

Forschung 1

Technikfolgenabschätzung aus
Arbeitnehmer:innenperspektive

Schaffareis

Forschung 1

Technikfolgenabschätzung aus
Arbeitnehmer:innenperspektive

**Publikation zur
Wissenschaftskonferenz der
Arbeiterkammer Vorarlberg
im September 2021**

Warum braucht es Forschung zur Arbeitskultur?

Die Gesellschaft braucht Forschung, weil sie auf Fakten, Daten und Experimenten beruhendes Wissen schafft, welches uns hilft, Probleme der Vergangenheit und der Gegenwart zu verstehen – und sie dadurch beiträgt, zukünftige zu meistern.

Arbeitskultur war in den letzten zwanzig Jahren durch große Umbrüche gekennzeichnet: Mit der „Globalisierung“ der 2000er-Jahre kam die internationale Arbeitsteilung, diese wurde von der „Globalisierung“ nochmals beschleunigt und vielfach erst ermöglicht. Mit ihren technologischen Entwicklungen ging der massive Verfall von Wissen einher. Mit „Corona“ kam das Homeoffice für die breite Masse und mit dem „Klimawandel“ wird es zu einem kompletten Umbruch von Arbeits- und Produktionsprozessen kommen.

Der Charakter der Arbeit verändert sich fortwährend. Das Arbeitsrecht definiert eine:n Arbeitnehmer:in als eine Person, die sich aufgrund eines privatrechtlichen Vertrags im Dienste eines anderen zu weisungsgebundener, fremdbestimmter Arbeit in persönlicher Abhängigkeit verpflichtet. „Lohnarbeit“ als Rechtsgegenstand unselbstständiger Erwerbsarbeit ist folglich dieser Transformationen nicht mehr was es einmal war.

Klassische Produktionsstätten fehlen. Durch Teleworking und Homeoffice ist der Arbeitsort nicht mehr klar bestimmbar und die Arbeitsmittel sind oft die eigenen. Immer mehr sind immaterielle Vermögensgegenstände für die Wertschöpfung verantwortlich und die physische Präsenz verliert an Bedeutung. Die ständige Erreichbarkeit durch Handy und Messaging-Apps bedeutet nicht nur Stress durch die Informationsflut, sondern einen Verlust der zeitlichen Selbstkontrolle der Arbeitnehmer:innen und Verdichtung der Arbeit. Die Aufgaben und Verantwortung der „unselbstständigen“ Arbeitnehmer:innen ähnelt heute in vielen Zügen denen von Selbstständigen in Abhängigkeit von einem oder zwei Auftraggebern. Die Ausbildung hat sich nach wenigen Jahren komplett überholt. Die Halbwertszeit von beruflichem Wissen liegt bei 5 Jahren, die von Hochschulwissen beträgt knapp 10 Jahre. Die ständige Reproduktion des Wissens, und damit des „Kapitals“ der Arbeitnehmer:innen, findet auf eigene Kosten statt.

All dies ist interessant, diese Tatsachen finden jedoch keinen Eingang in die Verteilung der Produktivitätsgewinne. Risiken, die früher vom Arbeitgeber getragen wurden, werden auf die Arbeitnehmer:innen abgewälzt. Arbeitsbereitschaft und Flexibilität, der Einsatz eigener Arbeitsmittel und das Wissen laufend auf dem neu-

esten Stand zu halten, sind mit Kosten verbunden, die der Arbeitnehmer:in nicht abgegolten werden.

Das Arbeitsrecht hält mit diesen Entwicklungen nur zum Teil Schritt. Fundierte Forschung und empirische Grundlagen können Anlass für die laufende Weiterentwicklung des Arbeitsrechts geben. Aus diesem Grund ist der Beitrag der Schaffarei Forschung und wissenschaftliche Austausch mit Forschern aus dem europäischen Ausland umso relevanter und kann ein Katalysator sein, um zukünftige/anstehende Transformationen erfolgreich im Sinne der arbeitenden Menschen zu meistern.

Eva King

Verantwortlich für die Schaffarei
Stv. Direktorin der AK Vorarlberg

DD	Projektbericht zur Technikfolgenabschätzung aus Arbeitnehmer:innenperspektive <i>Dominic Götz, Felix Jakob Jagg</i>	09
①	Plattformzentrierte Arbeitskoordination: Flexibilisierung, Standardisierung, Kontrolle <i>Jan-Felix Schrape</i>	31
②	Wer ist schuld, wenn Algorithmen irren? Entscheidungsautomatisierung, Organisationen und Verantwortung <i>Angelika Adensamer, Rita Gsenger, Lukas Daniel Klausner</i>	47
③	Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen <i>Nikolina Grgic, MSc</i>	75
④	SensiTrack - Privacy By Design Konzept zur Gestaltung eines Tracking- und Tracing-Anwendungsfalls im Arbeitsumfeld <i>Christian Jandl</i>	109
	Impressum	135

Dieser Forschungsband besteht aus einem zusammenfassenden Konferenzbericht des Schaffarei Foschung Teams und ausgewählten Beiträgen verschiedener Konferenzteilnehmer:innen. Die Wissenschaftler:innen beschreiben in den Beiträgen, unter besonderer Berücksichtigung der Arbeitnehmer:innenperspektive, jeweils aktuelle Forschungsergebnisse, Projekte und Leitfäden mit konkreten Handlungsempfehlungen.



Projektbericht zur Technikfolgenabschätzung aus Arbeitnehmer:innenperspektive

1	Projekt	11
	1.1 Inhalt	11
	1.2 Zielsetzung	12
	1.3 Rahmen	12
2	Technikfolgenabschätzung aus Arbeitnehmer:innenperspektive	13
3	Neue Technologien für betrieblichen Gesundheitsschutz	14
4	Betriebliche Datenmacht, Überwachung, Mitbestimmung und Datenschutz	15
5	Meine neue Kollegin KI	16
6	Algorithmen: Bias erkennen, Verantwortung übernehmen	17
7	Der neue Weg in die Beschäftigung	19
8	Digitalisierung in der Praxis gestalten	20
9	Schlussfolgerung	22
	Über die Autoren	24
	Quellenverzeichnis	25

1 Projekt

1.1 Inhalt

Die Zukunft der Arbeit ist durch Digitalisierung einem ständigen Wandel unterworfen. Die Technikfolgen der digitalen Transformation berühren die Arbeitnehmer:innen in allen Lebensphasen. Die Aspekte und Folgen des rasanten technischen Fortschritts am Arbeitsplatz müssen diskutiert werden, um Digitalisierung als Chance zu nutzen und Risiken zu antizipieren. Den Arbeitnehmer:inneninteressen stehen unter anderem folgende Herausforderungen gegenüber:

- Diskriminierungsrisiken durch automatisierte Entscheidungssysteme und Algorithmen (Allhutter u. a., 2020; Prietl, 2019; Orwat, 2020)
- Verarbeitung personenbezogener Beschäftigtendaten und Datenschutz (Auer-Mayer, 2020b; Riesenecker-Caba, Astleithner und Chlestil, 2021)
- Überwachung am Arbeitsplatz und Technikstress (Backhaus, 2019; Auer-Mayer, 2020a; Christl, 2021)
- Leistungskontrolle und Leistungserfassung in Arbeitsbeziehungen (Staab und Geschke, 2020; Gogola, 2020)
- Künstliche Intelligenz am Arbeitsplatz (Thieltges, 2020; Apt und Priesack, 2019)

Besonders aufgrund des rasanten Wachstums an Innovationen und Anwendungsbeispielen von neuen Technologien in den Unternehmen gibt es analog dazu eine große Chance für Regulierung und Mitbestimmung auf betrieblicher Ebene, beispielsweise durch das Instrument der Betriebsvereinbarung. Forschung in den Bereichen der Technikfolgenabschätzung, Arbeitssoziologie, Softwareentwicklung, Rechtswissenschaften und vielen anderen Disziplinen liefern die Grundlagen, um daraus entstehende Gefahren und Potenziale früh zu erkennen und Veränderungen aktiv mitzugestalten. Es besteht das Problem, dass großes Know-how vorhanden ist, aber um Mitbestimmungen zu ermöglichen, muss es auch an der

Großes Potenzial an Mitgestaltungsmöglichkeiten – aufgezeigt durch Forschung

richtigen Stelle verfügbar sein. So wünschen sich über 80 % der Betriebsräte fachliche Unterstützung durch die Gewerkschaft oder Arbeiterkammer (Riesenecker-Caba, Astleithner und Chlestil, 2021).

1.2 Zielsetzung

Um Mitbestimmung zu ermöglichen, braucht es einen Diskurs, der die arbeitspolitischen Konfliktfelder aus verschiedenen Disziplinen wissenschaftlich beleuchtet und die Interessen der Arbeitnehmer:innen einbindet. Damit der Wissenstransfer hin zur Praxis unterstützt wird und die Arbeitnehmer:innenvertretung die benötigte Expertise für proaktives Handeln bekommt, braucht es neue Kanäle und Möglichkeiten für Weiterbildung. Mit dem richtigen Know-how aus der Wissenschaft können Handlungsempfehlungen und Lehren zur Zukunft der Arbeit abgeleitet werden. Es müssen aktuelle wissenschaftliche Erkenntnisse gesammelt, diskutiert und aus Arbeitnehmer:innenperspektive angewendet werden. So kann die Grundlage zur Regulierung durch Betriebsvereinbarungen geschaffen werden und Betriebsrät:innen in ihrer Arbeit unterstützt werden.

Wissenschaftliche Erkenntnisse aus Arbeitnehmer:innenperspektive diskutieren und anwenden

1.3 Rahmen

Um eben diese Lücke zu schließen, haben wir dafür eine Plattform im Rahmen eines langfristigen Projekts geschaffen. Im Zuge einer wissenschaftlichen Konferenz sowie einer Publikation hat die AK Vorarlberg Referent:innen aus einem Netzwerk von Forschungs- und Bildungseinrichtungen, Unternehmen und Institutionen, Projekten des AK-Digifonds und anderen Arbeitnehmer:innenvertretungen eingeladen. Die Aufgabenstellung für die Teilnehmer:innen war es, die eigene Forschung aus einer Arbeitnehmer:innenperspektive zu beleuchten, die Folgen des technischen Fortschritts abzuschätzen und Handlungsempfehlungen für Mitbestimmung und Mitgestaltung abzuleiten. Mit der externen Expertise durch Kooperationen mit Referent:innen aus dem wissenschaftlichen Betrieb und der Expertise der Vertreter:innen der Arbeiterkammer hat die zweitägige Konferenz am 09. und 10. September 2021 in der Schaffarei, dem neu eröffneten Vorarlberger Haus für Arbeitskultur, stattgefunden. Die Konferenzteilnahme war nur für die eingeladenen Wissenschaftler:innen möglich.

Wissenstransfer durch eine interdisziplinäre Konferenz in der AK Vorarlberg

In den zwei Konferenztagen haben die Eingeladenen zu sieben verschiedenen Themenschwerpunkten Vorträge gehalten und anschließend im Plenum diskutiert. Um die Arbeitnehmer:innenperspektive in den Mittelpunkt zu stellen, haben wir für eine zielgerichtete Diskussion einen Moderationsleitfaden erstellt, um wichtige Fragen zu zukünftigen Veränderungen am Arbeitsmarkt zu beantworten und einen langfristigen Mehrwert für Betriebsrät:innen zu schaffen.

Nach zwei Tagen mit spannenden Vorträgen und intensiven Diskussionen stellen wir in diesem Praxisbericht einen inhaltlichen Überblick vor. Der Konferenzbericht folgt in seiner Struktur dem Ablauf der zweitägigen Konferenz. Die Inhalte aus den vorgestellten Beiträgen der einzelnen Sessions und der Diskussionen sind darin verarbeitet. Die Kapitel dieses Praxisberichts sind nach den Bezeichnungen der einzelnen Sessions der Konferenz benannt. Im Anhang findet sich eine Liste mit den Namen der Vortragenden.

2 Technikfolgenabschätzung aus Arbeitnehmer:innenperspektive

In der Eröffnungsrede der Konferenz äußerte der Keynote-Speaker Ben Wagner die Vermutung, dass alles was automatisiert werden kann, automatisiert wird. Alles was Arbeit bisher im Kern ausgemacht hat, löst sich zur Zeit immer mehr auf (Ort, Zeit, Bedingungen etc.) (Klammer u. a., 2017). Wagner plädierte, dass Technologie der Schlüssel ist, um Arbeit neu zu denken. Oft folgen Automatisierungsbestrebungen einem technik-zentrierten Design (Lutz, 1987) – der Mensch als Technik-Anhängsel, wie es Wagner überspitzt formuliert hat. Erst nachdem die Automatisierung stattgefunden hat, stellt sich die Frage, wozu der Mensch überhaupt noch gebraucht wird und ob er eigentlich nur als Kontrollorgan fungiert. Für eine nachhaltige Technikgestaltung und eine inklusive Arbeitswelt zwischen Mensch und Technik braucht es ein Umdenken hin zu einer Human Centered Automation (Mitchell, 1996).

Die digitale Transformation hat die klassische Organisation des Arbeitsplatzes in vielen Branchen schon stark verändert und die digitale Plattformökonomie hinsichtlich der Mitgliedschaft, Regeln, Kontrolle, Sanktionen und Hierarchie geprägt (Kirchner und Schüssler, 2019). Anhand der Plattformarbeit ist klar sichtbar, wie schnell und tiefgreifend die Veränderungen stattfinden. Vor allem für die Arbeitnehmer:innen führt das zu Destabilisierung von Er-

Der Mensch als
Technik-Anhängsel?

Plattform-Ökonomie
als Beispiel für
transformatives
Potenzial von Technik

werbsbiografien, Prekarisierung von Arbeitsverhältnissen, Intensivierung von Arbeit, Flexibilisierung von Arbeitsverteilung und Entgrenzung von Erwerbsarbeit (Schreyer und Schrape, 2021).

Nach der Diskussion mit den Expert:innen kann geschlussfolgert werden, dass es eine aktive wissenschaftliche Debatte braucht und in der Folge einen funktionierenden Wissenstransfer bzw. eine gute Kommunikation zwischen den richtigen Stellen und Handlungsträgern hergestellt werden muss. So können die Handlungsfelder erkannt werden und die Basis für Mitbestimmung geschaffen werden. Weiters lassen sich für eine nachhaltige Technikgestaltung unter anderem folgende arbeitspolitischen Implikationen ableiten: frühe Intervention, regulative Rahmensetzungen und betriebliche Mitbestimmung (Schreyer und Schrape, 2021).

Aus der Perspektive der Technikfolgenabschätzung stellen sich hinsichtlich der Zukunft der Arbeit und dem technologischen Fortschritt wichtige Fragen: Wie verändern digitale Technologien unterschiedliche Branchen und welche normativen Erwartungen werden an diese neuen Technologien geknüpft? Welche Forschungsmethoden braucht es und wie kann die Kommunikation und der Wissenstransfer in die Praxis gelingen? Wie können Beschäftigte, Arbeitnehmer:innenvertretung und andere Stakeholder-Gruppen in diese Fragen miteinbezogen werden und wie lässt sich Mitbestimmung in diesem Rahmen ermöglichen (Bösch u. a., 2021)?

Wissenstransfer und Kommunikation für Mitgestaltung essenziell

3 Neue Technologien für betrieblichen Gesundheitsschutz

Durch die Digitalisierung werden die Belastungsfaktoren diversifiziert und intensiviert. Neben starken Umwälzungen auf der organisationalen Ebene wie neue Formen und Aspekte der Arbeitsorganisation (z. B. Plattformökonomie), betrifft der technische Fortschritt die Arbeitsbedingungen direkt. Dabei lässt sich sagen, dass aus einer gesundheitlichen Perspektive die schon bekannten Themen der physischen (Arbeitszeitentgrenzung, Autonomieverlust, Kontrolle und Überwachung etc.) und körperlichen Belastung (automatisiertes Arbeitstempo, ergonomische Bedingungen etc.) neue Herausforderungen gegenüberstehen (Bretschneider u. a., 2020). Personen sollen nicht fit für die Arbeit gemacht werden und an eine sich rasch ändernde und zunehmend digitalisierte Arbeitswelt angepasst werden, sondern die Arbeit sollte menschengerecht gestaltet sein. Ebenso wie beim betrieblichen Gesundheitsschutz wird

Neue Herausforderung – altbekannte Problemfelder

ein gemeinsamer Prozess zwischen Arbeitnehmer:innen und der fortschreitenden Digitalisierung gefordert. Dazu braucht es laut Isabel Koberwein (Gewerkschaft GPA) vor allem eine Präzisierung der Vorgaben zu psychischen Belastungen im Arbeitsschutzgesetz. Die Kontrolle und Durchsetzung solcher Vorgaben gestaltet sich in der Praxis jedoch eher schwierig. Arbeits- und Organisationspsycholog:innen als verpflichtend einzusetzende Präventivfachkräfte sind daher gefragt.

4 Betriebliche Datenmacht, Überwachung, Mitbestimmung und Datenschutz

Die durch die Digitalisierung intensivierete Datenerfassung und -verarbeitung wird im betrieblichen Alltag immer wichtiger. Die Chancen und Risiken sind ungleich zwischen Arbeitgeber:innen und Arbeitnehmer:innen verteilt und Informationsasymmetrien verstärken diese Kluft. Die eingesetzten Systeme reichen von Leistungsauswertungen und Ablaufoptimierung über Überwachung und Kontrolle, bis hin zu automatisierten Management- und Personalentscheidungen. Während Arbeitgeber:innen sich durch die Technologien Produktivitätssteigerung und Kostensenkung durch einen digitalen Taylorismus erhoffen, geraten Arbeitnehmer:innen zunehmend unter Druck durch diese komplexen und oftmals intransparenten Systeme (Christl, 2021). Wie schon bei der Automatisierung scheint auch hier der Leitgedanke zu gelten, dass gemacht wird, was technisch möglich ist. Allerdings ist schon im aktuellen Rechtsrahmen vieles nicht erlaubt und klar gesetzlich geregelt. Neben offeneren Betriebsvereinbarungen für komplexe Anwendungen müssen Betriebsrät:innen unterstützt und weitergebildet werden.

