ist für den Mitbestimmungsprozess aus der Sicht der Arbeitswelt bei der kommenden Stufe der Digitalisierung wirklich neu?

In den letzten Jahrzehnten kam es immer wieder zu Neuerungen in der Arbeitswelt. Neue Modelle der Arbeitsorganisation, neue digitale Techniken und neue Möglichkeiten des Lernens forderten heraus. Oftmals war aber an manchen Neuerungen nur die Verpackung und die Bezeichnung neu, nicht aber der Inhalt. Doch hinter der bisherigen Digitalisierung hat sich etwas tatsächlich Neues entfaltet.

Sozialpartner:innen in den Betrieben, Dienstleistungszentren und Verwaltungen können erkennen, dass nicht nur einerseits die Arbeitswelt die Technik und andererseits die Technik die Arbeitswelt verändert. Menschen haben für Menschen eine überraschende Idee entwickelt. Diese Idee führte zu einem neuen Typ von Technik. Er unterscheidet sich von dem, was wir bisher kennen.

Die neue Software-Welt, die die Arbeitsplätze immer mehr erreicht, lässt sich in zwei große Richtungen gliedern: Handelt es sich um Software, die den Menschen unterstützt und die Entscheidungen beim Menschen belässt, oder steht die Nutzung von Software bevor, die selbst Entscheidungen zu treffen hat? Geht es um Unterstützungstechnik, also um „Assistenztechnik“ oder um Technik, auf die eine Entscheidungs-Vollmacht übertragen und delegiert werden kann, also um „Delegationstechnik“?

In der Fähigkeit, zwischen „Assistenztechnik“ und „Delegationstechnik“ unterscheiden zu können, liegt ein wesentlicher zukünftiger Erfolgsfaktor im Dialog der Sozialpartner:innen (Schröter, 2022). Dies wird zu einer Schlüsselfrage der Aushandlungsprozesse vor Ort.

Die Gestaltung, Einführung und Mitbestimmungspraxis für „Delegationstechnik“ verläuft zum Teil anders als jene Gestaltung, Einführung und Mitbestimmungspraxis für die „Assistenztechnik“.

Die Sozialpartner:innen können mit einer klug angelegten Mitbestimmung, Beteiligung und Mitwirkung der Beschäftigten bei diesem Umbau die gemeinsame Leistungsbilanz anheben. Der Wandel gelingt besser mit neuen Formaten arbeitsweltlicher Aushandlungen.

Empfehlungsimpuls 2:

Warum wir unterscheiden sollten, was „Nachholende Digitalisierung“ darstellt und was darüber hinaus geht?

In vielen mittelständischen Unternehmen, Handwerksbetrieben, Klein- und Kleinstfirmen hat sich seit den neunziger Jahren so etwas wie ein Digitalisierungs-Stau angesammelt. Entweder war die Netzanbindung, die Glasfaserverkabelung oder die Mobilfunkverbreitung nicht ausreichend belastbar. Oder es gab interne Gründe, die den digital gestützten Wandel der Arbeitskulturen verlangsamen.



Um sich im Ringen um Fachkräfte und Aufträge als besonders modern und „vorne“ darzustellen, greifen Wettbewerber nicht selten zu Marketingformulierungen. Sie bezeichnen ihre Technikimplementierungen als „smart“ oder „Vier-Null“ oder „KI“. In Wirklichkeit werden unter flotten Überschriften oftmals endlich jene software-technischen Lösungen eingeführt, die aus Sicht der Technikentwicklung schon zehn, fünfzehn oder zwanzig Jahre alt sind.

Ein solcher Marketing-Ansatz ist nicht zu empfehlen, denn er verursacht Misstrauen. Ein Impuls zum Wandel wirkt nicht sehr glaubwürdig, wenn ein tatsächlich neues komplexes Softwaresystem eingeführt wird, das den Namen „Künstliche Intelligenz“ verdiente, jedoch der Begriff vorher schon für etwas ganz anderes verwandt wurde.

Den Sozialpartner:innen ist nachdrücklich zu empfehlen, ganz uneitel und sprachlich nüchtern aufzutreten: Der Abbau des Digitalisierungs-Staus ist eine „Nachholende Digitalisierung“. Zur „Nachholenden Digitalisierung“ gehören unter anderem Anwendungen wie etwa Alternierende Telearbeit (HomeOffice), Mobiles Arbeiten, Nutzungen elektronischer Lernplattformen, Beschaffungsplattformen, GPS-Verfolgung, Techniken des Big Data, gängige Modelle des „Internets der Ding“ (IoT), Techniken der vorausschauenden Wartung, gängige Bots, elektronische Prozesssteuerungen, gängige Modelle der Mustererkennungstechnik, ältere AR/VR-Techniken, Cloud-Lösungen etc. Eine solche ehrliche Herangehensweise stärkt Vertrauen.

Zudem lehrt die Erfahrung: Die „Nachholende Digitalisierung“ schafft das Fundament für darauf aufbauende Anwendungen komplexer Software-Systeme. Ohne „Nachholende Digitalisierung“ wird der Sprung in komplexere Digitalisierungsprozesse kaum gelingen.

Empfehlungsimpuls 3:

Was ist unter dem Begriff der „Assistenztechnik“ zu verstehen?

Der Begriff der „Assistenztechnik“ entstammt der analogen, vordigitalen Zeit. Insbesondere in Fertigungs- und Produktionsverfahren wurde Technik gewünscht, die den Menschen beim Arbeitsvorgang im Maschinen-, Anlagen-, Werkzeug- und Automobilbau entlastet. Das Heben, Drehen und Wenden schwerer Bauteile sollte auf assistierende Geräte und mechanisch-automatisierte Systeme übergeben werden.

Mit der zunehmenden Einführung von digitalen Lösungen wurde der Begriff „Assistenztechnik“ im Bereich des Wohnens und der Unterstützung von älteren Menschen in ihrem Zuhause bei der Gesundheitsprävention neu aufgegriffen. Im „Smart Home“ sollte Assistenztechnik mehr Sicherheit und verbesserte Betreuung erbringen. Ein digitaler Notruftknopf und technische Lösungen bei Bewegungen wie zum Beispiel im Bad oder beim Treppensteigen geben dem Menschen mehr Wohnqualität.

In der Digitalisierung des Büro- und Verwaltungswesens trugen assistierende Software-Werkzeuge seit Jahrzehnten zur Erleichterung des Arbeitens bei. Assistenzsysteme haben sich auch in den Mobilitätsformen der KFZ-Branche durchgesetzt.



Kennzeichen jeder „Assistenztechnik“ ist es, dass die Entscheidung über die Nutzung und Handhabung der Technik beim handelnden Menschen liegt. Es gilt die Handlungsträgerschaft Mensch.

Gut und einvernehmlich gestaltete „Assistenztechnik“ dient der Humanisierung der Arbeit. Die zurückliegende Phase der Digitalisierung hat sich vor allem auf die Nutzung der „Assistenztechnik“ konzentriert.

Empfehlungsimpuls 4:

Was bedeutet „Delegationstechnik“?

Vor rund zwanzig Jahren begannen IT-Teams in ihren Labors mit der Entwicklung von „Softwareagenten“ und „mobilen Agentenplattformen“. Als Reaktion auf diese fachlichen Zusammenhänge entstand der Begriff „Delegationstechnik“. Aus dieser Zeit stammen die ersten Anpassungs- und Gestaltungsansätze.

Kennzeichen der „Delegationstechnik“ ist die softwaretechnisch entstandene Möglichkeit, einen rechtsverbindlichen Vorgang (Transaktion) auf eine Software zu übertragen, zu delegieren. Man könnte umgangssprachlich auch von einer „Vollmachtstechnik“ sprechen.

Dabei wechselt die Handlungsträgerschaft Mensch hinüber in eine Handlungsträgerschaft Softwaresystem. Das so bevollmächtigte algorithmische System wird ermächtigt, jenseits des Menschen verbindliche Entscheidungen zu treffen. Diese Handlungsbevollmächtigung unterscheidet „Delegationstechnik“ von „Assistenztechnik“.