Informationsasymmetrien bei personenbezogenen Daten

Ein sehr weit verbreitetes Produkt, das präsent für eine veränderte Arbeitsweise und neue Form der Organisation im Sinne von „New Work“ steht, ist Microsoft 365. Neben dem vielfältigen Potenzial zur Datenauswertung und Evaluierung durch eine große Menge gesammelter Daten rücken beim Marktführer in diesem Bereich besonders Datenschutzbedenken und die rechtlichen Anforderungen in den Mittelpunkt (Riesenecker-Caba, 2021). Besonders wegen der Umwälzungen durch die Pandemie sind die Produkte noch häufiger im Einsatz (Fritsch, 2021).

Beispiel: Microsoft 365

In einem weiteren Beispiel haben Sascha-Christopher Geschke und Philip Staab gezeigt, wie der Modehändler Zalando die Überwachung und Kontrolle auf die Spitze getrieben hat, indem er sei-

Beispiel: Zalando

ne eigenen Mitarbeiter:innen sich gegenseitig bewerten lassen hat. Dadurch entstand ein immenser innerbetrieblicher Wettbewerb um die Gunst der Arbeitgeber:innen, Gehaltserhöhungen und Beförderungen. In dieser Fallstudie wurde nicht nur dieses Evaluierungstool bis ins Detail aufgezeigt, sondern im Zuge von Interviews auch herausgefunden, dass dieses gegenseitige Kontrollsystem eindeutig dem Betriebsklima geschadet hat und das Ziel verfehlte (Staab und Geschke, 2020). Diese Studie oder selbst nur eine Statusnachricht auf Twitter (1) haben gezeigt, dass wissenschaftlicher Diskurs, mediale Aufmerksamkeit sowie Sichtbarmachung und Aufklärung wichtige Instrumente sind, um Mitbestimmung zu erzwingen und gewisse Praktiken der Unternehmen zu unterbinden. Im Fall von Microsoft 365 stehen dem Betriebsrat beispielsweise viele gesetzliche Möglichkeiten (Einhaltung der Gesetze überwachen, Anhörung/Intervention, Beratung/Information, Mitbestimmung, Vetorechte) zur Verfügung, um mitzugestalten und müssen auch genutzt werden. Dennoch bestehen viele Kritikpunkte aus Datenschutzsicht (Fritsch, 2021).

5 Meine neue Kollegin KI

Eine große Palette an verschiedenen Technologien (Artificial Intelligence, Machine Learning, Robotic Process Automation, Virtual- und Augmented Reality, Cyber Physical Systems, etc.), die hier als Künstliche Intelligenz zusammengefasst werden und zur Automatisierung dienen, können und werden entlang der gesamten Wertschöpfungskette von Unternehmen eingesetzt. Deren Funktionen unterstützen Unternehmen von der Planung über die Beschaffung, Produktion und Logistik bis hin zum Vertrieb und Kundenservice (Plattform Industrie, 2021). Die Plattform Industrie 4.0 unterstützt mit ihrem Projekt „AI for Good“ Unternehmen und Betriebsräte beim Umgang mit und der Einführung von KI-Systemen. Mithilfe von Interviews und Use Cases wurde so ein Praxisleitfaden erstellt. Die Expert:innengespräche haben gezeigt, dass bezüglich der Akzeptanz und des Datenschutzes der Systeme große Herausforderungen bestehen, aber auch schon Instrumente (gute Kommunikation, Betriebsvereinbarungen, DSGVO etc.) zur Lösung bestehen, die sehr aufwendig sein können. Die Mitarbeiter:innen sollten nicht erst bei der Implementierung eingebunden werden, sondern schon bei der Systemauswahl, sowie Mitspracherechte bis hin zur

Wie kommt die künstliche Intelligenz (erfolgreich) in den Betrieb?

Nicht-Einführung haben, wenn keine passenden Rahmenbedingungen für alle geschaffen werden können.(1)

Was in Österreich einzelne Projekte wie das hier vorgestellte sind, steht in Deutschland als Beratungsstelle für Betriebs-, Personalräte und Mitarbeitervertretungen (BEST der Arbeitskammer des Saarlandes) in einer institutionalisierten Form zur Verfügung. Dort werden viele Themen abgedeckt, die für die fachliche Unterstützung bei der betrieblichen Gestaltung von Arbeit und Technik von großer Bedeutung sind. Um diese technisch komplexen Handlungsfelder durch die Digitalisierung abdecken zu können, bietet die BEST unterschiedlichste Angebote und Kompetenzen wie Beratung, Gestaltung, Weiterbildung, Prozessbegleitung, Gutachten, Stellungnahmen, etc. an. Unter anderem führen sie langfristige Projekte wie „Gute Arbeit Saar 4.0“ (2) durch, um Mitbestimmungsgremien hinsichtlich der Digitalisierung und Transformation zu sensibilisieren. Gute Arbeitsbedingungen werden durch Analysen, Gestaltung und Wissenstransfer gefördert. Im Verbundprojekt „Regionales Zukunftszentrum für Künstliche Intelligenz und digitale Transformation Saarland/Rheinland-Pfalz“ (3) ist die BEST ebenfalls als Partnerin involviert. Hier liegt der Schwerpunkt auf der Gestaltung und Handhabung menschenzentrierter KI-Systeme für KMU. Der ganzheitliche Ansatz umfasst Einstieg, Lernmodule für verschiedene Zielgruppen (Arbeitnehmer:innenvertretung, Management etc.), vertiefende Beratung bis hin zur praktischen Erprobung.

Eine Institution für die betriebliche Gestaltung von Arbeit und Technik

6 Algorithmen: Bias erkennen, Verantwortung übernehmen

Die Verwendung von Algorithmen steigt stark an und diese Technologie bietet immer mehr Anwendungsmöglichkeiten und wird mittlerweile allgegenwärtig eingesetzt. Einerseits beeinflussen sie stark die Produkte und Dienstleistungen an sich, aber auch wie sie angeboten und verwendet werden. Kellogg, Valentine und Christin (2020) beschreiben, dass algorithmische Kontrolle am Arbeitsplatz über sechs Wirkungskanäle funktioniert: Mitarbeiter:innen werden durch Algorithmen angeleitet (restricting und recommending), evaluiert (recording und rating) und diszipliniert (replacing

(1) Tweet Wolfie Christl (24.11.2020): <https://twitter.com/WolfieChristl/status/1331221942850949121?s=19>

(2) Projekt Gute Arbeit Saar 4.0: <https://www.best-saarland.de/informationen/betriebsmonitor-gute-arbeit-saar-40/?L=0>

(3) RZzKI: <https://www.best-saarland.de/informationen/rzzki/>

und rewarding). Die Folgen sind Diskriminierung, Überwachung, Stress oder Prekarisierung (Kellogg, Valentine und Christin, 2020).

Ein prominentes Beispiel in der österreichischen Verwaltung ist der AMS-Algorithmus. Ziel des Arbeitsmarktservice war es, die Effektivität des Mitteleinsatzes und die Effizienz der Beratung sowie die Fördervergabe zu standardisieren. Allerdings wurde beispielsweise folgende Schwächen gezeigt: Viele Faktoren (Motivation, Auftreten, aktuelle Arbeitsmarktsituation) können durch den Algorithmus abgebildet werden; Gleichbehandlung, Standardisierung und Objektivität werden nicht erreicht bzw. sind nicht gewährleistet; die Daten für den Algorithmus haben die Ungleichheit am Arbeitsmarkt widergespiegelt (Daten, Aufbau, Funktionsweise); zuletzt verbietet die DSGVO in diesem Bereich grundsätzlich vollautomatisierte Entscheidungen. Anhand dieses Praxisbeispiels wurde umfangreich gezeigt, dass es klare rechtliche und politische Rahmenbedingungen und Gouvernancestrukturen sowie den Ausbau dieser braucht, um solche Systeme gerecht zu gestalten. Vor allem müssen Anti-Diskriminierungsmaßnahmen, System- und Datentransparenz und die Nachvollziehbarkeit bei der Einführung solcher Systeme erfüllt werden (Allhutter u. a., 2020).

Bisher hat sich die Forschung vor allem mit den Auswirkungen bei der Einführung algorithmusunterstützter Entscheidungssysteme auf Betroffene beschäftigt, weniger auf die Verantwortung und Auswirkungen für Entscheidungsträger:innen. Durch die Einführung können notwendige Informationen oder Zugriffsrechte fehlen oder neue Aufgabenbereiche (Datenschutz, etc.) entstehen. Mit „VERA“ haben die Autor:innen ein Tool für die Einführung von Algorithmen in der Entscheidungsfindung präsentiert. Dieser Leitfaden zu Verantwortlichkeit und Rechenschaft hilft einerseits geplante Systeme zur automatisierten Entscheidungsfindung frühzeitig zu reflektieren und unterstützt andererseits dabei mögliche Interessenskonflikte innerhalb des Unternehmens aufzuzeigen (Adensamer und Klausner, 2021).

Besonders Betriebsvereinbarungen sind ein wichtiges Instrument zur evaluativen und prozessorientierten Mitbestimmung. Die sogenannte Black Box (Nachvollziehbarkeit, Korrelation/Kausalität, Gewichtung, Modell im Hintergrund), Datenquellen und die kontinuierliche und dynamische Weiterentwicklung stellen offene Probleme für den betrieblichen Kontext dar und verlangen strenge Kontrollmechanismen. Hinzu kommt, dass die erwähnten Technologien noch in den Kinderschuhen stecken, sich aber rasant entwickeln und immer komplexer werden (Thieltges, 2020). Wie die menschl

Diskriminierung, Überwachung, Stress und Prekarisierung durch Algorithmen

Transparenz und präventive Maßnahmen zur Vermeidung von Diskriminierung sind dringend notwendig.

VERA:
Ein begleitendes Tool für die Einführung von Algorithmen und Verantwortlichkeiten

che Entscheidungsfindung wirft die automatisierte Entscheidungsfindung viele Fragen hinsichtlich Diskriminierung, Transparenz, Nachvollziehbarkeit und Verantwortung auf und kann anhand der hier vorgestellten Beispiele nicht alleine als Lösung dienen.

7 Der neue Weg in die Beschäftigung

Bei der Stellen- und Bewerber:innensuche sowie dem gesamten Bewerbungsprozess haben neue Strukturen und Prozesse durch die Digitalisierung Einzug gefunden. Einer der bedeutendsten Treiber für den Einsatz von KI ist dabei der Wettbewerb um Fachkräfte in einigen Branchen. Die Vision ist die Digitalisierung, Beschleunigung und Personalisierung des Bewerbungsprozesses, um dadurch nicht nur die Genauigkeit, sondern auch die Reichweite der Stellensuche zu erhöhen. Allerdings wird ebenfalls festgehalten, dass eine bedeutende Lücke zwischen der betrieblichen Praxis und den theoretisch technischen Möglichkeiten besteht (Holtgrewe, Lindorfer und Vana, 2020). Dennoch verändern digitalisierte Bewerbungsverfahren die Organisation der Stellen- und Bewerber:innensuche.

Durch den Einsatz von KI in Bewerbungsprozessen bestehen unter anderem bereits bekannte Probleme durch die Verzerrung aufgrund von Trainingsdaten (Dastin, 2018) oder der Benachteiligung durch Proxyvariablen. So können zum Beispiel Wohnadressen Aufschluss über die sozioökonomische Zugehörigkeit geben (Lopez, 2019), ähnlich dem zuvor erwähnten Problem der intransparenten Black Box. Beim Einsatz diskriminierender KI sind gesetzliche Handlungsmöglichkeiten beschränkt und die verschiedenen Rechtsnormen müssen besser aufeinander abgestimmt werden, um Rechtsprobleme aus dem Weg zu räumen. Der Arbeits- und Sozialrechtsexperte Elias Felten fordert deswegen unter anderem: Die Einführung einer generellen Offenlegungspflicht bei der Verwendung von KI weist aber gleichzeitig auf die wettbewerbsrechtliche Problematik hin; Klärung der Zuständigkeit der Betriebsrät:innen, weil es sich noch um Bewerber:innen im Recruitingprozess handelt und keine Arbeitnehmer:innen (Schlüselfunktion hinsichtlich der Informations- und Mitbestimmungsrechte inne haben); Zusammenarbeit und Austausch zwischen Techniker:innen und Arbeitsrechtler:innen.

Ein weiteres bestehendes Problem in Bewerbungsprozessen ist der unconscious Bias. Besonders bei der Mitarbeiter:innensuche reichen die Folgen von unfairer Behandlung bis hin zu schwer-

Stellensuche, Bewerbung und Wettbewerb um Fachkräfte voll digitalisiert?

Schutz vor Diskriminierung durch Algorithmen im Bewerbungsprozess

wiegender Diskriminierung durch verschiedenste Effekte oder den Prozess an sich. Eine vermeintlich „objektive“ Lösung dafür sollen Künstliche Intelligenzen sein. Algorithmen sind aber nie neutral und objektiv. Design- und Datenbasis verhindern keinen Bias, sondern lenken ihn und es bestehen Transparenz- und Accountability-Probleme. Mit dem Projekt „DeBias“ haben die Initiator:innen Florian Cech und Hilda Tellioglu den Mensch in den Mittelpunkt gestellt und nicht durch KI ersetzt. Mit den Strategien Anonymisierung (reduziert Bias aufgrund persönlicher Eigenschaften, erstes fachliches Assessment), Strukturierung (gleicher und vergleichbarer Ablauf, weitere Biasreduzierung) und Entscheidungsfindung (parallele und sequenzielle Bewertung, Entscheidung beim Mensch) haben sie versucht den Bias, mit dieser sozio-technischen Lösung zu verringern (Tellioglu und Cech, 2021).

Algorithmen
sind nie „objektiv“!

Sozio-technische
Lösungen können
auch helfen Bias zu
verringern, wie
im Projekt „DeBias“

8 Digitalisierung in der Praxis gestalten

Die letzte Session der Konferenz eröffnete Christian Jandl von der FH St. Pölten. Er stellte sein Privacy by Design-Konzept zur Gestaltung von Tracking- und Tracing-Anwendungen im Arbeitsumfeld vor (Jandl u. a., 2021). Ziel ist Maßnahmen zu entwickeln, die ermöglichen, die positiven Aspekte von Tracking-Technologie für Unternehmen und Benutzer:innen zu unterstützen, aber das Potenzial für Überwachung durch Arbeitgeber:innen bewusst einzuschränken. Risikoanalyse zur Mitarbeiter:innen-Privatheit, das Sicherstellen von Datenschutz durch Privacy by Design Strategien (Hoepman, 2014) und die von Jandl erarbeiteten Richtlinien sollen eine mitarbeiter:innenzentrierte Gestaltung ermöglichen. Die Richtlinien:

- Unterscheidung zwischen den Anwendungsfällen
Asset Tracking/MA-Monitoring
- Privatheit ist ein zwingendes Bedürfnis,
keine mögliche Option
- Einfluss des Stellenwerts von Privatheit in der Organisation,
Diskussion auch über unbequeme Themen
- Mitarbeiter:innen-Wahrnehmung ernst nehmen –
Akzeptanzfaktoren von Trackingsystemen beachten

Ein anderer Blickwinkel konnte durch den Vortrag des Geschäftsführers eines Unternehmens gewonnen werden, welches international Tracking- und Tracing-Systeme vertreibt. Erst durch den Bericht eines Mitarbeiters im Vertrieb wurde die Geschäftsführung über einen Fall von missbräuchlicher Verwendung ihrer Technologie aufmerksam. Die für Sicherheit in hoch riskanten Arbeitsumgebungen (wie Bohrinseln und Minen) entwickelten Tracking-Halsbänder wurden zur automatischen Zeiterfassung auf Baustellen verwendet. Auch wenn solche Beispiele der zweckentfremdeten Anwendung nicht im Sinne der Entwickler:innen sind, so kann ex post meist wenig Einfluss genommen werden ohne Vertragsbruch zu begehen. Denn Hard- und Software-Anbieter:innen ist demnach beim Design ihrer Produkte eine bedeutende Rolle für eine menschenzentrierte Gestaltung von Tracking- und Tracing-Systemen zuzuschreiben.

Ein positives Beispiel für die gemeinsame Gestaltung von Digitalisierungsprozessen durch Arbeitgeber und Betriebsrat rundete die Konferenz ab. Der Betriebsrat eines großen Vorarlberger Industrie-Unternehmens schilderte den Prozess, der es ihm erlaubte, die Digitalisierung im Betrieb mitzugestalten. Voraussetzungen dafür waren eine gute Gesprächsbasis und die Unterstützung durch die Geschäftsführung. Ein gemeinsamer Workshop mit externen Expert:innen (aus dem Feld der Technik-Beratung und Betriebsrät:innen-Beratung) wurde finanziert und darauf aufbauend eine Rahmen-Betriebsvereinbarung erstellt. Sie definiert die Prinzipien für den Umgang mit neuen Technologien im Betrieb und ermöglicht so die Mitgestaltung durch gezielte Betriebsvereinbarungen für jeden Anwendungsfall. An diesem Beispiel ist gut zu sehen, dass ein sozio-technischer Prozess wesentlich besser hinsichtlich der Akzeptanz und Wirksamkeit funktioniert als eine einseitige Technikzentrierung.

9 Schlussfolgerung

So unterschiedlich die diskutierten Themenfelder scheinen mögen, oft wurden in den Vorträgen und der Diskussion Parallelen und ähnliche Probleme aufgezeigt. Grundsätzlich konnten wir erkennen, dass, um den technischen Fortschritt menschenzentriert und nachhaltig mitzugestalten, die Technik eine Unterstützung für den Mensch darstellen muss und nicht umgekehrt. Ebenso sollten Mensch und Technik sich ergänzen und nicht ersetzen. Oft wird der Mensch durch „Maschinen“ ersetzt, aber dennoch muss der Mensch, wie wir in vielen Beispielen gesehen haben, bei Fehlern oder Problemen einschreiten und die Systeme überwachen, weil eine gänzliche Automatisierung nicht zulässig oder möglich ist (Bainbridge, 1983).

Digitalisierung darf nicht als Selbstzweck verstanden werden, sondern als steuerbarer Prozess, von dem wir profitieren können. Die Folgen und Risiken der Implementierung neuer Systeme und der Digitalisierung von Prozessen müssen neben den Chancen schon vor der Einführung analysiert werden. Transparenz, Schutz der Privatsphäre und Menschenzentrierung sollen schon in der Ausgestaltung stecken und Mitbestimmung in allen Phasen möglich sein.

Eine besondere Herausforderung ist die immer größere Komplexität der eingesetzten und zukünftigen Technologien sowie der rasante Fortschritt. Es wird immer schwieriger das Know-how gebündelt an einer Anlaufstelle oder bei einzelnen Personen zu bekommen, weil es immer spezifischeres Wissen zu den einzelnen Themenfeldern benötigt. Besonders bezüglich des Wissenstransfers muss auf eine konstruktive Debatte und Sensibilisierung gesetzt werden. Das großteils vorhandene Wissen muss an die richtigen Stellen gelangen und Expert:innen müssen ihr Know-how beispielsweise Betriebsrät:innen zur Verfügung stellen, um komplexe und umfangreiche Themen zu verstehen und früh sowie situationsangepasst darauf reagieren zu können. Gleichzeitig ist die Vernetzung von Institutionen und die Förderung der Zusammenarbeit von verschiedenen Disziplinen (Recht, Technik, Informatik, Arbeitssoziologie etc.) notwendig. Betriebsrät:innen brauchen bei der Ausarbeitung von Betriebsvereinbarungen die Unterstützung technischer und rechtlicher Experten vor Ort. Dafür müssen spezialisierte Anlaufstellen geschaffen werden, um eine Überforderung zu vermeiden. Selbst bei bestehenden Betriebsvereinbarungen kann es bei internationalen Unternehmen mit einer Konzernzentrale im Ausland zu Updates des Systems kommen, ohne dass der Betriebsrat es mit-

Digitalisierung nicht als Selbstzweck!

Wissenstransfer und persönliche Unterstützung bei der Mitgestaltung

bekommt. Gerade bei hochkomplexen Themen wie Künstlicher Intelligenz, Big Data oder Algorithmen braucht es Antizipation statt Reaktion, um mit den rasanten Entwicklungen Schritt zu halten. Es braucht Regeln und die Aktualisierung und Schaffung von Rechtsnormen, um gesetzliche Grundlagen für kommende Regulierungen zu haben. Der Entwurf der EU-Kommission ist daher grundsätzlich hinsichtlich einheitlicher Regelungen zu Künstlicher Intelligenz zu begrüßen (Kommission, 2021).

Um diese Ziele zu erreichen, muss die Arbeitnehmer:innenvertretung Ressourcen für die Aus- und Weiterbildung von Betriebsrät:innen bereitstellen. Im besten Fall sollten die Unternehmen Betriebsrät:innen finanziell unterstützen, um bei der Gestaltung von Betriebsvereinbarungen externe technische und rechtliche Expertise hinzuholen zu können. Es sollte unabhängig vom Gutwillen der Unternehmen im Interesse der Arbeitnehmer:innenvertretung sein, den Betriebsrät:innen eine adäquate Unterstützung zu ermöglichen. Nur so können in Zusammenarbeit einvernehmliche Lösungen ausgearbeitet werden, von denen alle profitieren. Parallel dazu braucht es alternative (kostenlose) Anlaufstellen, die für die oben genannten Zwecke von der Arbeitnehmer:innenseite, aber auch von der Arbeitgeber:innenseite genutzt werden können. Diese Stellen sollen aufsuchende Unterstützung anbieten und proaktiv auf Lücken aufmerksam machen, um den technischen Fortschritt menschenzentriert zu gestalten und Transparenz, Datenschutz und Mitbestimmung zu verinnerlichen.

Informierte und einvernehmliche Lösungen, sind die besseren Lösungen.

Dominic Götz

ist Referent für Grundlagenarbeit bei der Arbeiterkammer Vorarlberg und betreut das Projekt Schaffarei Konferenz sowie die dazugehörige jährliche Publikation, Schaffarei Forschung.

Felix Jakob Jagg

hat als Fachreferent die AK Vorarlberg beim Projekt Schaffarei unterstützt und die wissenschaftliche Konferenz 2021 organisiert. Er arbeitet im Sozialministerium in Wien.

Quellenverzeichnis

- Adensamer, Angelika und Klausner, Lukas Daniel (2021). „Algorithmen in der Entscheidungsfindung: Leitfaden zur Verantwortlichkeit und Rechenschaft“. In: URL: https://wien.arbeiterkammer.at/interessenvertretung/arbeitsdigital/DataPolitics/VerA_Leitfaden_Final.pdf (abgerufen am 23.09.2021).
- Allhutter, Doris, Cech, Florian, Fischer, Fabian, Grill, G. und Mager, Astrid (2020). „Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective“. In: Special Issue Critical Data and Algorithm Studies, S. 17.
- Apt, Wenke und Priesack, Kai (2019). „KI und Arbeit-Chance und Risiko zugleich“. In: Künstliche Intelligenz. Springer Vieweg, Berlin, Heidelberg, S. 221–238.
- Auer-Mayer, Susanne (2020a). „Dürfen Arbeitnehmer*innen im „Home-Office“ überwacht werden?“ In: CuRe 88. — (2020b). „Individualrechtliche Fragen des Datenschutzes“. DE. In: Arbeitsrechtliche Probleme des neuen Datenschutzrechts. Wien: new academic press, S. 42–88.
- Backhaus, Nils (März 2019). „Kontext-sensitive Assistenzsysteme und Überwachung am Arbeitsplatz: Ein meta-analytisches Review zur Auswirkung elektronischer Überwachung auf Beschäftigte“. In: Zeitschrift für Arbeitswissenschaft 73. doi: 10.1007/s41449-018-00140-z.
- Bainbridge, Lianne (1983). „Ironies of automation“. In: Analysis, design and evaluation of man-machine systems. Elsevier, S. 129–135.
- Böschen, Stefan, Grunwald, Armin, Krings, Bettina-Johanna und Rösch, Christine (2021). „Technikfolgenabschätzung – neue Zeiten, neue Aufgaben“. In: Technikfolgenabschätzung, S. 15.
- Bretschneider, Maximilian, Drössler, Stephanie, Magister, Selina, Zeiser, Maria, Kämpf, Daniel und Seidler, Andreas (2020). „Digitalisierung und Psyche: Rahmenbedingungen für eine gesunde Arbeitswelt. Ergebnisse des Projektes GAP“. In: Zeitschrift für Arbeitswissenschaft 74.2, S. 63–75.
- Christl, Wolfie (2021). „Digitale Überwachung und Kontrolle am Arbeitsplatz: Von der Ausweitung betrieblicher Datenerfassung zum algorithmischen Management?“ In: URL: <https://crackedlabs.org/daten-arbeitsplatz> (abgerufen am 23. 09. 2021).
- Dastin, Jeffrey (2018). „Amazon scraps secret AI recruiting tool that showed bias against women“. In: Reuters. URL: <https://www.reuters.com/article/usamazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (abgerufen am 23. 09. 2021).

Fritsch, Clara (2021). „Die wunderbare Welt von Microsoft 365 und wie der Betriebsrat sie mitbestimmen kann“. In: GPA Broschürenserie Gute Arbeit! URL: <https://www.gpa.at/themen/digitalisierung/die-wunderbare-weltvon-microsoft-365> (abgerufen am 23.09.2021).

Gogola, Michael (2020). „Digitale Ratings als rechtliche Herausforderung“. In: Arbeit in der Data Society: Zukunftsvisionen für Mitbestimmung und Personalmanagement. Hrsg. von Verena Bader und Stephan Kaiser. Wiesbaden: Springer Fachmedien Wiesbaden, S. 173–188. doi: 10.1007/978-3-658-32276-2_11.

Hoepman, Jaap-Henk (2014). „Privacy design strategies“. In: IFIP International Information Security Conference. Springer, S. 446–459.

Holtgrewe, Ursula, Lindorfer, Martina und Vana, Irina (2020). Bewerbungsprozesse im Wandel: Alte und neue Anforderungen an Bewerber*innen in Niederösterreich. Studie des Arbeitmarktservice Niederösterreich.

Jandl, Christian, Wagner, Markus, Moser, Thomas und Schlund, Sebastian (Juni 2021). „Reasons and Strategies for Privacy Features in Tracking and Tracing Systems - A Systematic Literature Review“. In: Sensors 21.13, S. 4501. doi: 10/gmf8sx.

Kellogg, Katherine C, Valentine, Melissa A und Christin, Angele (2020). „Algorithms at work: The new contested terrain of control“. In: Academy of Management Annals 14.1, S. 366–410.

Kirchner, Stefan und Schüßler, Elke (2019). „The organization of digital marketplaces: Unmasking the role of Internet platforms in the sharing economy“. In: Organization outside organization, S. 131–154.

Klammer, Ute, Steffes, Susanne, Maier, Michael F, Arnold, Daniel, Stettes, Oliver, Bellmann, Lutz und Hirsch-Kreinsen, Hartmut (2017). „Arbeiten 4.0—Folgen der Digitalisierung für die Arbeitswelt“. In: Wirtschaftsdienst 97.7, S. 459–476.

Kommission, Europäische (2021). Zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Gesetz über künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union. URL: <https://eur-lex.europa.eu/legalcontent/DE/TXT/?uri=CELEX%5C%3A52021PC0206> (abgerufen am 23.09.2021).

Lopez, Paola (2019). „Reinforcing inter-sectional inequality via the AMS algorithm in Austria“. In: Critical Issues in Science, Technology and Society Studies. Conference Proceedings of the 18th STS Conference (Graz: Verlag der Technischen Universität), S. 1–19.

Lutz, Burkart (1987). Das Ende des Technikdeterminismus und die Folgen: soziologische Technikforschung vor neuen Aufgaben und neuen Problemen. Campus Verlag.

Mitchell, Christine M (1996). „Human-Centered Automation: A Philosophy, Some Design Tenets, and Related Research“. In: Human Interaction with Complex Systems. Springer, S. 377–381.

Orwat, C. (2020). Risks of Discrimination through the Use of Algorithms. A study compiled with a grant from the Federal Anti-Discrimination Agency. Techn. Ber. Federal Anti-Discrimination Agency. 122 S. doi: 10.5445/IR/1000123477.

Plattform Industrie, 4.0 (2021). „Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen“. In: noch nicht veröffentlicht.

Priegl, Bianca (2019). „Algorithmische Entscheidungssysteme revisited: Wie Maschinen gesellschaftliche Herrschaftsverhältnisse reproduzieren können“. In: feministische studien 37.2, S. 303–319.

Riesenecker-Caba, Thomas (2021). „Microsoft 365: Digitale Anwendungen im Büro zwischen Kommunikation, Kollaboration u. Kontrolle“. In: Youtube. URL: <https://www.youtube.com/watch?v=mhJFwczBZjl> (abgerufen am 23.09.2021).

Riesenecker-Caba, Thomas, Astleithner, Franz und Chlestil, Martina (2021). Verarbeitung personenbezogener Beschäftigtendaten und Grenzen betrieblicher Mitbestimmung in einer digitalisierten Arbeitswelt. Techn. Ber. FORBA Forschungs und Beratungsstelle Arbeitswelt.

Schreyer, Jasmin und Schrape, Jan-Felix (2021). Digitale Plattformen in kommerziellen und gemeinwohlorientierten Arbeitszusammenhängen. 460. Study der Hans-Böckler-Stiftung.

Staab, Philipp und Geschke, Sascha-Christopher (2020). Ratings als Arbeitspolitisches Konfliktfeld: Das Beispiel Zalando. 429. Study der Hans-Böckler-Stiftung.

Tellioglu, Hilda und Cech, Florian (2021). „Projekt: DeBias“. In: URL: <https://debias.cisvienna.com/> (abgerufen am 23.09.2021).

Thieltges, Andree (2020). „Machine Learning Anwendungen in der betrieblichen Praxis: Praktische Empfehlungen zur betrieblichen Mitbestimmung“. In: Studie der Hans-Böckler Stiftung. URL: https://www.researchgate.net/publication/344556200_MACHINE_LEARNING_ANWENDUNGEN_IN_DER_BETRIEBLICHEN_PRAxis_Praktische_Empfehlungen_zur_betrieblichen_Mitbestimmung (abgerufen am 23.09.2021).

Vortragende der Konferenz (chronologisch)

- Ben **Wagner**, PhD – AI Fututres Lab (TU Delft)
Keynote: Anhängsel Mensch – Nachhaltige Technik zwischen Pandemie und Klimakatastrophe
- Johann **Cas**, Ing. Mag. – ITA – Institut für Technikfolgenabschätzung Wien
Künstliche Intelligenz und Arbeit – Chancen nutzen, Risiken minimieren
- Bettina-Johanna **Krings**, Dr.ⁱⁿ – Karlsruher Institut für Technik - ITAS
Die Technisierung der Arbeit und ihre Folgen
- Jan-Felix **Schrape**, PD. Dr. – Universität Stuttgart
Digitalisierung und Arbeit: Ambivalenzen und Herausforderungen für die Technikfolgenabschätzung
- Isabel **Koberwein**, Mag.a – GPA Grundlagenarbeit
Digitalisierung und der Faktor Mensch – Anforderungen und Möglichkeiten für den betrieblichen Gesundheitsschutz
- Thomas **Riesenecker-Caba**, Mag.
FORBA - Forschungs- und Beratungsstelle Arbeitswelt
Microsoft 365 New Work – Im Spannungsfeld von Digitalisierung und Unternehmenskultur
- Wolfie **Christl** – Cracked Labs
Gläserne Belegschaft 4.0 – Digitale Überwachung und Kontrolle im Betrieb
- Sascha-Christopher **Geschke** M.A. – Humboldt-Universität zu Berlin
Gläserne Mitarbeiter:innen und falsche Versprechen – Konfliktfelder softwaregestützter Arbeitskontrolle
- Tobias **Szygula**, Oliver **Müller** – BEST Arbeitskammer des Saarlandes
Herausforderungen an die Mitbestimmung im Zeitalter der digitalen Transformation – ein Einblick in die Projekte von BEST
- Nikolina **Grgic**, M.Sc. – Plattform Industrie 4.0
AI for GOOD – menschenzentrierter Einsatz von KI

- Bernhard **Moser**, Dr. – SCCH – Software Competence Center Hagenberg
Teamwork zwischen Mensch und KI – Das Konzept TEAMING.AI
- Lukas Daniel **Klausner**, Dipl. Ing.
St. Pölten University of Applied Sciences
Wer ist schuld, wenn Algorithmen irren? Entscheidungsautomatisierung,
Organisationen und Verantwortung
- Andree **Thieltges**, MA – Hochschule für Politik TU München
Immanente Leistungs- und Verhaltenskontrolle durch
Machine-Learning und KI?
- Fabian **Fischer**, Univ.-Ass. Dipl.-Ing.
Interdisziplinärer Sozial- und Computerwissenschaftler
AMS-Algorithmus: Signalwirkung für Algorithmen in der Verwaltung
- Florian **Cech**, Univ.-Ass. Dipl.-Ing. – TU Wien, C!S Vienna
Project DEBIAS: Unconscious Bias & Recruiting
- Irina **Vana**, Dr.ⁱⁿ – ZSI – Zentrum für soziale Innovation Wien
Tradition, Vorlieben und Algorithmen – Bewerbungsverfahren im Wandel?
- Elias **Felten**, Dr. – JKU Linz
Algorithmusbasierte Auswahl von neuen Mitarbeiter:innen
- Christian **Jandl**, Dipl.-Ing. – St. Pölten University of Applied Sciences
SensiTrack – Tracking vs Privacy in der Arbeitswelt 4.0
- Dietmar **Dünser**, Mag., B.Sc. – Zumtobel Group, Betriebsratsvorsitzender
- Urban **Siller**, Ph.D., MBA – Identec Solutions AG, Geschäftsführer



Plattformzentrierte Arbeitskoordination: Flexibilisierung, Standardisierung, Kontrolle

1	Digitale Plattformen:Ermöglichung und Kanalisierung	35
2	Ausprägungen plattformzentrierter Arbeit	36
3	Das Beispiel Essenslieferdienste	38
4	Plattformkoordinierte Arbeit in klassischen Betrieben	40
5	Arbeitspolitische Implikationen	40
	Über den Autor	43

Die digitale Transformation, also die Informatisierung gesellschaftlicher Koordinationsstrukturen sowie die damit einhergehenden sozioökonomischen Adaptionsprozesse, hat in den zurückliegenden 25 Jahren tiefgreifende Rekonfigurationen in fast allen Markt- und Organisationszusammenhängen angestoßen, die von mannigfaltigen Hoffnungen und Befürchtungen begleitet worden sind. Das gilt auch für das weite Feld der Erwerbsarbeit, die seit jeher auch auf dem Einsatz sach- und symboltechnischer Systeme gründet. Sie unterliegt seit der industriellen Revolution einer intensivierten Technisierung, die auf eine Steigerung der Effizienz und eine Substitution menschlicher Arbeitsleistung ausgerichtet ist (Pfeiffer 2018). Diese Technisierung führte ab dem 19. Jahrhundert im Verbund mit der weiteren Verfestigung kapitalistischer Wirtschaftsweisen zu der Entstehung neuartiger arbeitspolitischer Bewegungen und Konfliktlinien. Die dadurch angestoßenen Aushandlungsprozesse mündeten in der Einführung erster arbeitsschutzrechtlicher Regulierungen und insbesondere in Deutschland, Österreich und der Schweiz ab den 1950er-Jahren in der Verbreitung des sogenannten ‚Normalarbeitsverhältnisses‘, das „Korsettstangen von Gewißheit, Voraussehbarkeit und Frieden“ in die Arbeitswelt einzog (Mückenberger 1985: 17).