Der Begriff „Delegationstechnik“ dient heute in der Debatte über die Potentiale sogenannter „Künstlicher Intelligenz“ zur Differenzierung der Ziele und Anwendungen. „Delegationstechnik“ zielt auf die Automatisierung des virtuellen Raumes und unter anderem auf die Möglichkeit der „selbstständigen“ Steuerung von Wertschöpfungsketten.

Für die Gestaltung der „Delegationstechnik“ bedarf es zusätzlicher Aushandlungsformate zwischen den Sozialpartner:innen. Ein zu empfehlendes neues Format ist der „Moderierte Spezifikationsdialog“.

Empfehlungsimpuls 5:

Was ist das Besondere an „sich selbst verändernder Delegationstechnik“?

Aus den arbeitsweltlichen Erfahrungen in Produktion, Dienstleistung und Verwaltung können wir ableiten, dass sich in der Praxis der „Delegationstechnik“ zwei unterschiedliche Denkrichtungen entfalten. Zu dieser Wahrnehmung kommt das Personennetzwerk „Forum Soziale Technikgestaltung“ nach mehrjähriger empirischer Bewertung.

Die Softwarelösungen der „Delegationstechnik“ finden sich in bestimmten Personalmanagementsystemen, in der Nutzung von Lieferketten und Beschaffungen, in ausgewählten elektronischen Lernumgebungen und schrittweise auch in neuen



Verwaltungsplanungen. Es ist damit zu rechnen, dass sich die Nutzungen der „Delegationstechnik“ als Querschnittstechnik in alle wirtschaftlichen, arbeitsweltlichen, sozialen und gesellschaftlichen Bereiche ausbreiten werden.

Das „Forum Soziale Technikgestaltung“, das den Begriff „Delegationstechnik“ vor Jahren schuf, rechnet damit, dass dieser Techniktyp alsbald die Nutzungsbreite der „Assistenztechnik“ überflügeln wird.

Es ist zu erkennen, dass „Delegationstechnik“ einerseits Software-Werkzeuge hervorbringt, die sich durch Anwendung nicht verändern. Sie bleiben auch nach längerem Gebrauch in ihrer Basiszielbestimmung. Sie verfeinern sich. Jedoch bleiben sie in ihren gesetzten Regeln. Die Fachwelt bezeichnet sie daher als „regelbasierte“ Systeme mit vordefinierten algorithmischen Entscheidungen. Sie vollziehen Prozesse, ohne die Prozesskette zu verlassen.

Andererseits entsteht in der „Delegationstechnik“ eine Anwendungsrichtung, die sich dadurch auszeichnet, dass sich die Software-Werkzeuge durch Anwendungen „selbst verändern“. Durch stetige Datenaufnahme, deren Verarbeitung, neue Datenaufnahme, neue Verarbeitung kann sich das Werkzeug in seiner Zielbestimmung wandeln. Das Werkzeug kann gesetzte Prozessgrenzen überschreiten und kann beginnen, neue Prozessketten zu bilden. Diese „sich selbst verändernde Delegationstechnik“ gründet in der andauernden Datenverarbeitung. Die Fachwelt spricht von „datengetriebenen Systemen“.

„Datengetriebene Systeme“ sind für die Steuerung von Lieferketten, Wertschöpfungsketten und kommenden Verwaltungsprozessen von großer Bedeutung.

Für die Sozialpartner:innen wächst eine zusätzliche Herausforderung: Da es sich bei der „Delegationstechnik“ um „sich selbst verändernde“ Systeme handelt, reicht das bisherige Modell der Mitbestimmung und Gestaltung nicht aus. Es bedarf eines zusätzlichen Aushandlungsformates. Das „Forum Soziale Technikgestaltung“ empfiehlt das ergänzende neue Format „Moderierter Spezifikationsdialog“.

Empfehlungsimpuls 6:

Die Einführung komplexer „Delegationstechnik“ benötigt neue Aushandlungsformate

In den zurückliegenden vierzig Jahren der Erfahrung mit der Einführung digitaler Arbeitswelten haben die Sozialpartner:innen ein bestimmtes Grundmodell der Aushandlung praktiziert. Zwischen den Mitbestimmungsakteuren wurde in der Regel eine Phase des Ausprobierens und des Experimentierens einer neuen Softwarelösung oder eines neuen Organisationsgedankens als erster Schritt der Neuerung vereinbart. Danach wurden die praktischen Nutzungen über einen gewählten Zeitraum ausgewertet (evaluiert). Im Anschluss kam es zwischen den Sozialpartner:innen zur Einigung auf eine Betriebs- oder Dienstvereinbarung mit den passenden Standards zu Themen wie Datenschutz, Arbeitsschutz, Beschäftigungssicherung, Gleichberechtigung, Integration, Gesundheitsschutz, Inklusion, Vielfalt, Qualifizierung, Diskriminierungsfreiheit, Selbstbestimmung etc.



Eine solche Vorgehensweise ist in vielen Arbeitsumgebungen von Produktion, Dienstleistung und Verwaltung üblich. Diese Praxis bezieht sich auf die Welt der „Assistenztechnik“, also auf eine Technik, die sich durch Anwendung nicht verändert.

Doch die Einführung der „Delegationstechnik“ erfordert eine Änderung. Insbesondere jene Software-Werkzeuge, die sich durch Anwendung „selbst verändern“ können, benötigen ein umgebautes Vorgehensmodell.

Die Einfügung von „Delegationstechniken“ in den Arbeits- und Geschäftsalltag erfordert eine veränderte Qualität sozialpartnerschaftlicher Kommunikationen. Ein technik- und sozialinnovatives Vorgehensmodell setzt die Aushandlungen *vor* die Phase der Beschaffung, *vor* das Experimentieren und Erproben sowie *vor* den Technikstart.

Empfehlungsimpuls 7:

„Moderierter Spezifikationsdialog“ – Was ist darunter zu verstehen?

Für die Anpassung der sich „selbst verändernden“ „Delegationstechnik“ wurde vom „Forum Soziale Technikgestaltung“ ein neues Aushandlungsformat entwickelt. Das von Sozialpartner:innen rechtlich vereinbarte Modell kann die Einführung komplexer Technik in der Arbeitswelt erleichtern und Fehlinvestitionen vermeiden helfen.

Das Format ist ein Baustein für sozialpartnerschaftliche Prozesse der Anpassung algorithmischer Steuerungs- und Entscheidungssysteme an die Anforderungen von Arbeits-, Geschäfts- und Verwaltungsvorgängen. Diese Dialogform bildet Ort und Zeitpunkt der Anwendung von ethischen, rechtlichen und sozialen Gestaltungskriterien.

Das Format wird verbindlich vor der Beschaffung und vor der Implementierung von „Delegationstechnik“ als Aushandlungsvorgang zwischen Arbeitgebenden, dem IT-Team und der Beschäftigtenvertretung realisiert. Es wird extern moderiert.

Kennzeichnend für den „Moderierten Spezifikationsdialog“ ist die Gleichgewichtung von arbeitsweltlichem Erfahrungswissen der Mitarbeitenden und dem IT-Fach- und Sachwissen. Zudem tritt an die Stelle eines stark technikorientierten Handelns ein „Denken-In-Prozessen“.

Als Pionierleistung haben die Sozialpartner:innen der Landeshauptstadt Stuttgart – der öffentliche Arbeitgeber und der Gesamtpersonalrat – in ihrer Rahmendienstvereinbarung dieses Format rechtlich formalisiert und verankert. Für diese soziale Innovation erhielten die Stuttgarter Partner:innen den bundesweit renommierten „Personalrätepreis 2022“ in Silber.

Empfehlungsimpuls 8:

Erst Arbeitsabläufe verbessern, dann neue digitale Technik auswählen!