Schon ab den 1980er-Jahren wurden angesichts der ökonomischen Indienstnahme weiterer technischer Neuerungen wie Industrierobotern und EDV-Systemen, politischer Deregulierungsmaßnahmen und steigender Erwerbslosenzahlen allerdings erste Diagnosen zu einer Erosion des Normalarbeitsverhältnisses oder einem Ende der Arbeitsgesellschaft insgesamt formuliert. Obgleich die entsprechenden Vorhersagen in dieser Intensität keineswegs eingetroffen sind, entwickelte sich in der deutschsprachigen Arbeitsforschung schon ab dieser Zeit eine rege Debatte um die Folgen der Technisierung und die damit verknüpfte Relevanzverschiebung von der Industrie- zur Dienstleistungsarbeit (Überblick: Trinczek 2011). In diesen Diskussionen wurden bereits viele Dynamiken thematisiert, die im laufenden Diskurs um die durch die COVID-19-Pandemie beschleunigte Digitalisierung der Arbeitswelt erneut hervortreten, darunter vor allen Dingen (1) eine *Destabilisierung* von Erwerbsbiografien sowie eine Individualisierung von Beschäftigungsprofilen, (2) eine *Prekarisierung* von Arbeitsverhältnissen und eine damit einhergehende Ausweitung des Niedriglohnsektors, (3) eine Intensivierung von Arbeit durch technische und organisatorische Rationalisierungs- und Kontrollmaßnahmen, (4) eine an Marktgegebenheiten orientierte *Flexibilisierung* von Arbeitsabläu-

Frühe Debatten
zum Ende der
Arbeitsgesellschaft

fen, Arbeitsverteilungen und Arbeitszeiten sowie (5) eine zeitliche und räumliche *Entgrenzung* zwischen Erwerbsarbeit und Privatleben u.a. durch Telearbeit („Homeoffice“).

Diese seit mehreren Jahrzehnten beobachtbaren Trends gehen für viele Erwerbstätige in immer mehr Arbeitsbereichen mit deutlich erweiterten Entscheidungsspielräumen und zugleich mit einem erhöhten inneren Antrieb zur *Selbstkontrolle* (Steuerung eigener Tätigkeiten), *Selbstökonomisierung* (Selbstvermarktung, Weiterqualifikation) und *Selbstrationalisierung* (Durchorganisation des Alltags) einher. Um diesen neuen Erwerbstätigen-Typus zu beschreiben, haben Voß und Pongratz (1998: 131) früh das pointierte Konzept des „Arbeitskraftunternehmers“ entwickelt. Die skizzierten Diskursverläufe führen vor Augen, dass sich viele Verschiebungen auf dem Feld der Arbeit nicht alleine auf die soziale Aneignung digitaltechnischer Strukturen zurückführen lassen, sondern weit davor angestoßen worden sind.

Die fortschreitende digitale Transformation bzw. die damit verknüpfte *Plattformisierung der Arbeitskoordination* und die informationstechnische Automatisierung entsprechender Entscheidungsprozesse hat allerdings einen substantziellen intensivierenden Einfluss auf diese Veränderungsdynamiken. Da sich eine ebensolche zunehmende Plattformorientierung bereits auf vielen alltäglichen Feldern der Koordination von Arbeit beobachten lässt (Papsdorf 2019), während die umfassende Automatisierung von Arbeits- und Entscheidungsprozessen durch intelligente IT-Systeme („künstliche Intelligenz“) in der Praxis noch eher als ein bereichsspezifisches Phänomen beschrieben werden kann, stehen in den folgenden Abschnitten soziotechnische Dynamiken der Plattformisierung im Vordergrund.

Selbstkontrolle und
Selbstrationalisierung

Intensivierender
Einfluss der
Digitalisierung

1 Digitale Plattformen: Ermöglichung und Kanalisierung

Forschungsarbeiten der sozialwissenschaftlichen Technikanalyse und Technikfolgenabschätzung (Überblick: Bösch et al. 2021, Schrape 2021) weisen auf drei grundsätzliche Eigenheiten technischer Strukturen bzw. Infrastrukturzusammenhänge hin, die sich auch auf digitale Plattformen und deren Effekte in der Arbeitskoordination übertragen lassen:

Grundlegende
Eigenheiten

- Erstens haben digitale Plattformen *ermöglichende Eigenheiten*: Sie effektivieren Kommunikations- und Produktionszusammenhänge, vereinfachen Abstimmungsprozesse und erleichtern die Kollaboration.
- Zweitens bilden Plattformen *kanalisierende Eigenheiten* aus: Durch ihre Koordinationsleistungen strukturieren sie individuelle, kollektive und organisationale Beziehungsmuster und prägen Arbeitszusammenhänge.
- Drittens erweitern digitale Plattformen die Möglichkeiten *sozialer Kontrolle*: Durch ihre Strukturierungs- und Kanalisierungsleistungen erleichtern sie die systematische Überwachung, Steuerung sowie positive und negative Sanktionierung von Erwerbstätigen.

All diese Eigenheiten, ihre konkreten Gewichtungen und Konsequenzen sind allerdings nicht durch die Technik (Hardware- und Softwaresysteme) per se festgelegt, sondern das Ergebnis vielfältiger sozialer Implementations- und Aushandlungsprozesse innerhalb und außerhalb von Betrieben, die in äußerst langfristige Prozesse gesellschaftlichen Wandels eingelassen sind. Die digitale Transformation der Arbeitskoordination bezeichnet in diesem Sinne die sich wechselseitig verstärkende Verfestigung neuartiger oder veränderter soziotechnischer Prozesszusammenhänge durch die soziale Aneignung informationstechnischer Plattformen sowie die damit verknüpften Rekonfigurationen sozioökonomischer Ordnungsmuster. Diese betreffen die betriebliche Koordination, formalisierte und informelle Arbeitspraktiken, aber auch kollektive Problemwahrnehmungen, in deren Zuge viele bereits zuvor aufgespannte Diagnosen – z. B. eine Prekarisierung und Entgrenzung von Arbeit oder eine Agilisierung von Organisationen – neu konturiert werden.

Plattformisierung
als soziotechnischer
Prozess

2 Ausprägungen plattformzentrierter Arbeit

Insbesondere die vergleichsweise jungen Phänomene der außerbetrieblichen *Plattformarbeit* (Gig-, Click-, Cloud- und Crowdwork) wurden in der sozialwissenschaftlichen Arbeitsforschung zuletzt umfassend diskutiert und mitunter als Prototyp für künftige Arbeitszusammenhänge im Allgemeinen angesehen. Diskursprägende Autoren wie Jeremy Rifkin (2014: 17) schrieben der plattformkoordinierten Arbeit unter dem Stichwort „Sharing Economy“ das Potenzial zu, eingespielte ökonomische Ressourcenverteilungen aufzubrechen und eine „dritte industrielle Revolution“ auszulösen. Die so gefasste Sharing Economy galt zwischenzeitig als Sinnbild für künftige onlinebasierte Arbeits- und Geschäftsmodelle, in denen plattformbetreibende Unternehmen nicht mehr über eigene Produktionsmittel verfügen, sondern diese von den Erwerbstätigen selbst eingebracht werden.

Studien zu erfolgreichen Unternehmen in der kommerziellen Sharing Economy bzw. Gig Economy (z. B. Uber, Delivero Hero, Just Eat Takeaway) führen die ambivalenten Effekte vor Augen, die mit der Plattformisierung von ortsunabhängigen und ortsgebundenen Arbeitszusammenhängen einhergehen: Plattformbetreibende Unternehmen akquirieren Dienstleister:innen mit der Aussicht auf eine eigenverantwortliche Arbeit, die in vielen Fällen keine besonderen formalen Qualifikationen voraussetzt. Die Arbeitsabläufe sind in modulare Einheiten gegliedert, die schrittweise abgearbeitet werden. Plattformunternehmen definieren sich dabei in der Selbstbeschreibung oft weniger als klassische Arbeitgeber:innen, sondern stellen ihre Position als Intermediäre heraus. Mit den damit einhergehenden eingeschränkten Arbeitnehmer:innenschutzrechten ist eine Verlagerung der unternehmerischen Risiken auf die Arbeitskraftanbietenden verknüpft, die nur über wenige Mitbestimmungsoptionen verfügen und zugleich mit ausdefinierten Plattformregulativen konfrontiert werden (Lücking 2019).

In den zurückliegenden Jahren sind vier Ausprägungen plattformzentrierter Arbeit hervorgetreten, die gemessen an den verfügbaren empirischen Daten zwar nach wie vor eine eher geringe Relevanz für den Arbeitsmarkt insgesamt haben (je nach Studie lag z. B. der Anteil an Plattformarbeiter:innen in Deutschland jüngst zwischen 0,6 und 4,8 Prozent; vgl. als Übersicht dazu: Ellmer et al. 2019), aber mitunter als „ein Experimentierfeld für Formen der algorithmischen und netzwerkbasierten Steuerung von Arbeit“ insgesamt angesehen werden (Tabelle 1):

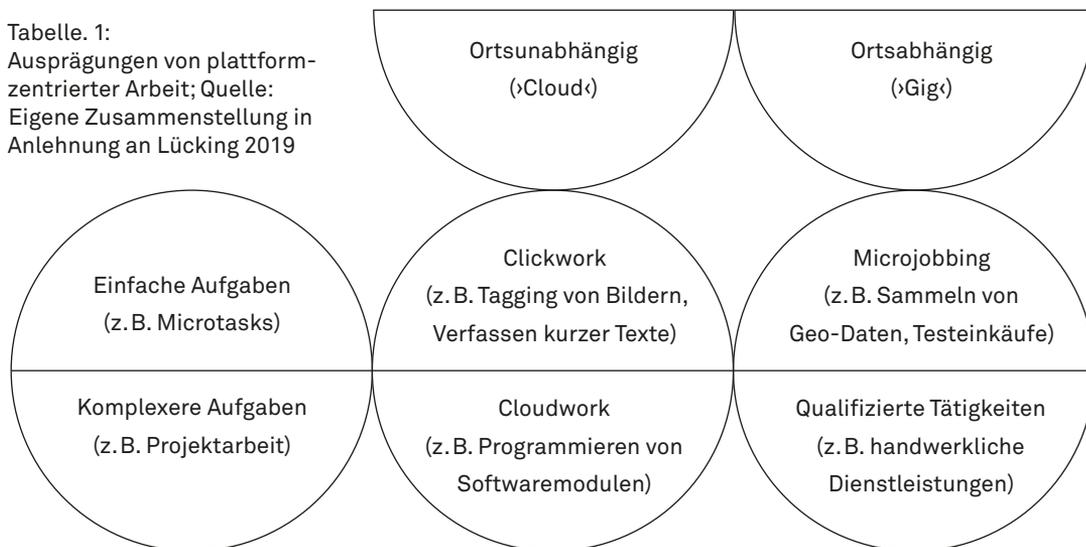
Plattformarbeit als neuer Prototyp?

Kennzeichen plattformzentrierter Arbeitsmärkte

Ausprägungen von Plattformarbeit

- *Clickwork* beschreibt einfache Tätigkeiten wie das Identifizieren von Objekten auf Bildern, das Testen von Websites, das Korrekturlesen von kurzen Texten oder niederschwellige Datenverarbeitungsaufgaben, die ortsunabhängig und direkt auf den vermittelnden Plattformen wie z. B. Clickworker oder Amazon Mechanical Turk durchgeführt werden. Die Dienstleister:innen erledigen ihre Aufträge auf freiberuflicher Basis.
- *Cloudwork* umschreibt komplexere ortsunabhängige Tätigkeiten, die ebenfalls über Plattformen vermittelt werden, darunter das Programmieren von Softwaremodulen und Websites, Datenanalysen oder kreative Aufgaben. Wichtige Plattformanbieter:innen in diesem Bereich sind z. B. Upwork, DesignCrowd, Fiverr und Gulp (betrieben durch Randstad). Diese Plattformen adressieren Freelancer, deren Auftragslage auch schon vor der Digitalisierung an die Marktentwicklung gekoppelt war, die aber nun auf diesen Plattformen in überregionaler Konkurrenz stehen.
- *Microjobbing* umfasst über Smartphone-Applikationen vermittelte ortsbezogene Kleinstaufgaben wie das Erstellen von anspruchlosen Fotos, das Evaluieren von Produktauslagen oder Öffnungszeiten, Geotagging oder Testeinkäufe, die für ein geringes Salär (oft 1 bis 3 Euro pro Job) ausgeführt werden. Führende Plattformanbieter:innen in diesem Bereich sind u. a. Streetspotr und appJobber, die mit ihren Services auch große Supermarktketten bedienen.
- *Gigwork* benennt ein weites Spektrum an komplexeren ortsgebundenen Arbeitsaufgaben – von handwerklichen Dienstleistungen (z. B. via MyHammer) über Mobilitätsdienstleistungen (z. B. Uber) bis hin zu Liefer- und Kurierfahrten (z. B. Lieferando.de, Gorillas), die über die Plattform an Solo-selbstständige oder an geringfügig Beschäftigte vermittelt werden. Nicht alle diese Tätigkeiten erfordern besondere Qualifikationen, aber sie gehen mit einem höheren Zeitaufwand als Microjobbing-Tasks einher.

Tabelle. 1:
 Ausprägungen von plattform-
 zentrierter Arbeit; Quelle:
 Eigene Zusammenstellung in
 Anlehnung an Lücking 2019



3 Das Beispiel Essenslieferdienste

Am Beispiel von Essenslieferdiensten, wie sie von *Just Eat Takeaway* (z.B. *Lieferando.de*) betrieben werden, lässt sich das ambivalente Zusammenspiel von Flexibilisierung, Standardisierung und Kontrolle in ortsgebundenen plattformzentrierten Arbeitszusammenhängen wie folgt darstellen (Schreyer/Schrape 2021):

Gigwork in der
 Plattformökonomie

- *Flexibilisierung*: Neben Möglichkeiten zur individuellen Einteilung der Arbeitszeiten ist der Zugang zu den jeweiligen plattformvermittelten Arbeitsmärkten niederschwellig angelegt. Arbeitskraft anbietende müssen neben ihrer allgemeinen Geschäftsfähigkeit in der Regel keine weiteren Qualifikationen angeben. Viele Arbeitsmittel (z.B. oft ein Smartphone mit Internetverbindung) müssen jedoch selbst eingebracht werden.
- *Standardisierung*: Dreh- und Angelpunkt für sämtliche Arbeitsabläufe ist die plattformspezifische Smartphone-Applikation. Die Liefervorgänge werden in standardisierte Schritte zerlegt, algorithmische Strukturen vergeben die Aufträge automatisiert, legen das Bestellvolumen auf die Fahrer:innen um, berechnen zu fahrende Strecken und schalten

sukzessive die jeweiligen Adressen frei. Die Möglichkeit, in den Ablauf einzugreifen, ist in der Regel nicht gegeben.

- *Kontrolle*: Arbeitszeiten, Interaktionen auf der Plattform und individuelle Leistungsdaten werden in unternehmensspezifischen Datenbanken gespeichert, integriert und in grafischer Form an die Fahrer:innen zurückgespielt. Die auf der Plattform aggregierten Daten eröffnen ab bestimmten Schwellenwerten höhere Entlohnungs- bzw. Aufstiegschancen. Ein situativer Verhandlungsspielraum ist kaum vorhanden.

Eine derartige plattformbasierte Integration technisch vermittelter Leistungen (u. a. eine vollautomatisierte Koordination und eine prozessumspannende Ablaufkontrolle) ist auf dem Feld der Arbeit erst seit der lebensweltlichen Verbreitung vernetzter mobiler IT-Endgeräte umsetzbar. Sie ermöglicht neben neuen Spielarten dezentral ausgeführter Plattformarbeit auch in der ortsgebundenen Arbeit so zuvor nur in Ausnahmefällen denkbare volatile Arbeitsverhältnisse, die durch ein geringeres Maß an wechselseitiger Bindung zwischen Unternehmen und Personal gekennzeichnet sind. Vor allen Dingen für Arbeitsmärkte, die vorrangig durch geringfügig Beschäftigte bespielt werden (z. B. im Bereich der Lieferdienste), lässt sich der strukturierende und kanalisierende Einfluss soziotechnischer Plattformen, die überwiegend proprietär betrieben werden, kaum überschätzen (vgl. dazu ferner: Pfeiffer 2021). Vor diesem Hintergrund hat sich mit Blick auf die plattformzentrierte Koordination von Arbeit ein reger sozialwissenschaftlicher Diskurs um Möglichkeiten der ordnungspolitischen Einhegung entwickelt, die einer Erosion „des historisch gewachsenen Systems der Regulation von Arbeit“ (Boes et al. 2014: 71) entgegenwirken können.

Einfluss
proprietärer
Plattformen

4 Plattformkoordinierte Arbeit in klassischen Betrieben

Auf korrespondierende Weise zeigt sich in klassischen betrieblichen Zusammenhängen das Wechselspiel von ermöglichenden und kanalisierenden Effekten der fortschreitenden Informatisierung, Automatisierung und Plattformisierung der Arbeitskoordination (Meyer et al. 2020): Einerseits eröffnet die durch die COVID-19-Pandemie noch einmal beschleunigte digitale Transformation der Arbeitswelt neue Optionen zur *Flexibilisierung und Individualisierung* von Beschäftigungsverhältnissen, womit für viele Arbeitnehmer:innen u. a. in der Büroarbeit) so zuvor kaum vorstellbare zeitliche und örtliche Spielräume wie auch neue Notwendigkeiten zur Selbstorganisation und -kontrolle einhergehen. Auf der anderen Seite bietet die Digitalisierung allerdings ebenso zuvor kaum realisierbare Möglichkeiten zur marktorientierten Arbeitsverteilung sowie zur hochauflösenden *Standardisierung und Kontrolle* von Arbeitsabläufen, die fallweise in einer erneut intensivierten Rationalisierung münden können.

In einer exemplarischen empirischen Untersuchung haben Andreas Boes et al. (2018) auf der Grundlage von 13 qualitativen Fallstudien zu Unternehmen aus unterschiedlichen Wirtschaftsbereichen (u. a. IT, Finanzdienstleistungen, Maschinenbau, Metall- und Elektroindustrie) in diesem Zusammenhang herausgearbeitet, dass mit der Einführung von informationstechnischen Plattformen sowie der damit erhofften ‚Agilisierung‘ der Arbeitskoordination in der Praxis ambivalente Effekte für die Beschäftigten einhergehen: Vor allen Dingen Kopfarbeiter:innen erlebten durch die betriebliche Implementation digitaler Arbeitsplattformen und agiler Softwaretools (z. B. Slack) auf der einen Seite zwar eine Intensivierung des arbeitsbezogenen kommunikativen Austauschs und neue Freiräume zur Selbstorganisation. Auf der anderen Seite gingen damit aber nach Einschätzung der Befragten auch erweiterte „Eingriffs- und Steuerungsmöglichkeiten für das Management“ und eine Zunahme von (wechselseitiger) Kontrolle einher. Boes et al. (2018: 205f.) kommen dahingehend zu dem Schluss, dass die technikvermittelte Plattformisierung bzw. Modularisierung von Organisationsabläufen und die automatisierte Protokollierung von Arbeitsprozessen für Erwerbstätige zu einem „erheblichen Anstieg von Belastungen“ führen kann.