In vielen betrieblichen und arbeitsweltlichen Zusammenhängen wird mit dem Stichwort „Prozessdenken“ lediglich der Vorgang der Technikeinführung gemeint. Eine neue Technik wird wahrgenommen und fasziniert eingekauft. Bei der Implementierung denken



Akteur:innen von den technischen Möglichkeiten her und sind enttäuscht, wenn sich bei den erhofften Anwender:innen Widerspruch äußert.

Dieser Widerspruch rührt nicht selten gar nicht von einer Technikdistanz oder gar von mangelnder Veränderungsbereitschaft. Der Ursprung des Widerspruchs liegt in einem anderen Verständnis von Prozessdenken. Dies zeigt sich gerade bei der Einführung von „Delegationstechnik“. Wer erst an Software denkt und dann an die Verbesserung der Arbeitsabläufe, bereitet häufig eine teure, unbrauchbare Technikruine vor.

Prozessdenken muss deutlich weiter gefasst werden. In einer guten Praxis des Prozessdenkens stehen nicht die Technik und deren Implementierung im Vordergrund, sondern die Analyse und Verbesserung der Arbeits- und Geschäftsprozesse im Sinne der Wertschöpfung (siehe dazu Kapitel 2.3 und Kapitel 4.2).

Das „Lernen-in-Prozessen-zu-denken“ gehört deshalb zu den Schlüsselherausforderungen einer konstruktiven Unternehmens- und Arbeitskultur. Eine Reduzierung des Prozessdenkens auf reine Techniknutzung schwächt die Produktivität und die Wissensbilanz.

Gerade beim Umstieg von „Assistenztechnik“ hin zu Anwendungspotenzialen der „Delegationstechnik“ ist es dringlich, von den Arbeits- und Geschäftsprozessen her zu agieren, um daraus den technischen Veränderungsbedarf abzuleiten. Wer seine Prozesse kennt und die Verbesserungsmöglichkeiten formuliert hat, weiß, welche technische Lösung zum begriffenen Ziel passt. Deshalb die Empfehlung: Erst Arbeitsabläufe verstehen und verbessern, dann digitale Technik auswählen!

Empfehlungsimpuls 9:

Notwendigkeit einer vorausschauenden, antizipierenden Arbeitsgestaltung

Daraus wird in diesem Leitfaden unter Hinweis auf die Diskussionsgrundlagen des „Forum Soziale Technikgestaltung“ der Schluss gezogen, dass die soziale, ethische und rechtliche Gestaltung dieses Typs „datengetriebener“ Software *vor* ihrer Implementierung realisiert werden muss. Das FST spricht deshalb von der Notwendigkeit „vorausschauender, antizipierender Arbeitsgestaltung“. Um diesen neuen, zusätzlichen Weg der Gestaltung operativ umsetzbar zu machen, entwickelte das FST das neue Format des „Modierten Spezifikationsdialoges“ zwischen Beschäftigten und Beschäftigtenvertretungen auf der einen Seite und den IT-Teams auf der anderen Seite (Schröter, 2023).

Die zentrale Aufgabe des „Modierten Spezifikationsdialoges“ besteht darin, das hohe Gut des Erfahrungs- und Prozesswissens der Beschäftigten auf gleicher Augenhöhe mit dem Fach- und Sachwissen der IT-Teams in einen Spezifizierungsprozess zu transformieren. Es geht also nicht primär um technische „Requirements“ sondern um die Einbettung des erfahrungsgeleiteten „Lernens-in-Prozessen-zu-denken“ und um die „Über-Setzung“ von Erfahrungskenntnissen in die Entwicklungswelt.

Dabei ist es nicht die Aufgabe der „Erfahrungs“-Seite, sich in Programmiersprachen und IT-Termini auszudrücken. Das Ziel des Dialoges ist es, dass die IT-Seite die erfahrungsgeleiteten



Anforderungen aus der Arbeitsprozessperspektive heraus verstehen lernt und in die IT-Sprachen übersetzt. Die Aufgabe der „Erfahrungs“-Seite besteht vor allem darin, ihre Anforderungen so präzise wie möglich zu artikulieren. Die bloße Formel einer „sozialen KI“ ist wenig hilfreich. Wenn zum Beispiel eine Abteilung den Arbeitsschutzgedanken unmittelbar im algorithmischen Steuerungs- und Entscheidungssystem softwaretechnisch verankern will, muss die Abteilung ihre arbeitsplatz- und arbeitsprozessbezogenen Anforderungen ganz präzise zur Sprache bringen, um deren Berücksichtigung zu ermöglichen.

Soziale Technikgestaltung erhält hierbei eine ergänzende Dimension, denn die Anforderungen werden nicht nur in Betriebs- oder Dienstvereinbarungen festgelegt, sondern zudem direkt im algorithmischen System verankert. IT-Teams, die diesen Denkansatz verstanden haben, reagieren sehr erfreut, da sie zwar über brillante Kenntnisse in der Technikentwicklung aber selten über Prozesskenntnisse bei den späteren Anwendenden verfügen.

Die späteren Anwendenden im Betrieb bzw. die späteren Kund:innen bestimmen zum frühesten möglichen Zeitpunkt mit. Es macht dabei Sinn, dass die Moderation von externen, unabhängigen Personen durchgeführt wird. Wichtig sind dabei eine aktive Kommunikationsbereitschaft und eine offene Kommunikationskultur, um diesen Weg einer ergänzenden Interessenaushandlung zu gewährleisten.

Das vorgeschlagene Format von „Moderierten Spezifikationsdialogen“, deren Anwendungen über Betriebs- oder Dienstvereinbarungen gestützt sind, hat im Projekt KARL eine interessierte Diskussion ausgelöst. Gesprächsteilnehmende betonten unter anderem die Relevanz folgender Eckpunkte:

- Gegenseitige Verständlichkeit zwischen IT-Expert:innen und Lai:innen sicherstellen (wegen unterschiedlicher Auffassung von Fachbegriffen);
- Wertvolle Erfahrungen der Mitarbeitenden über den Arbeitsprozess wertschätzen und mitnehmen;
- Starke instrumentelle Begründung des Wertes von Spezifikationsdialogen (Probleme vorab vermeiden, Erfahrungswissen nutzen);
- Kommunikation zwischen IT-Entwicklung, Geschäftsleitung und Belegschaft stärken;
- Partizipation von Mitarbeitenden im Design über User Centred Design-Ansätzen unmittelbar am Anfang ermöglichen;
- Stärkere Verbindlichkeit von Partizipation;
- Rechtliche Bedeutung der Mitbestimmung;
- Stakeholder:innenanalyse als wichtiges Instrument der Mitbestimmung;
- Mitbestimmung heißt Verbindlichkeit versus bloßer Beteiligung;
- Chancen der Vermeidung unnötiger Kosten;
- Notwendigkeit kontinuierlicher Reviews.

Grundsätzlich wurde das Format „Moderierter Spezifikationsdialog“ in KARL als ein „wirklich sehr überzeugendes Konzept“ begrüßt. Das Forum Soziale Technikgestaltung (FST) arbeitet an einem dafür hilfreichen Handlungsleitfaden. Dabei macht es Sinn, im Prozess der Software-Entwicklung zu unterscheiden zwischen Standardentwicklung sich selbst nicht verändernder



Software einerseits und komplexen algorithmischen Steuerungs- und Entscheidungssystemen andererseits, die sich durch Datenaufnahme ständig weiter wandeln können.

Aus betrieblichen Erfahrungen ist zu entnehmen, dass nicht selten sowohl Geschäftsleitungen wie auch Beschäftigtenvertretungen beim Umgang mit komplexen sich selbst verändernden Software-Werkzeugen überfordert sind. „Moderierte Spezifikationsdialoge“ können für beide Seiten bei der Kompetenzentwicklung sehr hilfreich sein.