Ambivalenzen der Plattformisierung in Betrieben

Freiräume und erhöhte Kontrollpotenziale

5 Arbeitspolitische Implikationen

Informationstechnische Plattformen und Infrastrukturen zeichnen sich durch ermöglichende, kanalisierende und kontrollierende Eigenheiten aus. In innerbetrieblichen wie außerbetrieblichen Arbeitszusammenhängen prägen die in die technischen Plattformarchitekturen eingeschriebenen sozialen Regelsetzungen die Koordinationsmuster und legen spezifische Handlungs- und Kommunikationsweisen nahe, während sie andere erschweren. Dies betrifft auch die Möglichkeiten zur Interaktion und kollektiven Abstimmung unter den Erwerbstätigen. Wie sich die Plattformisierung von Arbeit fallweise ausspielt, hängt insoweit stets von den *konkreten Implementationsweisen* ab: Eine offen gehaltene informationstechnische Infrastruktur, die von ihren Nutzer:innen angepasst werden kann und mehrere Kommunikationskanäle anbietet, eröffnet vielfältige Möglichkeiten zur Mitbestimmung. Geschlossene, vollautomatisierte und proprietär betriebene Koordinationsplattformen, wie sich in der Gig Economy finden, können hingegen den Austausch unter den Arbeitnehmenden behindern und in einem Kontrollregime münden.

Spielräume der
Implementation

Vor diesem Hintergrund sollten die in den jeweiligen Arbeitskontexten in Anschlag gebrachten Plattformstrukturen ein zentraler Ansatzpunkt der *arbeitspolitischen Intervention* werden. Eine solche Intervention setzt allerdings nicht nur ein grundlegendes Verständnis informationstechnischer Strukturen und algorithmischer Programmierweisen sowie ein hinreichendes Wissen über deren Folgen und Rückwirkungen voraus, sondern ebenso belastbare rechtlich-regulative Instrumente, die gewerkschaftlichen Vertreter:innen eine Einsicht in die Funktionsweisen kritischer informationstechnischer Strukturen garantieren. Die arbeitspolitische Intervention sollte überdies bereits in den frühen technischen Entwicklungs- und Prototypisierungsphasen ansetzen – denn einmal implementierte Plattformstrukturen lassen sich nur noch mit viel Aufwand reformieren.

Frühe arbeits-
politische
Intervention

Über plattformspezifische Regulierungsmaßnahmen hinaus, die bislang in Deutschland, Österreich und der Schweiz zumeist nachgelagert erfolgen, werden vor diesem Hintergrund übergreifende *rechtlich-regulative Rahmensetzungen* notwendig, die in einer gesellschaftspolitischen Debatte ausgehandelt werden sollten. Diese Rahmensetzungen könnten in einer gesetzlich festgeschriebenen Offenlegung informationstechnischer Strukturen oder in einer Zertifizierung der zugrundeliegenden Algorithmen durch arbeitspoli-

Regulative
Rahmensetzungen

tische Evaluationsstellen bestehen. Als solche Evaluationsstellen könnten auch Gewerkschaften fungieren, falls diese ihren Kompetenzkatalog um das dafür notwendige informationstechnische Bewertungswissen erweitern. Unter dieser Voraussetzung sähen sich Gewerkschaften in der Lage, auch in Arbeitskontexten konfliktvermittelnd aktiv zu werden, die über ihre angestammten Handlungsbereiche hinausgehen, wie das mit Blick auf die Arbeitskoordination in der Gig Economy zum Teil bereits geschehen ist.

Mitbestimmung in plattformvermittelten Arbeitszusammenhängen setzt zudem Anpassungen in den jeweils länderspezifischen Betriebs- und Arbeitsverfassungsgesetzen voraus, die derzeit auf viele Dynamiken der Plattformisierung der Arbeitskoordination noch keine belastbaren Antworten bieten. So lassen sich beispielsweise im deutschen Betriebsverfassungsgesetz entsprechende Begrifflichkeiten wie ‚Betriebsstätte‘ nicht mehr umstandslos mit den Gegebenheiten volatiler digitaler Arbeit in Bezug setzen, weshalb Neugründungen von Betriebsräten in der Gig Economy in Deutschland immer wieder verhindert werden konnten. Zum anderen sehen etablierte Betriebs- und Arbeitsverfassungsgesetze zwar vor, dass Arbeitnehmer:innen über die Einführung neuer technischer Anlagen und Verfahren rechtzeitig unterrichtet werden müssen. Dieser Informationspflicht kommen die Unternehmen allerdings oft nicht angemessen nach – insbesondere, wenn es nicht um die Einführung neuer Hardware, sondern um die Neugestaltung von Software und neuer algorithmischer Logiken geht. Davon abgesehen reicht es aus Sicht der Arbeitnehmer:innenseitiger Interessenvertretungen mit Blick auf die strukturierenden Eigenheiten digitaler Plattformen nicht aus, lediglich über die Einführung veränderter Koordinations- und Automatisierungslogiken informiert zu werden. Vielmehr müsste die gesamte Technikentwicklung und -gestaltung ein gesetzlich garantierter *Gegenstand betrieblicher Mitbestimmung* werden.

Mitbestimmung

Schrape, PD. Dr., Jan-Felix

Jg. 1979, Senior Researcher am Lehrstuhl für Organisations- und Innovationssoziologie des Instituts für Sozialwissenschaften der Universität Stuttgart; Gastprofessor am Lehrstuhl für Technik und Gesellschaft der RWTH Aachen (2021–2022). Arbeitsschwerpunkte: Technik-, Arbeits-, Innovations- und Organisationssoziologie.

Quellenverzeichnis

Boes, Andreas/Kämpf, Tobias/Langes, Barbara/Lühr, Thomas (2018): ›Lean‹ und ›agil‹ im Büro. Neue Organisationskonzepte in der digitalen Transformation und ihre Folgen für die Angestellten. Bielefeld.

Boes, Andreas/Kämpf, Tobias/Langes, Barbara/Lühr, Thomas/Steglich, Steffen (2014): Cloudworking und die Zukunft der Arbeit. Kassel.

Böschen, Stefan/Grunwald, Armin/Krings, Bettina-Johanna/Rösch, Christine: (Hrsg.): Technikfolgenabschätzung. Handbuch für Wissenschaft und Praxis. Baden-Baden.

Ellmer, Markus/Herr, Benjamin/Klaus, Dominik/Gegenhuber, Thomas (2019): Platform Workers Centre Stage! Taking Stock of Current Debates and Approaches for Improving the Conditions of Platform Work in Europe. Forschungsförderung HBS Working Paper 140. Düsseldorf.

Lücking, Stefan (2019): Arbeiten in der Plattformökonomie. HBS Forschungsförderung Report 5. Düsseldorf.

Meyer, Uli/Schaupp, Simon/Seibt, David (Hg.) (2020): Digitalization in Industry. Between Domination and Emancipation. Cham.

Mückenberger, Ulrich (1985b): Die Krise des Normalarbeitsverhältnisses. In: Mitteilungsblatt Arbeit und Betrieb 11/12, 3–36.

Papsdorf, Christian (2019): Digitale Arbeit: Eine soziologische Einführung. Frankfurt (Main).

Pfeiffer, Sabine (2021): Digitalisierung als Distributivkraft. Über das Neue am digitalen Kapitalismus. Bielefeld.

Pfeiffer, Sabine (2018): Technisierung von Arbeit. In: Böhle, Fritz/Voß, Gerd-Günter / Wachtler, Günther (Hrsg.): Handbuch Arbeitssoziologie. Bd. 1. 2. Auflage. Wiesbaden, 321–358.

Rifkin, Jeremy (2014): Zero Marginal Cost Society. New York.

Schrage, Jan-Felix (2021): Digitale Transformation. Reihe: Einsichten. Themen der Soziologie. Bielefeld.

Schreyer, Jasmin/Schrage, Jan-Felix (2021): Digitale Plattformen in kommerziellen und gemeinwohlorientierten Arbeitszusammenhängen. HBS Study 460. Düsseldorf.

Trinczek, Rainer (2011): Überlegungen zum Wandel von Arbeit. In: WSI-Mitteilungen 64(11), 606–614.

Voß, Gerd-Günther/Pongratz, Hans J. (1998): Der Arbeitskraftunternehmer. Eine neue Grundform der Ware Arbeitskraft? In: Kölner Zeitschrift für Soziologie und Sozialpsychologie 50 (1), 131–158.



Wer ist schuld, wenn Algorithmen irren? Entscheidungs-automatisierung, Organisationen und Verantwortung

Abstract	49
1 Einleitung	49
2 Rechenschaft und Verantwortung in Organisationen	52
3 Entscheidungsfindung und algorithmische „Unterstützung“	53
4 Vershobene und veränderte Verantwortung	56
5 Leitfaden und Verantwortungszuordnung mittels VERA	58
6 Grenzen und Ausblick	64
7 Fazit	66
Danksagung	67
Über den Autor	68

Abstract

Algorithmenunterstützte Entscheidungsfindung (*algorithmic decision support*, ADS) kommt in verschiedenen Kontexten und Strukturen vermehrt zum Einsatz und beeinflusst in diversen gesellschaftlichen Bereichen das Leben vieler Menschen. Ihr Einsatz wirft einige Fragen auf, unter anderem zu den Themen Rechenschaft, Transparenz und Verantwortung. Im Folgenden möchten wir einen Überblick über die wichtigsten Fragestellungen rund um ADS, Verantwortung und Entscheidungsfindung in organisationalen Kontexten geben und einige offene Fragen und Forschungslücken aufzeigen. Weiters beschreiben wir als konkrete Hilfestellung für die Praxis einen von uns entwickelten Leitfaden samt ergänzendem digitalem Tool, welches Anwender:innen insbesondere bei der Verortung und Zuordnung von Verantwortung bei der Nutzung von ADS in organisationalen Kontexten helfen soll.

1 Einleitung

Algorithmenunterstützte Entscheidungsfindung (*algorithmic decision support*, ADS; manchmal auch algorithmische Entscheidungsfindung, *algorithmic decision-making*, ADM) kommt in immer mehr Anwendungsbereichen zum Einsatz, z. B. im Personalwesen (Dreyer und Schulz 2019, S. 7), beim Zugang zu Sozialleistungen oder Krediten (O’Neil 2016), in Polizeiwesen (Bennett Moses und Chan 2018; Ferguson 2017) und Justiz (Christin 2017), mit möglicherweise schwerwiegenden Folgen für Betroffene. Diese Art der Technologie wird oft als besonders objektiv und neutral empfunden, weil für computerunterstützte Entscheidungen mehr Rechenleistung und Dateninput herangezogen werden kann als durch rein menschliche Entscheidungen (Christin 2017; Gillespie 2016). Die Verwendung von ADS-Systemen kann Vorteile mit sich bringen, etwa in der medizinischen Diagnostik (Castellucia und Métayer 2019), gefährdet Gleichbehandlung aber auf verschiedene Arten, von Verzerrungen in den zugrundeliegenden Daten (Barocas und Selbst 2016; Hao 2019) oder durch implizite (Friedman und Nissenbaum 1996) und explizite Modellannahmen (Bennett Moses und Chan 2018, S. 809ff.) bis hin zur Erklärung und Darstellung der „Empfehlungen“ für Endbenutzer:innen (Goddard, Roudsari und Wyatt 2011). Versprechungen von höherer Effizienz und „besseren“ Entscheidungen gehen jedoch oft mit neuen Problemen wie Intransparenz und un-

Algorithmenunterstützte oder algorithmische Entscheidungsfindung

Gefahr verschiedener Arten von Ungleichbehandlung

klarer Verantwortungsverteilung einher. In vielen Kontexten bleibt die Letztverantwortung für Entscheidungen unter Zuhilfenahme von Algorithmen formal bei den menschlichen Entscheider:innen, auch wenn diese in den meisten Fällen kaum oder gar kein Mitspracherecht dabei haben, ob und in welcher Form sie ADS-Systeme für ihre Arbeit nutzen möchten. Dies führt zu Interessenskonflikten zwischen verschiedenen Hierarchieebenen innerhalb von Organisationen.

Interessenskonflikte zwischen verschiedenen Hierarchieebenen einer Organisation

Es gibt bisher vergleichsweise wenig Forschung zu den Auswirkungen von ADS auf die Arbeitsbedingungen der menschlichen Entscheider:innen. Viele leitende Angestellte und öffentliche Bedienstete mit Entscheidungsbefugnissen (im Folgenden Entscheider:innen) werden sich in der nahen Zukunft in Situationen wiederfinden (oder tun dies bereits jetzt), in denen sie nach der Einführung von ADS-Systemen nicht nur weiterhin verantwortlich für die Folgen der von ihnen getroffenen Entscheidungen bleiben, während von ihnen erwartet wird, dass sie diese effizienter (Vieth und Wagner 2017) und schneller treffen (Zweig, Fischer und Lischka 2018, S. 15) als ohne ADS-Nutzung, sondern sie darüber hinaus auch die vom Computer „vorgeschlagenen“ Entscheidungen kritisch prüfen müssen, obwohl sie dessen Funktionsweise nicht genau kennen.

Die zunehmende Verwendung von Algorithmen bei der Entscheidungsfindung verschärft auch das Risiko von Diskriminierung und Ungleichbehandlung sowohl auf einer individuellen Ebene (da die erwartete schnellere Entscheidungsfindung den Entscheider:innen weniger Zeit einräumt, um mögliche diskriminierende Auswirkungen ihrer Entscheidungen zu reflektieren) als auch auf institutioneller Ebene, da Verzerrungen und Unfairness eines ADS-Systems deutlich mehr Menschen betreffen können als Vorurteile einer einzelnen Entscheider:in. In einer früheren Arbeit zu algorithmischen Systemen in Polizeiwesen und Justiz (Adensamer und Klausner 2021a) haben wir hierfür die Unterscheidung zwischen „*retail bias*“ und „*wholesale bias*“ eingeführt. Ersteres bezeichnet die diskriminierenden Folgen von menschlichen Einzelentscheidungen; ihre Auswirkungen bleiben, bei aller Problematik, dennoch im möglichen Betroffenenkreis sowie zeitlich und örtlich eingeschränkt. Zweiteres bezieht sich darauf, dass bei automatisierter Entscheidungsfindung eine sehr viel größere Zahl an Menschen (potenziell um Größenordnungen mehr) denselben Biases unterworfen ist. All das stellt zusätzliche und neuartige Hindernisse für Gerechtigkeit, Inklusion und Schutz vor Diskriminierung in Entscheidungsprozessen dar. Diese Herausforderungen betreffen in besonderer Wei-

Verschärfung des Risikos von Diskriminierung und Ungleichbehandlung: *retail bias* vs. *wholesale bias*

se Arbeitnehmer:innen, die ADS-Systeme verwenden müssen. In diesem Artikel befassen wir uns darum mit der Frage der Arbeitsbedingungen von Entscheider:innen bei Nutzung von ADS-Systemen und insbesondere mit der Frage der organisationalen Verantwortung für ihre Einführung und Nutzung.

Vorerst gilt es jedoch die Frage der Begrifflichkeit zu klären. In der englischsprachigen Forschungsliteratur werden *accountability* und *responsibility* oft nicht klar voneinander abgegrenzt und unter Verweis auf den jeweils anderen Begriff definiert (McGrath und Whitty 2018). Auf Deutsch kommt zusätzlich die Übersetzungsproblematik hinzu wir haben uns in diesem Text (meistens) für die Übersetzungen „Rechenschaft“ (*accountability*) und „Verantwortung“ bzw. „Verantwortlichkeit“ (*responsibility*) entschieden. Die deutschen Ausdrücke sind für unser Sprach- und Begriffsgefühl allerdings nicht bedeutungsgleich mit den englischen. Die Begriffsdiskussion ist weiters auch alles andere als abgeschlossen für eine ausführlichere Begriffsdiskussion möchten wir an dieser Stelle nur kurz auf Abschnitt 1 unseres englischsprachigen Artikels (Adensamer, Gsenger und Klausner i. E.) verweisen.

Wir analysieren im Folgenden die verschiedenen Hierarchieebenen von Organisationen und die damit verknüpften Grade an Handlungsfähigkeit und Verantwortung menschlicher Entscheider:innen in Bezug auf die Einführung, Anwendung und Evaluierung von ADS-Systemen. Wir beschreiben dann eine Methode zur Zuweisung von Verantwortung und Rechenschaft an die involvierten Akteur:innen und Rollen; dies soll dabei helfen, Diskrepanzen und Probleme in der Zuordnung zu identifizieren und zu vermeiden. Zur praktischen Lösung der von uns herausgearbeiteten Probleme präsentieren wir schlussendlich einen Leitfaden für Arbeitgeber:innen und Betriebsräte (Adensamer und Klausner 2021b), der die Problemfelder zu Verantwortungsfragen bei der Nutzung von ADS-Systemen niedrigschwellig erläutert. Außerdem wird das ergänzende digitale Tool „VERA“ (1) vorgestellt, welches bei der Verantwortungszuordnung assistieren und mögliche Interessenskonflikte und Verantwortungslücken aufzeigen soll. Sowohl der Leitfaden als auch VERA wurden in Zusammenarbeit mit der Arbeiterkammer Wien sowie in enger Absprache mit Expert:innen aus Wissenschaft und Industrie entwickelt, um ihre praktische Anwendbarkeit und Nützlichkeit sicherzustellen.

Leitfaden für Arbeitgeber:innen und Betriebsräte sowie digitales Tool „VERA“

(1) <https://vera.arbeiterkammer.at> (letzter Zugriff: 31.07.2021)

2 Rechenschaft und Verantwortung in Organisationen

Rechenschaft und Verantwortung von Algorithmen (*algorithmic accountability*) hat als Forschungsgebiet in den letzten Jahren zunehmend an Bedeutung gewonnen. Die Forschung stützt sich einerseits auf die Verantwortlichkeitstheorie aus den Organisationswissenschaften, andererseits auf Konzepte zu Fairness, Verantwortlichkeit und Transparenz in der Informatik. Wieringa (2020) bietet einen systematischen und strukturierten Überblick über den aktuellen Forschungsstand. Sie baut ihren Begriff von Rechenschaft (*accountability*) auf der weit verbreiteten Definition und Ontologie von Bovens (2007) auf, die *accountability* als eine soziale Beziehung mit (in Wieringas Darstellung) fünf Komponenten versteht: (1) ein:e Akteur:in, (2) ein Forum (dem gegenüber die Akteur:in Rechenschaft ablegen muss), (3) die Beziehung zwischen Akteur:in und Forum, (4) der Inhalt, die Parameter und der Rahmen der Rechenschaft und zuletzt (5) die Art der möglichen Konsequenzen für die Akteur:in. Nach ihrer Analyse ist *algorithmic accountability* als „vernetzte Rechenschaft für ein soziotechnisches algorithmisches System gemäß der verschiedenen Abschnitte im Lebenszyklus des Systems“ (2) zu verstehen, in der mehrere Akteur:innen verpflichtet sind, ihre Handlungen in Zusammenhang mit dem fraglichen System „zu erklären und zu rechtfertigen“ (Wieringa 2020, S. 10). Unser Zugang ist insbesondere im Einklang mit ihrem Befund, dass Verantwortung zwischen diesen verschiedenen Akteur:innen verteilt ist und dass es Rechenschaft sowohl innerhalb der Organisation wie auch gegenüber externen Foren bedarf. (Wir kehren zu dieser Thematik in Abschnitt 5 zurück und bauen dort auf Wieringas Arbeit auf.)

In den Organisationswissenschaften gibt es relativ wenig Forschung, die sich mit der Frage der organisationalen Verantwortung bei der Nutzung von ADS (oder Algorithmen allgemein) befasst. Exemplarisch genannt seien an dieser Stelle neben Faraj, Pachidi und Sayegh (2018) und Kellogg, Valentine und Christin (2020) insbesondere Moradi und Levy (2020). Moradi und Levy stellen fest, dass sich die Auswirkungen der Algorithmisierung insgesamt als Risikoverlagerung von den Unternehmen und Arbeitgeber:innen hin zu den Angestellten verstehen lassen. Hierbei werden bestehende unwirtschaftliche Strukturen und Prozesse nicht durch technische Innovation reduziert oder ausgeräumt, sondern nur die daran geknüpften Risiken und Kosten zu den Angestellten verlagert (Moradi und Levy 2020, S. 278).

Algorithmic accountability:
vernetzte Rechenschaft mehrerer Akteur:innen für ein soziotechnisches algorithmisches System

Die Forschung und der öffentliche Diskurs zu den Auswirkungen von ADS auf die Entscheidungsfindung und die Arbeitsbedingungen, zu Risiko und Nutzen und zu grundlegenden ethischen Fragestellungen rund um den Gebrauch von ADS (wie etwa Christen et al. 2020; Fjeld et al. 2020; Zweig, Fischer und Lischka 2018) beleuchten somit zwar sehr wohl auch Verantwortung, Haftbarkeit und Rechenschaft, in den meisten Fällen wird aber nicht genauer auf die Schwierigkeiten eingegangen, die geteilte Kompetenzen in komplexen und oft stark hierarchisch strukturierten Organisationen und deren Auswirkungen auf die (inner-)organisationale Verantwortung mit sich bringen. In diesem Kontext ist auch das Problem der vielen Hände (*many-hands problem*; siehe z. B. Poel, Royackers und Zwart 2015; Thompson 2017) zu nennen, bei dem Verantwortung einer Gruppe an Akteur:innen kollektiv zugeschrieben werden kann, die individuelle Zuordnung aber unmöglich bleibt. In unserer Konzeption von organisationaler Verantwortung entspricht dies der Existenz von Verantwortungslücken, auf die wir in Abschnitt 4 genauer eingehen.

Bisherige Forschung legt wenig Augenmerk auf Schwierigkeiten geteilter Kompetenzen und die Auswirkungen auf organisationale Verantwortung.

3 Entscheidungsfindung und algorithmische „Unterstützung“

Wir wenden uns nun den praktischen Auswirkungen von ADS-Systemen auf Entscheidungsfindungsprozesse zu und insbesondere den Folgen für die Entscheider:innen, die solche Systeme verwenden. Die Einführung von ADS kann zu Interessenskonflikten zwischen verschiedenen Hierarchieebenen einer Organisation führen und gravierende Auswirkungen auf das Arbeitsleben von Angestellten haben, die zugleich oft wenig Einfluss auf Veränderungen ihrer Arbeitsumgebung durch Automatisierung ausüben können. Zentral ist hierbei, dass die Entscheidung über die Einführung von Automatisierung nicht von den davon unmittelbar betroffenen Angestellten gefällt wird (vgl. Faraj, Pachidi und Sayegh 2018, S. 366f.). Diese Veränderung der Arbeitsprozesse kann allerdings die Arbeitsbedingungen und die Erwartungen an die von den Entscheider:innen geleistete Arbeit verändern, manchmal sogar grundlegend: Oft ohne ausreichend darin geschult zu werden oder genügend Dokumentation zur Verfügung zu haben, müssen sie nun die automatisch erstellten Entscheidungs„vorschläge“ eines algorithmischen Systems kritisch beurteilen und akzeptieren oder verwerfen – und dabei zumeist mehr Fälle als bislang in

Entscheidung über Einführung von Automatisierung wird meist nicht von unmittelbar betroffenen Angestellten gefällt.

(2) Dieses sowie alle weiteren wörtlichen Zitate aus englischsprachigen Quellen in diesem Beitrag wurden durch die Autor:innen übersetzt.

derselben Zeit bearbeiten. Weiters kann die Nutzung von ADS das Potenzial für diskriminierende Entscheidungen massiv erhöhen (durch den von uns beschriebenen „*wholesale bias*“, siehe Abschnitt 1 sowie Adensamer und Klausner 2021a). (3)

Die von ADS-Systemen vorgeschlagenen Entscheidungen beeinflussen die menschliche Entscheidungsfindung und können verschiedenartige Effekte und Verhaltensweisen hervorrufen wie *complacency*, Aversion und diverse Arten von kognitiven Verzerrungen. *Complacency* (Duldsamkeit oder Gefälligkeit) beschreibt beispielsweise den Effekt, dass bei Beobachtung eines automatisierten Systems oft zu spät oder gar nicht eingeschritten wird, obwohl die Situation dies eigentlich erfordern würde (Zerilli et al. 2019). *Aversion* (Widerwillen) beschreibt die Tendenz zur Verweigerung bzw. Nichtnutzung gewisser Systeme, obwohl ihre Anwendung insgesamt positive Effekte hätte (Dietvorst, Simmons und Massey 2015; Dietvorst, Simmons und Massey 2018). Der für unseren Kontext zentrale Automatisierungsbias, welcher sich spezifisch auf kognitive Verzerrungen bei menschlichem Umgang mit automatisierten Systemen bezieht, bezeichnet die Situation, dass ein ADS-System falsche oder suboptimale Vorschläge macht, die menschlichen Entscheider:innen diese aber nicht (ausreichend) prüfen oder hinterfragen (Parasuraman und Manzey 2010).

Erklärungsmuster für den Automatisierungsbias liefert die Sozialpsychologie (unter dem Begriff *cognitive miser*, in etwa „geistiger Geizhals“): Menschen bevorzugen (unabhängig von Intelligenz oder Bildung) einfachere Erklärungen und Lösungen und sind bemüht, mentale Ressourcen möglichst sparsam einzusetzen (Clarke 2007; Dunn und Risiko 2019). Dieser ökonomische Umgang mit der eigenen mentalen Energie führt dazu, dass die von ADS-Systemen angebotenen „Abkürzungen“ häufig dankbar angenommen werden, und in der Folge eben zum Automatisierungsbias führen (Parasuraman und Manzey 2010).

Die beschriebenen Effekte können fallweise auch gleichzeitig, gegenläufig und widersprüchlich auftreten: Entscheider:innen versuchen einerseits oft, ADS-Systeme zu umgehen, sofern das möglich ist (Christin 2017), bei als zu gering empfundener Handlungsfähigkeit akzeptieren sie algorithmische Entscheidungsvorschläge aber wiederum oft, ohne sie zu überprüfen (Schäufele 2017). Ihre individuellen Wahrnehmungen von und Einstellungen gegenüber ADS-Systemen können dabei einen maßgeblichen Einfluss auf ihr Verhalten im praktischen Umgang damit haben.

Einfluss auf menschliche Entscheidungsfindung: *complacency*, Aversion und diverse Arten kognitiver Verzerrungen

Effekte können gleichzeitig, gegenläufig und widersprüchlich auftreten.

Weiters sind die Wahrnehmung und Anwendung solcher Systeme stark von den Erwartungshaltungen geprägt (Burton, Stein und Jensen 2020). Die Aversion sinkt mit steigender Kenntnis über die Funktionsweise des Systems (Yeomans et al. 2019), und selbst geringfügige Einflussmöglichkeiten auf den Output des Systems reduzieren Aversion (Dietvorst, Simmons und Massey 2018).

Ebenfalls anzumerken ist, dass wir Menschen zwar häufig kognitiven Verzerrungen, Heuristiken und geistigen Abkürzungen erliegen, uns dessen aber oft zumindest teilweise bewusst sind (De Neys, Rossi und Houdé 2013). Es gibt durchaus auch Positivbeispiele, in denen die Entscheider:innen überraschende oder unpassende Vorschläge hinterfragen und ihr eigenes Urteilsvermögen einsetzen (siehe z. B. De-Arteaga, Fogliato und Chouldechova 2020 für eine ausführliche Diskussion eines konkreten Falls). Umgekehrt hat allerdings Kolkman (i. E.) durch ethnographische Arbeit zur praktischen Anwendung verschiedener algorithmischer Modelle feststellen müssen, dass Transparenz von Algorithmen selbst für Expert:innen (also Menschen, die beruflich mit solchen Modellen arbeiten) „bestenfalls fraglich und schlimmstenfalls unerreichbar“ ist.

Zuletzt wollen wir noch darauf hinweisen, dass die Forschung zur Entwicklung der Entscheidungsqualität durch die Einführung von ADS nach wie vor zu wünschen übrig lässt. Obwohl eines der häufigsten Argumente für den Einsatz von ADS verbesserte Entscheidungsfindung ist, gibt es laut einer (bereits etwas älteren) Studie von Skitka et al. „auffallend wenige“ Studien dazu, ob die Einführung von ADS die Fehlerrate reduziert; im Gegenteil räumt die bisherige Faktenlage zumindest die Möglichkeit ein, dass es nicht zu einer reinen Reduktion menschengemachter Fehler kommt, sondern menschliche Fehler zumindest teilweise durch neuartige, von Algorithmen verursachte Fehler ersetzt werden (Skitka, Mosier und Burdick 1999, S. 992).

Die Forschungslage zur Entwicklung der Entscheidungsqualität durch die Einführung von ADS ist dürftig.

(3) Siehe auch Loi und Spielkamp (2021) für einen Überblick über KI-Regulierung im öffentlichen Sektor; ihre einführende Analyse der Herausforderungen bzgl. Rechenschaft und Verantwortung kommt unabhängig von uns zu sehr ähnlichen Schlüssen.

4 Vershobene und veränderte Verantwortung

Die Einführung von ADS-Systemen in Arbeitsumgebungen kann substantielle Auswirkungen auf die Verteilung von Verantwortung, Handlungsfähigkeit und Haftbarkeit haben. Oft wird Macht (und manchmal auch Verantwortung) von den Angestellten, die bislang Entscheidungen ohne ADS getroffen haben, hin zu externen Entwickler:innen verschoben, die weniger Einblick in und Wissen über die tagtägliche Anwendung und die Auswirkungen der Entscheidungen haben (vgl. Moradi und Levy 2020, S. 278 für weitere Ausführungen zu Risikoverschiebung sowie Wagner 2019 zu Haftung und Kriterien für relevante Handlungsfähigkeit in quasi-automatisierten Systemen). Diese Verschiebungen können zu Problemen führen, wenn keine angemessenen Begleitmaßnahmen getroffen werden.

Erstens kann die Verschiebung von Aufgaben und Verantwortung zu Angestellten problematisch sein, wenn diese nicht ausreichend dafür geschult werden. Die Aufgabenbeschreibung von Sachbearbeiter:innen kann sich durch die Einführung eines ADS-Systems drastisch verändern, etwa von der Beurteilung einzelner Fälle hin zur Kontrolle und kritischen Prüfung der Vorgänge in einem Computersystem. Bei der Einführung von ADS in einer Organisation ist es daher unumgänglich, die Arbeitsbedingungen der Entscheider:innen mitzubedenken, die jetzt durch den Algorithmus „unterstützt“ werden sollen. Ihre Aufgabenbeschreibung verändert sich höchstwahrscheinlich (zumindest stillschweigend und implizit), aber ihre Kompetenzen, ihr Wissen und ihre Ausbildung entwickeln sich nicht automatisch im gleichen Tempo und zur gleichen Zeit. Wenn die betroffenen Angestellten keine oder nicht ausreichend Unterstützung bekommen, können sie sich in der unhaltbaren Position wiederfinden, dass der praktische Einfluss auf die Entscheidungen, die sie treffen (oder eben meist der Algorithmus für sie), schwindet, während gleichzeitig ihre Verantwortung für die Entscheidungen bestehen bleibt.

Zweitens können Angestellte nicht die Letztverantwortung für Entscheidungen zugeschrieben bekommen, wenn sie nicht berechtigt sind, das ADS-System zu überstimmen oder notwendige Änderungen im ADS-System vorzunehmen oder zu veranlassen. Generell können Angestellte von ihren Arbeitgeber:innen verantwortlich gehalten werden, wenn sie (unrechtmäßig) diskriminierende Entscheidungen vornehmen. Wenn allerdings ein algorithmisches System diskriminiert und Angestellte von ihren Arbeitgeber:innen

Problem 1:
unzureichende
Schulung und
Unterstützung der
Entscheider:innen

Problem 2:
mangelnde Befugnisse
der Entscheider:innen
über das ADS-System

Wer ist schuld, wenn Algorithmen irren?

angewiesen werden, ADS in ihren Arbeitsabläufen zu nutzen, sie aber nicht ausreichend befähigt werden, das Computermodell, die zugrundeliegenden Daten etc. zu verstehen und zu hinterfragen, ist die Situation grundlegend anders. Aus unserer Sicht können Entscheider:innen in solchen Fällen nicht verantwortlich gemacht werden, da sie keine signifikante Mitsprache beim Wirken des algorithmischen Systems haben. Die zugewiesene Verantwortung kann niemals das Maß der praktischen Entscheidungsfähigkeit übersteigen.

Drittens kann es dazu kommen, dass die Zuordnung von Verantwortung nicht oder nicht ausreichend vorgenommen wird, so dass es zu Verantwortungslücken kommt. Es gibt eine Vielzahl von Rollen und involvierten Personen(Gruppen), bei denen die Verantwortung für teilweise algorithmisch, teilweise menschlich getroffene Entscheidungen liegen kann – die Führungsebene, die letztendlich über die Einführung oder weitere Nutzung von ADS entscheidet, die für die Einbindung der ADS-Systeme in die Prozesse und Abläufe der Organisation Zuständigen, die Entwickler:innen, die Tester:innen und Qualitätsprüfer:innen der Software und der zugrundeliegenden Datenbasis, unabhängige Gutachter:innen und Auditor:innen, die für die Systemsicherheit Verantwortlichen und nicht zuletzt die eigentlichen Anwender:innen. Insbesondere möchten wir darauf hinweisen, dass die Verantwortung somit nicht ausschließlich innerhalb der Organisation geteilt sein muss, sondern sogar teilweise außerhalb der Organisation verortet werden kann. Insbesondere für von den Entscheidungen betroffene Menschen kann die Unmöglichkeit, die für die Entscheidung in ihrem Fall verantwortliche Person oder Personen zu identifizieren, zumindest frustrierend, fallweise aber sogar gefährlich sein, und kann jedenfalls komplexe und schwierige Situationen erzeugen.

Um diese Art von Problemen zu vermeiden, müssen Verantwortungsfragen bereits vor der Einführung eines ADS-Systems behandelt werden. Wir haben hierfür ein Tool entwickelt (siehe unten), welches bei der Zuordnung von Verantwortung und Identifikation möglicher Probleme helfen kann. Jedenfalls muss bei dieser Frage unbedingt ein strukturierter Zugang verfolgt werden, in dem alle Akteur:innen, Aufgaben, Verantwortungsbereiche und Kommunikationswege berücksichtigt und behandelt werden, um nichts zu übersehen und keine Verantwortungslücken entstehen zu lassen.

Problem 3:
Verantwortungslücken

5 Leitfaden und Verantwortungszuordnung mittels VERA

Die zuvor beschriebenen Verschiebungen und Veränderungen der Verantwortung finden oft ohne die notwendigen Begleitmaßnahmen oder vorhergehende Planung statt. Nach unserer Analyse ist der erste Schritt zur Behandlung der möglichen negativen Effekte die Bewusstseins-schaffung und Information über die grundlegende Problematik. Es gibt bereits einige Leitfäden, Berichte und Handbücher (wie etwa Engelmann und Puntschuh 2020; O'Neil und Gunn 2020; Puntschuh und Fetic 2020a; Puntschuh und Fetic 2020b; Reisman et al. 2018), die einen ausgezeichneten Überblick über die vielgestaltigen möglichen Probleme bei der Nutzung von ADS-Systemen bieten, von Kosten und Zielkonflikten über Transparenz, Sicherheit, Datenschutz, Diskriminierung und Dokumentation bis hin zu Evaluierung und Ergebnisanalyse. Das Verantwortungsthema wird allerdings in all diesen Texten nur am Rande behandelt; um diese Lücke zu füllen, haben wir in Zusammenarbeit mit der Arbeiterkammer Wien und in enger Abstimmung mit Expert:innen aus Wissenschaft und Industrie einen Leitfaden speziell zur Frage der organisationalen Verantwortung erstellt. Wir haben die verschiedenen Herausforderungen zum Problemkreis der organisationalen Verantwortung aus der existierenden Forschungsliteratur sowie aus Anwendungsfällen gesammelt und gruppiert; parallel dazu haben wir sowohl Aufgabenbereiche und Funktionen als auch Verantwortungsbereiche identifiziert, gegliedert und differenziert. In der Folge haben wir die verschiedenen Problemstellungen in Form der Beziehungen und Verhältnisse zwischen Aufgaben/Funktionen und Verantwortungsbereichen formuliert; dieser Zugang erlaubte uns eine strukturierte Analyse der verschiedenen Aufgaben/Funktionen und Verantwortungsbereiche, was sie jeweils beinhalten sollten, sowie wie die wechselseitigen Verknüpfungen aussehen. Pläne zum Praxistest unseres Leitfadens und des ergänzenden digitalen Tools in Form von Fallstudien mussten leider wegen der COVID-19-Pandemie und ihrer Folgen vertagt bzw. aufgeschoben werden.

In unserem Leitfaden erläutern wir typische Probleme, die durch Verantwortungsverschiebungen entstehen können (siehe Abschnitt 4), und legen dar, wie man die verschiedenen Arten von Verantwortung sowie ihre Verteilung vor und nach der Einführung eines ADS-Systems erfassen und organisieren kann. Um sicherzustellen, dass alle Aufgaben und Zuständigkeiten zugewiesen und er-

Leitfaden Algorithmen in der Entscheidungsfindung: kurzer Überblick über das Thema, strukturierte Aufstellung von Aufgaben und Zuständigkeiten, Leitfragen zur Reflexion

füllbar sind, ist es entscheidend, sich zuerst einmal diesen grundlegenden und umfassenden Überblick zu verschaffen. Weiters machen wir eine strukturierte Aufstellung von Aufgaben und Zuständigkeiten zur Unterstützung einer Verantwortungszuordnung, namentlich die grundlegende Entscheidung für oder gegen die Einführung eines ADS-Systems; die Integration des Systems in die existierenden Prozesse, Strukturen und Abläufe der Organisation; die Entwicklung des Systems; seine praktische Anwendung; die Systemsicherheit; die Datenverwaltung; sowie zuletzt die (unabhängige) Evaluierung des Systems. Unser Leitfaden befürwortet weiters proaktives Fehlermanagement und die Formalisierung von Beschwerde- und Feedbackwegen (intern wie extern), um die schnellstmögliche Fehlererkennung und -behebung zu ermöglichen. Des Weiteren enthält der Leitfaden einen kurzen Abriss der relevanten Aspekte aus dem Datenschutzrecht (insbesondere zu Haftungsfragen, Betroffenenrechten und spezifischen Rechtsmaterien zu automatisierten Entscheidungen über Menschen) sowie gewisse technische Anforderungen an das System (wie Verständlichkeit, Transparenz, Dokumentation und Anpassbarkeit). Abgeschlossen wird der Leitfaden von einer Reihe an thematisch gruppierten Leitfragen, die bei der praktischen Umsetzung unserer Empfehlungen helfen sollen.

Ergänzend zum Leitfaden haben wir auch ein digitales Tool namens „VERA“ entwickelt, welches bei der Zuordnung von Aufgaben- und Verantwortungsbereichen assistiert. Der automatisierte Output eines solchen Tools kann natürlich nur einen Teil der möglichen Probleme aufzeigen und die Notwendigkeit einer genaueren Prüfung dieser und anderer Aspekte verdeutlichen; es kann und will dezidiert keine umfassende und vollständige Problemanalyse liefern, sondern soll als erster Schritt für tiefergehende Begutachtung und Prüfung dienen.

VERA besteht aus fünf Eingabeblocken, in denen die Benutzer:innen Fragen zum ADS-System beantworten, gefolgt von sechs Abschnitten mit automatisiertem Output zu möglichen Problemen hinsichtlich der Verantwortungsbereiche (siehe Abbildungen 1 und 2). Die Inputblöcke sind wie folgt strukturiert:

Struktur von VERA:
fünf Eingabeblocke,
sechs Ausgabe-
abschnitte

- In Block 1 werden die Namen aller Personen und/oder Personengruppen abgefragt, die in irgendeiner Form in das ADS-System und seine Anwendung involviert sind; gemeinsam mit der (Standard-)Option „niemand“ sind diese dann die möglichen Antworten für die folgenden Blöcke.
- In Block 2 sollen die Benutzer:innen folgende Aufgaben und Funktionen an Personen(-gruppen) zuweisen: die grundlegende Entscheidung für oder gegen die Einführung des ADS-Systems, die organisationale Implementierung, die Entwicklung, die praktische Anwendung, die Systemsicherheit, die Datenverwaltung sowie die Evaluierung.
- Block 3 enthält Fragen nach der Verantwortung für einige Bereiche, d. h. wer konkret gewisse Arten von Problemen lösen muss bzw. die Folgen dafür tragen muss, wenn (1) das System seine Zielvorgaben nicht erfüllt, (2) das System nicht ordentlich in die organisatorischen Prozesse und Strukturen integriert wurde, (3) es zu Datenschutzbeschwerden kommt, (4) die Systemsicherheit verletzt wird oder (5) das ADS-System falsch angewendet wird.
- Block 4 befasst sich Fragen über Befugnisse und Handlungsfähigkeit, nämlich: (1) Wer kann die Nutzung des ADS-Systems stoppen? (2) Wer kann die Prozessintegration und praktische Anwendung des Systems verändern? (3) Wer kann Daten in der zugrundeliegenden Datenbasis abändern und korrigieren? (4) Wer kann Sicherheitsmaßnahmen einrichten und vorschreiben?
- Der letzte Block 5 fragt das Vorhandensein von Kommunikationswegen und internen Beschwerdemechanismen zwischen den verschiedenen Personen(-gruppen) ab.

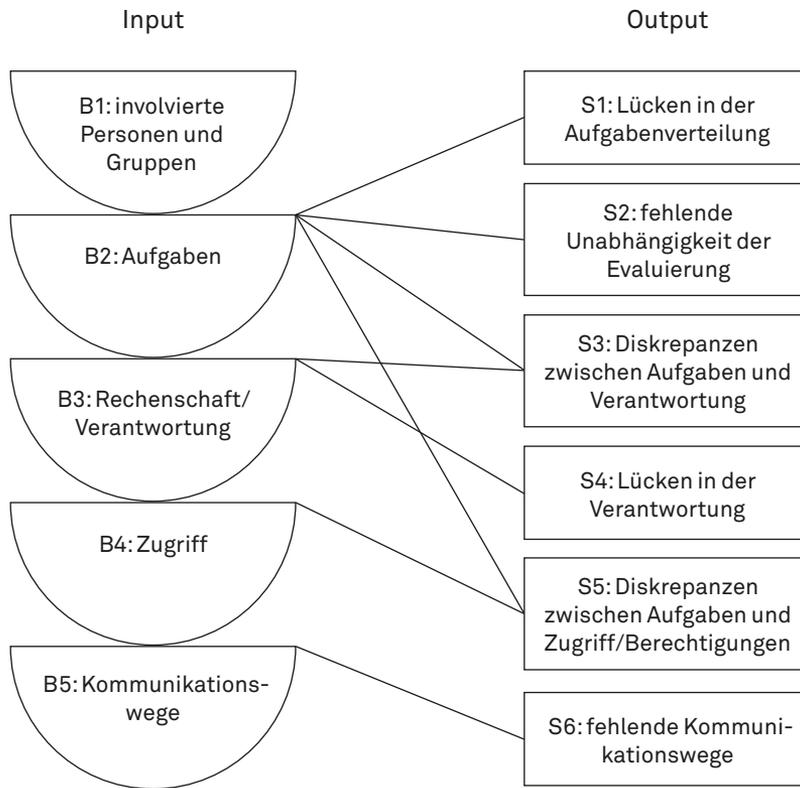


Abbildung 1:
Flowchart-Diagramm
zu den Verknüpfungen
zwischen Input
und Output von VERA

Der Output ist in sechs Abschnitte gegliedert, wovon sich jeder mit einem anderen Problemkreis befasst.

- Abschnitt 1 zeigt mögliche Lücken in der Zuweisung von Aufgaben auf (wenn in Block 2 die Antwort „niemand“ gegeben wurde).
- Abschnitt 2 befasst sich mit mangelnder Unabhängigkeit der Evaluierung, falls bei den für die Evaluierung Zuständigen durch andere, ebenfalls ihnen zugewiesene Aufgaben mögliche Interessenskonflikte bestehen könnten.
- Abschnitt 3 weist darauf hin, wenn jemand für einen Bereich verantwortlich ist (Block 3) ohne es als Aufgabenbereich zugeordnet zu haben (Block 2).
- Ähnlich wie Abschnitt 1 zeigt Abschnitt 4 Lücken in der Zuordnung von Verantwortung auf (wenn in Block 3 „niemand“ als Antwort aufscheint).

- In Abschnitt 5 hebt VERA mögliche Probleme durch mangelnde Übereinstimmung zwischen Aufgaben und Befugnissen hervor, wenn eine Person(-engruppe) für etwas verantwortlich ist (Block 3), ohne in diesem Bereich Änderungen veranlassen zu können (Block 4).
- Abschnitt 6 schließlich beinhaltet Probleme durch fehlende (aber notwendige) Kommunikations- und Beschwerdewege zwischen verschiedenen Gruppen.

Zur Illustration möchten wir an dieser Stelle kurz schlaglichtartig illustrieren, wie VERA in einem fiktiven Fallbeispiel funktionieren könnte. Die Gutes Beispiel KG, ein mittelständischer Personalvermittler, plant die Einführung von Algorithmenunterstützung bei der Verknüpfung von Stellenangeboten und Personen in ihrer Datenbank. Intern für dieses Projekt verantwortlich zeichnen zwei der vier Eigentümer:innen, Azra Jašarević und Deniz Nacar; die technische Entwicklung wird aus Ressourcengründen (Gutes Beispiel hat nur eine kleine IT-Abteilung, die nur für Administration und Wartung im kleineren Rahmen vorgesehen ist) extern an die TechSolve GmbH vergeben. Der Algorithmus soll zunächst in einem Pilotbetrieb nur von zwei Angestellten (Patrick Felderer und Eunice Oumarou) eingesetzt werden; nach einem halben Jahr sollen die Ergebnisse evaluiert werden.

Fiktives Fallbeispiel:
Gutes Beispiel KG

Zur internen Projektabwicklung wird eine Arbeitsgruppe („AG Algorithmen“) eingerichtet, die sich entscheidet, zum Verantwortungstracking das digitale Tool VERA einzusetzen. Die AG sammelt demnach alle relevanten Daten (sofern verfügbar/bekannt) und gibt diese in die Eingabemaske ein (vgl. Abbildung 2); VERA zeigt auf Basis dieses Inputs dann einige potenzielle Probleme auf, zum Beispiel:

- Zum Zeitpunkt der Eingabe in VERA ist noch unklar, wer für die Implementierung des Algorithmus in die existierenden Prozesse und Abläufe bei Gutes Beispiel sorgen wird.
- Niemand ist explizit für datenschutzrechtliche Beschwerden zuständig.
- Jašarević ist als studierte Informatikerin für Datenverwaltung und Systemsicherheit verantwortlich (Zweiteres zusammen mit der IT-Abteilung), soll aber gemeinsam mit Nacar auch

Wer ist schuld, wenn Algorithmen irren?

Schritt 2: Zuständigkeiten

A) Aufgaben

Geben Sie hier bitte jeweils die Personen oder Gruppen an, die für den jeweiligen Bereich die **Entscheidungsbefugnis** haben.

Wer entscheidet darüber, ob ein Algorithmensystem eingeführt oder (weiter-)betrieben wird?

Zuständige Person 1: **Azra Jašarević** | Zuständige Person 2: **Deniz Nacar** | WEITERE PERSON

Wer entwickelt das System?

Zuständige Person 1: **TechSolve GmbH** | WEITERE PERSON

Wer implementiert das System?

Zuständige Person 1: **AG Algorithmen** | WEITERE PERSON

Wer wendet das System an?

Zuständige Person 1: **Azra Jašarević** | WEITERE PERSON

Zuständige Person 2: **Deniz Nacar** | Sicherheit des Systems zuständig?

Zuständige Person 3: **Patrick Felderer** | WEITERE PERSON

Zuständige Person 4: **Eunice Oumarou** | WEITERE PERSON

Zuständige Person 5: **AG Algorithmen** | WEITERE PERSON

Zuständige Person 6: **TechSolve GmbH** | WEITERE PERSON

Zuständige Person 7: **niemand** | WEITERE PERSON

Zuständige Person 8: **niemand** | WEITERE PERSON

Zuständige Person 9: **niemand** | WEITERE PERSON

Zuständige Person 10: **niemand** | WEITERE PERSON

Zuständige Person 11: **niemand** | WEITERE PERSON

Zuständige Person 12: **niemand** | WEITERE PERSON

Abbildung 2: Screenshots von Teilen des Inputs und Outputs von VERA

Schritt 3: Auflösung – identifizierte Probleme

Auf Basis Ihrer Antworten wurden folgende **potenzielle Probleme** identifiziert. Diese Problemliste ist als **Anregung** zu verstehen, genauer darüber nachzudenken, ob diese Probleme im **konkreten Fall** wirklich bestehen und welche **Auswirkungen** sie haben können, und erhebt insbesondere **keinen Anspruch auf Vollständigkeit**.

Problemkreis 1

Lücken in der Aufgabenverteilung

- Es ist möglichst bald zu klären, wer das **Algorithmensystem implementiert**.

Problemkreis 2

Fehlende Unabhängigkeit der Evaluierung

- Da die gleiche Person oder Gruppe für die Evaluierung wie für die Sicherheit des Algorithmensystems zuständig ist, kann eine **unabhängige Evaluierung nicht gewährleistet** werden.
- Da die gleiche Person oder Gruppe für die Evaluierung wie für die Datenverwaltung zuständig ist, kann eine **unabhängige Evaluierung nicht gewährleistet** werden.

Problemkreis 3

Jemand ist für Aufgaben verantwortlich, die nicht ihr/ihm zugeteilt sind

Wir konnten keine offensichtlichen Probleme dieser Art identifizieren.

Problemkreis 4

Lücken in der Verantwortung

- Da für Probleme mit **datenschutzrechtlichen Beschwerden** niemand zuständig ist, besteht hier potenziell eine Verantwortungslücke.
- Potenziell kommt es zu einer Überschneidung von Verantwortung, weil für die **Erfüllung der Zielvorgaben** mit Azra Jašarević und Deniz Nacar mehrere verantwortlich sind.

die Evaluierung verantworten – ein möglicher Interessenskonflikt, da die Evaluierung möglichst unabhängig von anderen Funktionen im Prozess erfolgen sollte.

- Mit Jašarević und Nacar sind zwei Personen für die Erfüllung der Zielvorgaben verantwortlich – derart geteilte Verantwortung kann potenziell problematisch sein (muss es aber nicht).

An diesem Beispiel soll weiters deutlich werden, dass ein solches einfaches Werkzeug natürlich nicht alle potenziellen Probleme identifizieren kann. So ist etwa fraglich, wie die ad hoc eingerichtete Arbeitsgruppe in die internen Verantwortungsstrukturen der Gutes Beispiel eingebunden ist.

6 Grenzen und Ausblick

Dieser Artikel bietet einen Überblick über die Problemfelder, die mit dem Einsatz von ADS in Form von Biases und fehlerhaften Entscheidungen einhergehen. Einige Aspekte sprengen allerdings den Rahmen dieses Beitrags, etwa die Frage des Vertrauens in Organisationen und Algorithmensysteme und der Einfluss der Expertise und des Vorwissens der Entscheider:innen. Die Resultate einer Studie des Pew Research Centers von 2018 (Smith 2018) weisen auf eine grundlegende Skepsis gegenüber der Nutzung persönlicher Daten in ADS-Systemen hin; generell hatten die Befragten wenig Vertrauen in die richtige Anwendung solcher Systeme durch die Entwicklerfirmen. Besonders wichtig sind auch die Einstellungen der Entscheider:innen und Anwender:innen sowie das Vorwissen über ADS-Systeme, welche die Meinung bzgl. der Nutzung von ADS-Systemen beeinflussen kann (siehe z. B. Alexander, Blinder und Zak 2019; Burton, Stein und Jensen 2020; Lee und Baykal 2017). Das Vertrauen der Angestellten einer Organisation, die ADS-Systeme einsetzt, ist von zentraler Bedeutung für den verantwortungsvollen Umgang damit.

Manche Problemkreise und Herausforderungen von ADS-Systemen konnten hier nicht im Detail beleuchtet werden, sollten aber in Zukunft weiter untersucht werden. So ist bereits mehrfach belegt, dass die den ADS-Systemen zugrundeliegenden Daten nicht neutral oder objektiv sind und oft den Status quo bestärken und reproduzieren (Barocas und Selbst 2016; Christin 2017; Eubanks 2018; O’Neil 2016). In vielen Fällen werden gewisse Datensätze pri-

Grundlegende Fragen der Datenqualität, Repräsentativität, Modellierung und Vorhersagbarkeit

mär deshalb verwendet, weil sie einfach verfügbar und zugänglich sind, obwohl bekannt ist, dass sie verzerrt oder nicht repräsentativ sind. Diese Probleme können schwerwiegende Folgen für die Entscheidungsvorschläge von ADS-Systemen haben. In einer groß angelegten Vergleichsstudie (Salganik et al. 2020) hat sich weiters herausgestellt, dass Langzeitvorhersagen (etwa über Lebensverläufe) unabhängig von der verwendeten Methode sehr ungenau sind, was grundlegende Fragen über die Anwendbarkeit von Vorhersagetechnologien außerhalb eines enorm eingeschränkten Nutzungsbereichs aufwirft. Zuletzt gibt es auch Forschung über Algorithmen aus dem Blickwinkel der Performativität, d. h. unter Berücksichtigung der wechselseitigen Beeinflussung zwischen den Algorithmen und der sozialen Realität statt unter der Annahme, dass Algorithmen (größtenteils) statische Gebilde sind (Glaser, Pollock und D'Adderio 2021). Diese (wichtige) Perspektive hätte allerdings den Rahmen unseres Artikels gesprengt.