Für die Organisation von diesem Typ „Spezifikationsdialog“ sollten verschiedene Intensivierungsvarianten berücksichtigt werden:

- Dialog Beschäftigtenseite/Beschäftigtenvertretung mit IT-Seite;
- Dialog Beschäftigtenseite/Beschäftigtenvertretung mit IT-Seite und mit Geschäftsleitung;
- Dialog Beschäftigtenseite/Beschäftigtenvertretung mit IT-Seite und mit Sachverständigen; (Datenschutz, Arbeitsschutz, Gesundheitsschutz, Gefährdungsbeurteilung, Schwerbehindertenvertretung, Weiterbildung, ...);
- Gemeinsame Austausch- und Kommunikationsräume schaffen;
- Prinzipien der Agilität (z.B. SCRUM) als Methode nutzen,
- Jeweils mit externer, „neutraler“ Moderation.

Die Spezifikationsdialoge sollten dazu beitragen, in einem interessengruppenübergreifenden Aushandlungsvorgang eine gemeinsam vereinbarte Liste von Gestaltungs- und Zulassungskriterien zu erarbeiten, nach denen komplexe Software-Systeme gehandhabt werden können. Die Spezifikation muss kontextbezogen erfolgen. Das „Forum Soziale Technikgestaltung“ hat dazu vierzig generische Kriterien vorbereitet, veröffentlicht und in Diskussionsprozesse eingebracht. Diese Kriterienliste kann als einer der Ausgangspunkte für weitere Beratungen dienen. Die Sammlung von Gestaltungs- und Zulassungskriterien konnten aus den Erfahrungen von Belegschaften und Beschäftigtenvertretungen abgeleitet werden. Diese Sammlung gilt es fortzuentwickeln.

Die aus dem FST entwickelten „Spezifikationsdialoge“ wurden durch eine rechtsgültige Rahmendienstvereinbarung zwischen dem Gesamtpersonalrat und dem öffentlichen Arbeitgeber der Landeshauptstadt Stuttgart bekräftigt. In Anlage 6 der Rahmen-DV wurde von den Sozialpartner:innen unter anderem festgelegt:

In den Spezifikationsdialogen werden verbindlich Anforderungen an die Gestaltung von „vertrauenswürdiger Künstlicher Intelligenz“ formuliert. Es soll sichergestellt werden, dass:

- diese Systeme auf ihre Anwendung als Assistenztechnik beschränkt bleiben und nicht selbst rechtsverbindliche Entscheidungen treffen,
- die Nutzung dieser Systeme eine Verbesserung der Arbeitsbedingungen und der Arbeitsqualität gemäß sozialer Standards der Rahmendienstvereinbarung ermöglicht.

Aus den gemeinsamen Erfahrungen der Spezifikationsdialoge wird eine Anforderungs-Checkliste für die Gestaltung und Zulassung von KI-Systemen erstellt, um die Einführung zu erleichtern. Jeweils an einem konkreten Fall/Beispiel soll Klarheit geschaffen und Anforderungen am konkreten Vorhaben entwickelt werden. Für jedes Vorhaben wird



festgelegt, ob eine befristete Experimentierphase zur Prüfung der Software durchgeführt werden soll.

Vor dem Beschaffungsvorgang: Sollen komplexe Software-Systeme beschafft werden, deren algorithmische Vorgaben durch die Landeshauptstadt Stuttgart nicht verändert und angepasst werden können, findet vor dem Beschaffungsvorgang ein Spezifikationsdialog mit dem Gesamtpersonalrat statt, bei dem dieser seine Anforderungen an die Software und deren Leistungsbreite artikulieren kann. Die Anforderungen werden verbindlich geprüft und einvernehmlich als Anforderung in den Beschaffungsvorgang eingefügt.

Vor dem Implementierungsvorgang: Sollen komplexe Software-Systeme implementiert werden, deren algorithmische Vorgaben durch die Landeshauptstadt Stuttgart verändert und angepasst werden können, findet vor dem Implementierungsvorgang ein Spezifikationsdialog mit dem Gesamtpersonalrat statt, bei dem dieser seine Anforderungen an die Software und deren Leistungsbreite artikulieren kann. Die Anforderungen werden verbindlich geprüft und einvernehmlich als Anforderung in den Beschaffungsvorgang eingefügt. Gegebenenfalls wird in einem begrenzten Experimentierfeld die Vorgaben getestet und überprüft.

Die praktische Lösung der Landeshauptstadt Stuttgart ist vom Grundsatz her auf andere Sektoren und Branchen übertragbar (siehe dazu auch den Lesehinweis „Deutsche Normungsroadmap Künstliche Intelligenz Ausgabe 2“, Wahlster und Winterhalter, 2022).



Hilfsmittel für Sozialpartner:innen in der betrieblichen Praxis:

- Handlungsempfehlungen an beide Sozialpartner:innen. Siehe: <https://kompetenzzentrum-karl.de/karl-angebot/kurzimpulse-fur-sozialpartner/> oder <http://www.blog-zukunft-der-arbeit.de/handlungsempfehlungen-fuer-sozialpartner/>
- „Rahmendienstvereinbarung zur Digitalisierung und Informationstechnik bei der Landeshauptstadt Stuttgart“ vom 25. Mai 2022 mit acht gültigen Anlagen http://www.blog-zukunft-der-arbeit.de/wp-content/uploads/2022/07/GPRLHS_Rahmen-DV_2022_Gesamtfassung_mit_Anlagen.pdf
- Schröter, W. (2023), Der mitbestimmte Algorithmus – Impulse für eine Ergänzung bisheriger Mitbestimmungspraxis. Aus den Erfahrungen des Projektes PROTIS-BIT und des „Forum Soziale Technikgestaltung“ (FST). Onlinepublikation für das „Mitbestimmungsportal“ der Hans-Böckler-Stiftung (www.boeckler.de) (im Erscheinen).
- Der „Moderierte Spezifikationsdialog“ https://kompetenzzentrum-karl.de/wp-content/uploads/2022/12/KARL_Periodika_4.pdf
- Factsheet der Offensive Mittelstand: KI erkennen, bewerten und einführen – Kriterien der Arbeitsgestaltung https://www.offensive-mittelstand.de/fileadmin/KI-Factsheet/KI_Kriterien_factsheet_2022_12_1_web.pdf
- Forum Soziale Technikgestaltung (FST): Mitbestimmungsplan und Checkliste für Betriebs- und Personalräte <http://www.blog-zukunft-der-arbeit.de/mitbestimmungsplan-und-checkliste-zur-gestaltung-sogenannter-kuenstlicher-intelligenz/>
- Foliensatz „Vierzig FST-Kriterien zur Gestaltung der sogenannten ‚KI‘“. Siehe: <http://www.blog-zukunft-der-arbeit.de/vierzig-fst-kriterien-zur-gestaltung-der-sogenannten-ki/>
- Für fachlich interessierte Expert:innen: „Deutsche Normungsroadmap Künstliche Intelligenz Ausgabe 2“ <http://www.din.de/go/normungsroadmapki>



6. Forschungs- und Gestaltungsdesiderate

Die in diesem Leitfaden dargestellten Vorgehensweisen sollen im Rahmen der Use Cases von KARL in der Praxis überprüft und erprobt werden. Damit einher geht die Generierung weiterer empirischer Einblicke über die Form und Folgen der Einführung von KI im betrieblichen Kontext. Neben der bereits angezeigten Notwendigkeit der iterativen Weiterentwicklung der beschriebenen Vorgehensmodelle und Prozesse, verlangt eine umfassende Befassung mit den sozialen und ethischen Implikationen des Einsatzes von KI danach, weitergehende **gesellschaftspolitische und ethische Fragen** zu stellen. Diese umfassen im weitesten Sinne Forschungsfragen nach der normativen Ausrichtung und Gestaltung unserer zukünftigen Arbeitswelt:

- **KI als kontinuierliche Weiterentwicklung der Automatisierung von Arbeitsprozessen und Arbeitsorganisation. Forschungsfragen:** Wie gestalten sich Mensch-Maschine-Interaktionen in KI-gestützten Arbeitsumgebungen? Wer sind die Gewinner:innen, wer die Verlierer:innen dieser Automatisierungsprozesse? Können integrative Modelle entwickelt werden? Welchen Zweck erfüllen die Automatisierungsprozesse im Hinblick auf gesellschaftliche Transformationserwartungen im Hinblick auf einen sozial-ökologischen Umbau? Wie könnte Automatisierung genutzt werden, um eine Humanisierung der Arbeit zu ermöglichen?
- **Institutionelle und organisationale Gestaltung des Einsatzes von KI. Forschungsfragen:** Wer entscheidet über die Zwecke bzw. Ziele, die mit dieser neuen Technisierungswelle verbunden werden? Bleibt die Verfügungsmacht über Investitionsentscheidungen und die Verteilung der ökonomischen Gewinne bei einer kleinen Zahl wirtschaftlicher Entscheidungsträger:innen konzentriert – oder gelingt ein demokratischer Umbau von Arbeits- und Produktionsformen im Interesse gerechter Arbeitsbeziehungen? Wie ließen sich die Früchte technologischer Entwicklung aber auch jenseits der betrieblichen Ebene inklusiv verteilen?
- **Demokratisierung zukünftiger KI-Gestaltung in Arbeitsumgebungen. Forschungsfragen:** Wie können Formen erweiterter Mitbestimmung aussehen, die betrieblichen Interessenvertretungen eine aktive, gestaltende Rolle im Innovationsgeschehen sichern? Wie können auch jenseits klassischer Mitbestimmungsmechanismen sogenannte einfache Beschäftigte an der Gestaltung ihrer zukünftigen Arbeitsmittel partizipieren? Und wie lässt sich eine inklusive Gestaltung von KI auch in Betrieben absichern, in denen es keine betrieblichen Interessenvertretungen gibt?

Die Demokratisierung der Entscheidung über die Ausgestaltung der Digitalisierung, insbesondere der Einführung von KI, im betrieblichen wie auch im breiteren gesellschaftlichen Kontext stellt dabei in praktischer wie normativer Hinsicht eine Voraussetzung gelingender Innovation dar. In Anlehnung an einen der ersten großen Pioniere im Feld der KI, Joseph Weizenbaum (Weizenbaum, 1976), gilt dabei: Ein souveräner Umgang mit Technologie schließt dabei auch explizit die Möglichkeit ein, technologische Potentiale ungenutzt zu lassen. Dies kann beispielsweise sinnvoll sein, wenn Effizienzsteigerungen technologisch nur auf Kosten prioritärer normativer Werte wie der Würde des Menschen zu gewinnen wären, etwa im Pflegesetting. Gerade die Pfadabhängigkeiten, die die Einführung technischer Systeme



schaftt und die potentielle Eingriffstiefe der KI, verlangt nach einem Innehalten zur antizipierenden ethischen Reflexion.

Gleichzeitig ist es notwendig, dass die Gesellschaft die potenziellen Folgen von KI sachlich aber wertorientiert diskutiert, hierzu bedarf es mehr guter populärwissenschaftlicher Materialien (wie z.B. das im Erscheinen befindliche Buch Zweig, 2023¹²).

Aus **Sicht der Sozialpartnerschaften** und insbesondere aus den Praxiserfahrungen des Forum Soziale Technikgestaltung ergeben sich für die nahe Zukunft gesellschaftswissenschaftlicher Forschung und der Arbeitsforschung drei wesentliche Anforderungen:

- Wir benötigen neue gesellschaftswissenschaftliche Analysen zum Wandel des impliziten und expliziten Erfahrungswissens.
- Die Verschiebungen der Handlungsträgerschaften innerhalb der Arbeitswelten bedürfen der intensiveren Erforschung.
- Es gilt die Potenziale der „Spezifikationsdialoge“ zu untersuchen und handlungsorientiert aufzubereiten.

Diese Anforderungen des Forum Soziale Technikgestaltung wurden in die Abschlusskonferenz des BMBF-Forschungszusammenhangs „KI – Mensch – Gesellschaft“ (KIMeGe) am 14./15. November 2022 in München eingebracht¹³.

Auch die Berücksichtigung der **rechtlichen Aspekte** im Zusammenhang mit dem Einsatz von Künstlicher Intelligenz (KI) in den Bereichen Arbeit und Bildung hat wesentliche Auswirkungen auf die Forschungs- und Entwicklungstätigkeiten:

Forschende und Entwickelnde müssen sicherstellen, dass ihre Aktivitäten den geltenden rechtlichen Bestimmungen entsprechen. Dies könnte bedeuten, dass die Prozesse in der Praxis angepasst werden müssen, um den rechtlichen Rahmenbedingungen gerecht zu werden. Die Entwicklung von KI-Systemen müsste somit nicht nur technologische Innovationen berücksichtigen, sondern auch ethische und rechtliche Anforderungen von Anfang an einbeziehen. Darüber hinaus könnte die Forschung dazu beitragen, klare rechtliche Leitlinien für den verantwortungsbewussten Einsatz von KI zu definieren.

Eine bedeutende Aufgabe könnte darin bestehen, umfassende rechtliche Analysen durchzuführen, um zu ermitteln, wie bestehende Gesetze und Vorschriften auf den Einsatz von KI in diesen Bereichen anwendbar sind. Hieraus könnten Empfehlungen für notwendige Anpassungen oder sogar neue Rechtsvorschriften resultieren, um den spezifischen Herausforderungen des KI-Einsatzes gerecht zu werden.

Ein weiteres Forschungsgebiet könnte darin liegen, potenzielle Haftungsrisiken zu identifizieren, die mit dem Einsatz von KI in Arbeit und Bildung einhergehen könnten. Die Forschung könnte zur Entwicklung von Richtlinien und Methoden beitragen, die dazu dienen, diese Risiken zu minimieren und verantwortungsvolle Entscheidungen bei der Nutzung von KI-

¹² Die Autor:innen des vorliegenden KARL GF5 Whitepaper haben die Informationen über die bevorstehende Publikation von Katharina Zweig von der Webseite <https://www.penguin.de/ebook/Die-KI-wars/Katharina-Zweig/Heyne/e618108.rhd>

¹³ Videobeitrag (Screencast) von Welf Schröter hierzu <https://youtu.be/OYRKBx7AJr0>



Technologien zu fördern. Die Einhaltung des Datenschutzes ist ebenfalls von entscheidender Bedeutung, insbesondere wenn KI-Systeme personenbezogene Daten verarbeiten. Die Forschung könnte dazu beitragen, die Auswirkungen von KI auf den Datenschutz zu bewerten und Mechanismen zu entwickeln, um sicherzustellen, dass Daten angemessen geschützt werden.

Durch die Bereitstellung von Bildungsmaterialien, Schulungen und Ressourcen könnten Forscher:innen dazu beitragen, Fachleute im Bereich Arbeit und Bildung über die relevanten rechtlichen Aspekte von KI aufzuklären. Dies könnte zur Sensibilisierung beitragen und sicherstellen, dass relevante Akteur:innen über das notwendige Wissen verfügen, um verantwortungsvolle Entscheidungen zu treffen.

Schließlich könnte die Forschung zur Harmonisierung internationaler rechtlicher Standards im Zusammenhang mit KI in Arbeit und Bildung beitragen. Dies könnte dazu beitragen, einheitliche Regelungen zu schaffen und sicherzustellen, dass der KI-Einsatz global kohärent und ethisch verantwortlich geregelt wird.

Zusammenfassend könnte und sollte die Einbindung rechtlicher, sozialer und ethischer Dimensionen in die Forschung zu KI und in die Gestaltung von KI in den Bereichen Arbeit und Bildung zu einer verantwortungsvolleren und nachhaltigeren Entwicklung und Nutzung von KI-Technologien führen.



7. Quellenverzeichnis

- Adamo, G., Ghidini, C. und Di Francescomarino, C. (2021): What is a process model composed of? In Software and Systems Modeling 20 (4), S. 1215–1243. <https://doi.org/10.1007/s10270-020-00847-w>
- Adams-Prassl, J. (2022): Regulating algorithms at work: Lessons for a “European approach to artificial intelligence”, European Labour Law Journal, 13(1), S. 30–50. <https://doi.org/10.1177/20319525211062558>
- Alpar, P., Alt, R., Bensberg, F. und Czarnecki, C. (2023): Phasenmodelle in der Softwareentwicklung. In: Alpar, P. et al. (Hrsg.): Anwendungsorientierte Wirtschaftsinformatik: Strategische Planung, Entwicklung und Nutzung von Informationssystemen. Wiesbaden: Springer Fachmedien, S. 389–448. https://doi.org/10.1007/978-3-658-40352-2_13
- Alpers, S. (2022); Notwendigkeit der Integration von ethischen, rechtlichen und sozialen Aspekten in die gängigen Vorgehensmodelle für IT-Projekte, In Fazal-Baqaie, M., et al. (Hrsg.) Projektmanagement und Vorgehensmodelle 2022 - Virtuelle Zusammenarbeit und verlorene Kulturen? Trier. <https://dl.gi.de/handle/20.500.12116/39672>
- Alpers, S. und Krings, B.-J. (2023): Ethische, rechtliche und soziale Aspekte (ELSA) der Gestaltung von KI-Systemen und Integration der verschiedenen Akteure in den Prozess. In Gesellschaft für Arbeitswissenschaft (Hrsg.): 69. GfA Frühjahrskongress. Nachhaltig Arbeiten und Lernen – Analyse und Gestaltung lernförderlicher und nachhaltiger Arbeitssysteme und Arbeits- und Lernprozesse. <https://doi.org/10.5445/IR/1000156501>
- Auer-Reinsdorff, A. und Conrad, I. (2019): Handbuch IT- und Datenschutzrecht. 3. Auflage. München: C.H. Beck.
- Auer-Reinsdorff, A. (2023): Digitales Kaufrecht B2B. MMR Zeitschrift für IT-Recht und Recht der Digitalisierung, 2023, S. 6-10
- Bao, M., Vugrincic, A., Dreher, A., Vonderau, D., Tran, H., Rill, M. und Balaban, S. (2023): Datenschutz bei Künstlicher Intelligenz. Whitepaper August 2023, KITOpen, <https://doi.org/10.5445/IR/1000161675>
- Beierle, B. (2021): Die Produkthaftung im Zeitalter des Internet of Things. Nomos Verlag. <https://doi.org/10.5771/9783748924708>
- Beyerer, J., Müller-Quade, J. und Reussner, R. (2018): Karlsruher Thesen zur Digitalen Souveränität Europas, Datenschutz und Datensicherheit, 42(5), S. 277–280. <https://doi.org/10.1007/s11623-018-0940-2>
- Bijker, W. und Pinch, T. (1987): The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit of Each Other. In: Bijker, W. et al. (Hrsg.): The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology. MIT Press, Cambridge 1987, S. 17–50.
- Bongers, F., Byers, P., Eylert, M., Fuhlrott, M., Geißler, K., Hoppe, C., Kramer, S., Kramer, S., Neu, J., Oberthür, N., Petri, D., Raif, A., Schulze Zumkley, K., Solmecke, C., Straten, A., Tiedemann, J., Törkel, S. und Wenzel, K. (Hrsg.). (2023). IT-Arbeitsrecht: Handbuch Digitalisierung, Homeoffice, KI, Virtuelle Betriebsratsarbeit. C.H.Beck.
- Bosch, G. (2017): Berufliche Bildung. In: Hirsch-Kreinsen, H.; Minssen, H. (Hrsg.) (2017): Lexikon der Arbeits- und Industriosozologie. 2. überarb. Auflage Baden-Baden: Nomos, S. 82-85. <https://doi.org/10.5771/9783845276021-82>
- Brinkkemper, S. (1996): Method engineering: engineering of information systems development methods and tools. In: Information and Software Technology 38 (4), S. 275–280. [https://doi.org/10.1016/0950-5849\(95\)01059-9](https://doi.org/10.1016/0950-5849(95)01059-9)



- Byers, P., Winkler, M., Stelter, S. (2023): Zulässigkeit von biometrischen Kontrollen am Arbeitsplatz. Neue Zeitschrift für Arbeitsrecht (NZA), 2023, S. 457-464.
- Curedale, R. (2019): Design thinking: process & methods. 5th edition. Los Angeles, CA: Design Community College
- Dalton, J. (2019): Definition of Ready Great Big Agile: An OS for Agile Leaders, S. 163–164. <https://doi.org/10.1007/978-1-4842-4206-3>
- Düwell, F.J. und Brink, S. (2017): Beschäftigtendatenschutz nach der Umsetzung der Datenschutz-Grundverordnung: Viele Änderungen und wenig neues, Neue Zeitschrift für Arbeitsrecht (NZA), 2017, S. 1081–1085.
- Dzida, B. (2017): Big Data und Arbeitsrecht, Neue Zeitschrift für Arbeitsrecht (NZA), 9, S. 541–546.
- Ebers, M., Heinze, C., Krügel, T. und Steinrötter, B. (2020): Künstliche Intelligenz und Robotik. Verlag CH BECK Recht-Wirtschaft-Steuern. <https://doi.org/10.17104/9783406769818>
- Epping, T. (2011): Begriffsklärungen. Kanban für die Softwareentwicklung, S. 13–21. <https://doi.org/10.1007/978-3-642-22595-6>
- Fiedeler, U. und Nentwich, M. (2009): Begleitforschung: Zur Klärung eines politischen Begriffs. TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis, 18(2), S. 94–102.
- Fischer, M.; Krings, B.-J.; Moniz, A.; Zimpelmann, E. (2017): Herausforderungen der Mensch-Roboter-Kollaboration. Lernen & Lehren, 32(125), S. 8–14. http://lernenundlehren.de/heft_dl/Heft_125.pdf
- Fischer, T., Biskup, H. und Müller-Luschnat, G. (1998): Begriffliche Grundlagen für Vorgehensmodelle, Vorgehensmodelle für die betriebliche Anwendungsentwicklung, S. 13–31. https://doi.org/10.1007/978-3-663-05994-3_1
- Flecker, J. (2018): TINA und die technologische Revolution. sozialpolitik.ch, <https://doi.org/10.18753/2297-8224-101>
- Franken, S. (2019): Unternehmen, Unternehmenskultur und Unternehmensethik. In Franken, S., Verhaltensorientierte Führung. Springer Fachmedien Wiesbaden, S. 191–234. https://doi.org/10.1007/978-3-658-25270-0_7
- Freund, J. und Rücker, B. (2019): Praxishandbuch BPMN. 6., aktualisierte Auflage. Carl Hanser Verlag München. <https://doi.org/10.3139/9783446461123>
- Frey, C. und Osborne, M. (2017): The future of employment: How susceptible are jobs to computerisation? Technological forecasting and social change, S. 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Gerlsbeck, F. und Herzog, L. (2020): The epistemic potentials of workplace democracy. Review of Social Economy, 78:3, 307-330, <https://doi.org/10.1080/00346764.2019.1596299>
- Giermendl, L., Strich, F. Christ, O., Leicht-Deobald, U. und Redzepi, A. (2022): The dark sides of people analytics: reviewing the perils for organisations and employees. European Journal of Information Systems, 31(3), S. 410–435. <https://doi.org/10.1080/0960085X.2021.1927213>
- Gola, P. und Klug, C. (2023): Die Entwicklung des Datenschutzrechts. Neue Juristische Wochenschrift NJW, 2023, S. 658-663.
- Grambow, T. (2021): Das Betriebsrätemodernisierungsgesetz im Überblick. Neue Juristische Wochenschrift NJW, 2021, S. 2074 - 2080.