Ein weiterer wichtiger Aspekt außerhalb des Rahmens dieses Artikels sind Datenschutzfragen sowie generell der rechtliche Rahmen, insbesondere mit Hinblick auf die Datenschutz-Grundverordnung (DSGVO), sowie die Frage, wie Organisationen bei Nutzung von ADS-Systemen den Datenschutz gewährleisten können. Die verwendeten Daten und die Weiterverarbeitung der von Benutzer:innen eingegebenen Informationen kann bzgl. Privatsphäre und Einwilligung problematisch sein. Einerseits kann auch die Nutzung von Daten, deren Verarbeitung die Betroffenen zugestimmt haben, fragwürdig sein, wenn die Einwilligung unter Verwendung sogenannter Dark Patterns eingeholt wurde. Andererseits nutzen viele ADS-Systeme ergänzend Datensätze aus externen Quellen (wie Daten aus den sozialen Medien oder aus der Nutzung eines Smartphones), um die Entscheidungsvorschläge ihrer Systeme zu verbessern (Castellucia und Métayer 2019; Lohokare, Dani und Sontakke 2017; Wei et al. 2016). Hier bietet die DSGVO einen rechtlichen Rahmen für die Regulierung von ADS-Systemen, z. B. durch das Erfordernis einer informierten, aktiven Einwilligung (Art. 6–7), das Recht auf Vergessenwerden und Datenlöschung (Art. 17) und das Recht auf Korrektur falscher Daten (Art. 16). Weiters reguliert die DSGVO die automatisierte Verarbeitung von Nutzer:innendaten. Insbesondere räumt sie das Recht ein, nicht rein automatisierten Entscheidungen ohne Beteiligung menschlicher Entscheider:innen unterworfen zu werden (Art. 22). Die Verordnung schreibt weiters auch die Option auf menschliche Überprüfung der Entscheidungen eines ADS-Systems vor (Dreyer und Schulz 2019).

Datenschutzfragen
(DSGVO und darüber
hinaus)

Allerdings lässt die Rechtslage einige Fragen offen, so etwa den Schutz Betroffener vor weitergehender Nutzung ihrer Daten durch Datenverantwortliche (wie z. B. durch Verkauf der Daten an Dritte). Auch die Frage der Transparenz komplexer Systeme, sowohl gegenüber betroffenen Einzelpersonen als auch gegenüber der Gesamtöffentlichkeit, ist ungelöst (Castets-Renard 2019). Krafft, Zweig und König (i. E.) plädieren dafür, dass spezifische Rechtsmaterie für den Anwendungsbereich algorithmischer und algorithmengestützter Entscheidungsfindung verabschiedet werden sollte (in Ergänzung zum allgemeinen Datenschutzrecht), nicht zuletzt um die große Bandbreite an verschiedenen bereits existierenden und noch möglichen Anwendungsbereichen von Automatisierung in diesem Kontext adäquat zu berücksichtigen. Eine Weiterentwicklung des Rechtsbestands erscheint nicht zuletzt vor dem Hintergrund der COVID-19-Pandemie notwendig, im Zuge derer die Anwendung und Verbreitung von ADS in Europa massiv zugenommen hat (Chiusi 2020). Der „European Approach to Artificial Intelligence“ der Europäischen Kommission (4) sowie der jüngst vorgestellte Entwurf zur Regulierung von KI (5) könnten potenziell Schritte in diese Richtung darstellen.

7 Fazit

Die Einführung von ADS-Systemen kann erhebliche Auswirkungen auf die effektive Verteilung und Zuweisung von Verantwortung innerhalb einer Organisation haben. Generell bedeutet der Einsatz von ADS-Systemen einige Herausforderungen für die Entscheidungsprozesse der menschlichen Entscheider:innen.

Wir haben drei grundlegende Typen von Problemen identifiziert, die mit Änderungen durch die Einführung von ADS-Systemen einhergehen können. Erstens könnte Angestellten nach Veränderungen ihrer Aufgaben die notwendige Information oder Schulung fehlen, um diese verantwortungsvoll auszuführen. Zweitens könnten Entscheider:innen nicht die notwendigen Befugnisse oder Zugriffsrechte haben, um bei auftretenden Problemen in ihrem Wirkungsbereich Änderungen am ADS-System zu veranlassen oder systemische Fehler zu korrigieren. Drittens kann es durch die Umschichtung bestehender und die Einführung neuer Aufgabenbereiche (Systemsicherheit, Datenschutz etc.) dazu kommen, dass gewisse Verantwortlichkeiten nicht ausreichend klar oder gar nicht zugewiesen werden, sodass Verantwortungslücken entstehen. Um

solche Probleme zu vermeiden und schon vorab auf diese Veränderungen vorbereitet zu sein, schlagen wir eine umfassende Zuordnung aller Aufgaben und Verantwortlichkeiten innerhalb des relevanten Organisationsbereichs vor. Wir haben dazu einen Leitfaden für Organisationen erstellt, der bei solchen Vorhaben helfen soll; weiters haben wir ein interaktives digitales Tool entwickelt, welches die Beziehungen zwischen (1) der Aufgabenverteilung, (2) der Verantwortung und (3) den Berechtigungen und Zugriffsrechten darlegt. Dieses Tool soll dabei helfen, potenzielle Probleme in einem spezifischen Organisationskontext zu erkennen und mögliche Problemfelder zur genaueren Prüfung aufzuzeigen und dadurch den verantwortungsbewussten und transparenten Einsatz von ADS-Systemen in Organisationen unterstützen.

Danksagung

Die erste Autorin und der dritte Autor wurden vom Digitalisierungsfonds Arbeit 4.0 der Arbeiterkammer Wien durch das Projekt B-01 „Algorithmen, Recht und Gesellschaft: Entscheider*innen zwischen algorithmischer Steuerung und persönlicher Verantwortung“ unterstützt.

(4) <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> (letzter Zugriff: 31.07.2021)

(5) <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence> (letzter Zugriff: 31.07.2021)

Lukas Daniel Klausner

(l17r.eu) ist Mathematiker und Informatiker und forscht als Researcher an der FH St. Pölten in den Bereichen Security, Privacy, Data Science und Science and Technology Studies.

Gemeinsam mit Paola Lopez hat er kürzlich den Arbeitskreis Mathematik trans- und interdisziplinär (AK MatriX) gegründet. Seine aktuellen Forschungsschwerpunkte umfassen Critical Algorithm and Data Studies, Ethik und Bias von Algorithmen und alle Fragestellungen, bei denen Technik und Gesellschaft aufeinandertreffen.

Quellenverzeichnis

- Adensamer, Angelika/Gsenger, Rita/
Klausner, Lukas Daniel (im Erscheinen).
“Computer Says No”: Algorithmic Decision
Support and Organisational Responsibility.
- Adensamer, Angelika/Klausner, Lukas
Daniel (2021a). “Part Man, Part Machine,
All Cop”: Automation in Policing, in
Frontiers in Artificial Intelligence 4.
DOI: 10.3389/FRAI.2021.655486.
- Adensamer, Angelika/Klausner, Lukas
Daniel (2021b). Algorithmen in der
Entscheidungsfindung – Leitfaden zu
Verantwortlichkeit und Rechenschaft.
Wien: Arbeiterkammer Wien.
- Alexander, Veronika/Blinder, Collin/Zak,
Paul J. (2019). Why Trust an Algorithm?
Performance, Cognition, and Neurophysio-
logy, in Computers in Human Behavior 89,
S. 279–288. DOI: 10.1016/J.CHB.2018.07.026.
- Barocas, Solon/Selbst, Andrew D. (2016).
Big Data’s Disparate Impact, in California
Law Review 104(3), S. 671–732.
DOI: 10.15779/Z38BG31.
- Bennett Moses, Lyria/Chan, Janet (2018).
Algorithmic Prediction in Policing:
Assumptions, Evaluation, and Accountabili-
ty, in Policing and Society 28(7), S. 806–822.
DOI: 10.1080/10439463.2016.1253695.
- Bovens, Mark (2007). Analysing and
Assessing Accountability: A Conceptual
Framework, in European Law Journal 13(4),
S. 447–468. DOI: 10.1111/J.1468-0386.
2007.00378.X.
- Burton, Jason W./Stein, Mari-Klara/Jensen,
Tina Blegind (2020). A Systematic Review
of Algorithm Aversion in Augmented
Decision Making, in Journal of Behavioral
Decision Making 33(2), S. 220–239.
DOI: 10.1002/BDM.2155.
- Castellucia, Claude/Le Métayer, Daniel
(2019). Understanding Algorithmic
Decision-Making: Opportunities and
Challenges. Luxembourg: Publications
Office of the European Union. ISBN: 978-
92-846-3506-1. DOI: 10.2861/536131.
- Castets-Renard, Céline (2019).
Accountability of Algorithms in the GDPR
and Beyond: A European Legal Framework
on Automated Decision-Making, in
Fordham Intellectual Property, Media &
Entertainment Law Journal 30(1), S. 91–137.
URL: [https://ir.lawnet.fordham.edu/iplj/
vol30/iss1/3](https://ir.lawnet.fordham.edu/iplj/vol30/iss1/3).
- Chiusi, Fabio (2020). Automated Decision-
Making Systems in the COVID-19 Pandemic:
A European Perspective (Special Issue of
the Automating Society Report 2020).
Berlin: Algorithm Watch. URL: [https://
algorithmwatch.org/en/wp-content/
uploads/2020/08/ADM-systems-in-the-
Covid-19-pandemic-Report-by-AW-BSt-
Sept-2020.pdf](https://algorithmwatch.org/en/wp-content/uploads/2020/08/ADM-systems-in-the-Covid-19-pandemic-Report-by-AW-BSt-Sept-2020.pdf).
- Christen, Markus/Mader, Clemens/Čas,
Johann/Abou-Chadi, Tarik/Bernstein,
Abraham/Braun Binder, Nadja/Dell’Aglia,
Daniele/Fábián, Luca/George, Damian/
Gohdes, Anita/Hilty, Lorenz/Kneer, Markus/
Krieger-Lamina, Jaro/Licht, Hauke/Scherer,
Anne/Som, Claudia/Sutter, Pascal/

Thouvenin, Laurent (2020). Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstliche Intelligenz (Hrsg.: TA-SWISS). TA Publikationsreihe 72. Zürich: vdf Hochschulverlag. ISBN: 978-3-7281-4002-9. DOI: 10.3218/4002-9.

Christin, Angèle (2017). Algorithms in Practice: Comparing Web Journalism and Criminal Justice, in *Big Data & Society* 4(2), S. 1–14. DOI: 10.1177/2053951717718855.

Clarke, Jennifer A. (2007). Representativeness Heuristic, in Baumeister, Roy F. / Vohs, Kathleen D. (Hrsg.): *Encyclopedia of Social Psychology*. Thousand Oaks, CA: SAGE Publishing, S. 747. DOI: 10.4135/9781412956253.N448.

De-Arteaga, Maria/Fogliato, Riccardo/Chouldechova, Alexandra (2020). A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores, in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI: ACM, S. 1–12. DOI: 10.1145/3313831.3376638.

De Neys, Wim/Rossi, Sandrine/Houdé, Olivier (2013). Bats, Balls, and Substitution Sensitivity: Cognitive Misers Are No Happy Fools, in *Psychonomic Bulletin & Review* 20(2), S. 269–273. DOI: 10.3758/S13423-013-0384-5.

Dietvorst, Berkeley J./Simmons, Joseph P./Massey, Cade (2015). Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err, in *Journal of Experimental Psychology: General* 144(1), S. 114–126. DOI: 10.1037/XGE0000033.

Dietvorst, Berkeley J./Simmons, Joseph P./Massey, Cade (2018). Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them, in *Management Science* 64(3), S. 1155–1170. DOI: 10.1287/MNSC.2016.2643.

Dreyer, Stephan/Schulz, Wolfgang (2019). The General Data Protection Regulation and Automated Decision-Making: Will It Deliver? Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals, Groups and Society as a Whole. Volume #5. Discussion Paper Ethics of Algorithms. Gütersloh: Bertelsmann Stiftung. DOI: 10.11586/2018018.

Dunn, Timothy L./Risko, Evan F. (2019). Understanding the Cognitive Miser: Cue-Utilization in Effort-Based Decision Making, in *Acta Psychologica* 198, S. 1–10. DOI: 10.1016/J.ACTPSY.2019.102863.

Engelmann, Jan/Puntschuh, Michael (2020). KI im Behördeneinsatz: Erfahrungen und Empfehlungen. Berlin: Kompetenzzentrum Öffentliche IT. ISBN: 978-3-948582-04-3. URL: <https://www.oeffentliche-it.de/publikationen?doc=177719>.

Eubanks, Virginia (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York, NY: St. Martin's Press. ISBN: 978-1-250-07431-7.

Faraj, Samer/Pachidi, Stella/Sayegh, Karla (2018). Working and Organizing in the Age of the Learning Algorithm, in *Information and Organization* 28(1), S. 62–70. DOI: 10.1016/J.INFOANDORG.2018.02.005.

- Ferguson, Andrew Guthrie (2017). Policing Predictive Policing, in *Washington University Journal of Law & Policy* 94(5), S. 1109–1189. URL: https://openscholarship.wustl.edu/law_lawreview/vol94/iss5/5/.
- Fjeld, Jessica/Achten, Nele/Hilligoss, Hannah/Nagy, Adam/Srikumar, Madhulika (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*. Volume 2020-1. Berkman Klein Center Research Publication. Cambridge, MA: Berkman Klein Center for Internet & Society. URL: <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420>.
- Friedman, Batya/Nissenbaum, Helen (1996). Bias in Computer Systems, in *ACM Transactions on Information Systems* 14(3), S. 330–347. DOI: 10.1145/230538.230561.
- Gillespie, Tarleton (2016). #trendingis-trending: when algorithms become culture, in Seyfert, Robert/Roberge, Jonathan (Hrsg.): *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*. Abingdon-on-Thames: Routledge, S. 52–75. ISBN: 978-1-315-65869-8. DOI: 10.4324/9781315658698.
- Glaser, Vern L./Pollock, Neil/D'Adderio, Luciana (2021). *The Biography of an Algorithm: Performing Algorithmic Technologies in Organizations*, in *Organization Theory* 2(2). DOI: 10.1177/26317877211004609.
- Goddard, Kate/Roudsari, Abdul/Wyatt, Jeremy C. (2011). Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators, in *Journal of the American Medical Informatics Association* 19 (1), S. 121–127. DOI: 10.1136/AMIAJNL-2011-000089.
- Hao, Karen (2019). Police Across the US are Training Crime-Predicting AIs on Falsified Data, in *MIT Technology Review* (13.02.2019). URL: <https://www.technology-review.com/s/612957/predictive-policing-algorithms-ai-crime-dirty-data/>.
- Heimstädt, Maximilian/Dobusch, Leonhard (2020). Transparency and Accountability: Causal, Critical and Constructive Perspectives, in *Organization Theory* 1(4). DOI: 10.1177/2631787720964216.
- Kellogg, Katherine C./Valentine, Melissa A./Christin, Angèle (2020). Algorithms at Work: The New Contested Terrain of Control, in *Academy of Management Annals* 14(1), S. 366–410. DOI: 10.5465/ANNALS.2018.0174.
- Kolkman, Daan (im Erscheinen). *The (In)Credibility of Algorithmic Models to Non-Experts*, in *Information, Communication & Society*. DOI: 10.1080/1369118X.2020.1761860.
- Krafft, Tobias D./Zweig, Katharina Anna/König, Pascal David (im Erscheinen). *How to Regulate Algorithmic Decision-Making: A Framework of Regulatory Requirements for Different Applications*, in *Regulation & Governance*. DOI: 10.1111/REGO.12369.
- Lee, Min Kyung/Baykal, Su (2017). *Algorithmic Mediation in Group Decisions: Fairness Perceptions of Algorithmically Mediated vs. Discussion-Based Social Division*, in *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing*. CSCW '17. Portland, OR: ACM, S. 1035–1048. DOI: 10.1145/2998181.2998230.

Lohokare, Jay/Dani, Reshul/Sontakke, Sumedh (2017). Automated Data Collection for Credit Score Calculation Based on Financial Transactions and Social Media, in 2017 International Conference on Emerging Trends & Innovation in ICT. ICEI 2017. Pune: IEEE Computer Society, S. 134–138. DOI: 10.1109/ETIICT.2017.7977024.

Loi, Michele/Spielkamp, Matthias (2021). Towards Accountability in the Use of Artificial Intelligence for Public Administrations, in Proceedings of the Fourth AAAI/ACM Conference on AI, Ethics, and Society. AIES '21. Virtual Event: ACM, TBD. DOI: 10.1145/3461702.3462631.

Lyell, David/Coiera, Enrico (2017). Automation Bias and Verification Complexity: A Systematic Review, in Journal of the American Medical Informatics Association 24(2), S. 423–431. DOI: 10.1093/JAMIA/OCW105.

McGrath, Stephen Keith/Whitty, Stephen Jonathan (2018). Accountability and Responsibility Defined, in International Journal of Managing Projects in Business 11(3), S. 687–707. DOI: 10.1108/IJMPB-06-2017-0058.

Moradi, Pegah/Levy, Karen (2020). The Future of Work in the Age of AI: Displacement or Risk-Shifting?, in Dubber, Markus Dirk/Pasquale, Frank/Das, Sunit (Hrsg.): The Oxford Handbook of Ethics of AI. Oxford: Oxford University Press, S. 271–288. ISBN: 978-0-1900-6739-7. DOI: 10.1093/OxfordHB/9780190067397.013.17. URL: <http://www.karen-levy.net/wp-content/uploads/2020/07/oxfordhb-9780190067397-e-17.pdf>.

O'Neil, Cathy (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Publishing Group. ISBN: 978-0-553-41881-1.

O'Neil, Cathy/Gunn, Hanna (2020). Near-Term Artificial Intelligence and the Ethical Matrix, in Liao, S. Matthew (Hrsg.): Ethics of Artificial Intelligence Oxford: Oxford University Press, S. 237–270. ISBN: 978-0-190-90503-3. DOI: 10.1093/oso/9780190905033.003.0009.

Parasuraman, Raja/Manzey, Dietrich H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration, in Human Factors 52(3), S. 381–410. DOI: 10.1177/0018720810376055.

Poel, Ibo van de/Royackers, Lambèr/Zwart, Sjoerd Douwe (2015). Moral Responsibility and the Problem of Many Hands. Abingdon-on-Thames: Routledge. ISBN: 978-1-1383-4671-0.

Puntschuh, Michael/Fetic, Lajla (2020a). Handreichung für die digitale Verwaltung: Algorithmische Assistenzsysteme gemeinwohlorientiert gestalten. Gütersloh: Bertelsmann Stiftung. DOI: 10.11586/2020060.

Puntschuh, Michael/Fetic, Lajla (2020b). Praxisleitfaden zu den Algo