- Grunwald A. und Hillerbrand, R. (2021): Handbuch Technikethik. 2., aktualisierte und erweiterte Auflage. Berlin: J.B. Metzler Verlag. <https://doi.org/10.1007/978-3-476-04901-8>
- Hanser, E. (2010): Scrum. Agile Prozesse: Von XP über Scrum bis MAP, S. 61–77. https://doi.org/10.1007/978-3-642-12313-9_5
- Hirsch-Kreinsen, H. (2023): Industrie 4.0. In Bohn, R. et al. (Hrsg.) Lexikon der Arbeits- und Industriesoziologie. 3. überarb. Auflage Baden-Baden: Nomos. <https://doi.org/10.5771/9783845276021-170>
- Hoeren, T., Sieber, U., & Holznagel, B. (2021): Handbuch Multimedia-Recht: Rechtsfragen des elektronischen Geschäftsverkehrs. Handbuch 57. Auflage. C.H. Beck.
- Holthausen, J. (2021): Big data, people analytics, KI und Gestaltung von Betriebsvereinbarungen. Grund-, arbeits- und datenschutzrechtliche An- und Herausforderungen. Recht der Arbeit, 74(1), S. 19–32.
- Huff, J. und Götz, T. (2019): Evidenz statt Bauchgefühl? Möglichkeiten und rechtliche Grenzen von Big Data im HR-Bereich. Neue Zeitschrift für Arbeitsrecht, 36(24), S. 73–78.
- Jarrahi, M., Newlands, G., Lee, M., Wolf, C., Kinder, E. und Sutherland W. (2021): Algorithmic management in a work context. Big Data & Society, 8(2). <https://doi.org/10.1177/20539517211020332>
- Jiang, Y., Li, X., Luo, H., Yin, S. und Kaynak, O. (2022): Quo vadis artificial intelligence? Discover Artificial Intelligence, 2(1), 4. <https://doi.org/10.1007/s44163-022-00022-8>
- Johannsen, F., Mang, F., Fill, H. und Hofmann, Sara (2023): Tool-based Codification in Business Process Improvement and the Impact on Problem-Solving. Enterprise Modelling and Information Systems Architectures (EMISA), Vol. 18. <https://doi.org/10.18417/emisa.18.1>
- Jonas, H. (2020): Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation. Berlin: Suhrkamp. https://doi.org/10.1007/978-3-476-05728-0_9657-1
- Joos, D. (2020): Einsatz von künstlicher Intelligenz im Personalwesen unter Beachtung der DS-GVO und des BDSG. Neue Zeitschrift für Arbeitsrecht, 2020, S. 1216 - 1222.
- Klöpfer, M. (2023): Every Break You Take, Every Click You Make-Empirical Insights on Employee's Perception of People Analytics. ECIS 2023 Research Papers. https://aisel.aisnet.org/ecis2023_rp/356
- Klöpfer, M. und Köhne, S. (2023): Shifting Structures-A Systematic Literature Review on People Analytics and the Future of Work. ECIS 2023 Research Papers. https://aisel.aisnet.org/ecis2023_rp/360
- Kneis, L, Rill, M., Alpers, S., Friedewald, M., Beckert, B., Brückmayr, J., Denninger, O., Krauß, K., Metzger, F. und Wins, A. (2023): Anonymität und Mobilität - Whitepaper zum Begriffs- und Domänenverständnis des Kompetenzcluster ANYMOS – Anonymisierung für vernetzte Mobilitätssysteme. August 2023, KITOpen, <https://doi.org/10.5445/IR/1000161584>
- Kort, Michael (2018): Die Bedeutung der neueren arbeitsrechtlichen Rechtsprechung für das Verständnis des neuen Beschäftigtendatenschutzes, Neue Zeitschrift für Arbeitsrecht, 2018, S. 1097 - 1106.
- Krings, B.-J. (2011) Brain drain or brain gain?: changes of work in knowledge-based societies. edition sigma.
- Krings, B.-J., Moniz, A. B., Frey, P. (2021): Technology as enabler of the automation of work? Current societal challenges for a future perspective of work. Revista Brasileira de Sociologia, 9 (21), S. 206–229. <http://doi.org/10.20336/rbs.806>
- Kühling, J., Buchner, B. (Hrsg.) (2020): Datenschutz-Grundverordnung, BDSG: Kommentar (3. Auflage). C.H. Beck.



- Kuhrmann, M.; Linssen, O. (2014): Welche Vorgehensmodelle nutzt Deutschland? In: Engstler, M. et al. (Hrsg.): Projektmanagement und Vorgehensmodelle 2014. Soziale Aspekte und Standardisierung. Köllen. <https://dl.gi.de/items/24829cd2-551b-4e70-b985-318a2c0b079b>
- Loi, M. (2021): People Analytics muss den Menschen zugutekommen. Eine ethische Analyse datengesteuerter algorithmischer Systeme im Personalmanagement. Hans Böckler Stiftung https://www.boeckler.de/fpdf/HBS-007994/p_study_hbs_450.pdf
- Marjanovic, O., Cecez-Kecmanovic, D. und Vidgen, R. (2021): Algorithmic pollution: Making the invisible visible. Journal of Information Technology, 36(4), S. 391–408. <https://doi.org/10.1177/02683962211010356>
- Maschmann, F. (2018): Führung und Mitarbeiterkontrolle nach neuem Datenschutzrecht, Neue Zeitschrift für Arbeitsrecht NZA-Beilage, 2018, S. 115 - 125.
- McCarthy, J., Minsky, M. L., Rochester, N. und Shannon, C. E. (2006): A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-12. <https://doi.org/10.1609/aimag.v27i4.1904>
- Moldaschl, M. (2010): Kapitel IV Rationalisierung von Arbeit: Organisation und Organisation von Arbeit', in F. Böhl et al. (Hrsg.) Handbuch Arbeitssoziologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 263–299.
- Molnar, C. (2022). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (2nd ed.) https://doi.org/10.1007/978-3-531-92247-8_9
- Müller, A. (2023) Forschungsmonitoring Arbeit der Zukunft. Ausgabe 21: Januar bis März 2023 (Berichtszeitraum). Working paper Forschungsförderung. https://www.boeckler.de/fpdf/HBS-008663/p_fofoe_WP_299_2023.pdf
- Müller-Glöße, R., Preis, U., Schmidt, I., Dieterich, T., Hanau, P., Schaub, G. und Franzen, M. (2023): Erfurter Kommentar zum Arbeitsrecht (23., neu bearbeitete Auflage). C.H. Beck.
- Münchener Kommentar zum Bürgerlichen Gesetzbuch Bd. 3: Schuldrecht -Allgemeiner Teil II §§ 311 - 432, (9. Auflage). (2022). C.H.Beck.
- Münchener Kommentar zum Bürgerlichen Gesetzbuch Bd. 7: Schuldrecht -Besonderer Teil IV §§ 705-853, Partnerschaftsgesellschaftsgesetz, Produkthaftungsgesetz (9. Auflage). (2022). C.H.Beck.
- Nachtwey, O. und Seidl, T. (2017): Die Ethik der Solution und der Geist des digitalen Kapitalismus. IFS Working Paper #11 Institut für Sozialforschung an der Johann Wolfgang Goethe-Universität.
- Oehlke, P. (2017): Humanisierung des Arbeitslebens. In: Hirsch-Kreinsen, H. et al. (Hrsg.) (2017): Lexikon der Arbeits- und Industriesoziologie. 2. überarb. Auflage Baden-Baden: Nomos, S. 167-169. <https://doi.org/10.5771/9783845273099-263>
- Pfeiffer, S. (2015): Industrie 4.0 und die Digitalisierung der Produktion–Hype oder Megatrend. Aus Politik und Zeitgeschichte, 65(31/32), S. 6–12.
- Piore, M. und Sabel, C. (1984) The second industrial divide: possibilities for prosperity. New York: Basic Books. <https://doi.org/10.2307/1956301>
- Raif, A. und Swidersky, M. (2017): Arbeit 4.0 – Typische Fehler in der digitalen Arbeitswelt vermeiden, Gesellschafts- und Wirtschaftsrecht, 2017, S. 351 - 355.
- Reisig, W. (2010): Petrinetze. Modellierungstechnik, Analysemethoden, Fallstudien. Vieweg+Teubner Verlag, Wiesbaden. <https://doi.org/10.1007/978-3-8348-9708-4>



Rohde, F., Wagner, J. Reinhard, P., Petschow, U., Meyer, A., Voß, M. und Mollen, A. (2021): Nachhaltigkeitskriterien für künstliche Intelligenz, Schriftenreihe des IÖW. <https://nbn-resolving.de/urn:nbn:de:kobv:109-1-15457517>

Rosenthal, K., Ternes, B., Strecker, S. (2021): Business Process Simulation on Procedural Graphical Process Models. In: Business & Information Systems Engineering 63 (5), S. 569–602. <https://doi.org/10.1007/s12599-021-00690-3>

Rudkowski, L. (2019): „Predictive policing “am Arbeitsplatz, Neue Zeitschrift für Arbeitsrecht (NZA), S. 72 - 78.

Russell, S. und Norvig, P. (2016): Artificial Intelligence: a modern approach. 3 Auflage, Englewood Cliffs, N. J. Prentice Hall.

Schafheitle, S., Weibel, A. und Rickert, A. (2020): The Bermuda Triangle of Leadership in the AI Era? Emerging Trust Implications From ‘in Two-Leader-Situations’ in the Eyes of Employees.” In: Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS 2021). <https://doi.org/10.24251/HICSS.2021.665>

Schrader, P. (2022): Die neue vertragliche Haftung für „smarte Produkte“, Juristische Arbeitsblätter JA, 2022, S. 1-10.

Schröter, W. (2019): Der mitbestimmte Algorithmus. Arbeitsweltliche Kriterien zur sozialen Gestaltung von Algorithmen und algorithmischen Entscheidungssystemen. In: W. Schröter (Hrsg.): Der mitbestimmte Algorithmus. Gestaltungskompetenz für den Wandel der Arbeit. Talheimer Verlag Mössingen, S. 101–150.

Schröter, W. (2022): Gestaltungskompetenz für Betriebsräte im Spannungsverhältnis von Assistenz- und Delegationstechnik. Erfahrungen aus dreißig Jahren „Forum Soziale Technikgestaltung“. In: E. Gruber et al. (Hrsg.): Magazin erwachsenenbildung.at – Das Fachmedium für Forschung, Praxis und Diskurs. Ausgabe 44-45, 2022. Digitalisierung und Erwachsenenbildung. Reflexionen zu Innovation und Kritik. Wien 2022, S. 15-0 bis 15-8. https://erwachsenenbildung.at/magazin/22-44u45/15_schroeter.pdf

Schröter, W. (2023a): Vierzig FST-Kriterien zur Gestaltung der sogenannten „KI“. Blogbeitrag vom 11.05.2023. <http://www.blog-zukunft-der-arbeit.de/vierzig-fst-kriterien-zur-gestaltung-der-sogenannten-ki/>, zuletzt abgerufen 28.08.2023

Schröter, W. (2023b): Zehn KARL-kurzimpulse (Videos) – Praxistipps in Alltagssprache zur erfolgreichen Einführung komplexer Technik – Vorschläge an die Sozialpartner. <https://kompetenzzentrum-karl.de/karl-angebot/kurzimpulse-fur-sozialpartner/>, zuletzt abgerufen 28.08.2023 oder alternativ unter <http://www.blog-zukunft-der-arbeit.de/handlungsempfehlungen-fuer-sozialpartner/> (Blogbeitrag vom 31.07.2023), zuletzt abgerufen 28.08.2023

Schütte, R., Seufert, S. und Wulfert, T. (2019): Das Wertbeitragscontrolling als Anreicherung bestehender Vorgehensmodelle des Software Engineering. In Linssen, O. et al. (Hrsg.) Projektmanagement und Vorgehensmodelle 2019. Neue Vorgehensmodelle in Projekten – Führung, Kulturen und Infrastrukturen im Wandel. Lörrach. <https://dl.gi.de/handle/20.500.12116/34858>

Sesing, A. und Tschech, A. (2022): AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, Einschätzung aus der Perspektive des (Anti-)Diskriminierungsrechts. Zeitschrift für IT-Recht und Recht der Digitalisierung, 2022, S. 24 - 31.

Silva, A., Araújo, T., Nunes, J., Perkusich, M., Dilenzo, E., Almeida, H. und Perkusich, A. (2017): A systematic review on the use of Definition of Done on agile software development projects. In Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering, S. 364–373. <https://doi.org/10.1145/3084226.3084262>



Stachowiak, H. (1973): Allgemeine Modelltheorie. Springer, Wien, New York.

Stehr, N. (2001): Wissen und Wirtschaften: die gesellschaftlichen Grundlagen der modernen Ökonomie. <https://doi.org/10.1007/s11577-002-0050-9>

Turing, A. (1948): Intelligent machinery. In Copeland, B. (Hrsg.) (2004): The Essential Turing. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198250791.003.0016>

Tursunbayeva, A., Pagliari, C., Di Laura, S. und Antonelli, G. (2022): The ethics of people analytics: risks, opportunities and recommendations. Personnel Review, 51(3), S. 900–921. <https://doi.org/10.1108/pr-12-2019-0680>

von Schomberg, R. (2013): A Vision of Responsible Research and Innovation. In Owen R. (Hrsg.): Responsible Innovation. Chichester, UK: John Wiley & Sons, Ltd, S.51–74. <https://doi.org/10.1002/9781118551424.ch3>

Wahlster, W. und Winterhalter, C. (Hrsg.) (2022): Deutsche Normungsroadmap Künstliche Intelligenz. Ausgabe 2. DIN und DKE. <http://www.din.de/go/normungsroadmapki>, zuletzt abgerufen 28.08.2023

Weinreuter, M., Alpers, S. und Oberweis, A. (2022): Method for Eliciting Requirements in the area of Digital Sovereignty (MERDigS). In: Yang, X. et al. (Hrsg.): Proceedings of Eighth International Congress on Information and Communication Technology (ICICT) London. https://doi.org/10.1007/978-981-99-3243-6_76

Weizenbaum, J (1976): Computer Power and Human Reason: From Judgment to Calculation. W. H. Freeman, New York, San Francisco.

Witteler, M. (2023): ChatGPT – ein Thema für den Betriebsrat? Rechte des Betriebsrat bei der Einführung oder dem Verbot von ChatGPT. Zeitschrift für Datenschutz, 2023, S. 377 - 383.

Wolff, H., Brink, S., von Ungern-Sternberg, A. (2023): Beck'scher Online Kommentar Datenschutzrecht, 44. Edition, Stand 01.05.2023., C.H. Beck.

Zweig, K. (2023): Die KI war's! Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz. Wilhelm Heyne Verlag, München. Im Erscheinen¹⁴.

¹⁴ Die Autor:innen des vorliegenden KARL GF5 Whitepaper haben die Informationen über die bevorstehende Publikation von Katharina Zweig von der Webseite <https://www.penguin.de/ebook/Die-KI-wars/Katharina-Zweig/Heyne/e618108.rhd>. Der Inhalt ist entsprechend nur in den dort zum Zeitpunkt der Publikation des Whitepapers dargestellten Auszügen bekannt.



DOI: 10.5445/IR/1000161674



The work is licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0). You can read the license terms here: <http://creativecommons.org/licenses/by/4.0>

Dieses Forschungs- und Entwicklungsprojekt wird durch das Bundesministerium für Bildung und Forschung (BMBF) im Programm „Zukunft der Wertschöpfung – Forschung zu Produktion, Dienstleistung und Arbeit“ (Förderkennzeichen: 02L19C250) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor:innen.



www.kompetenzzentrum-karl.de



Künstliche Intelligenz
für Arbeit und Lernen



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung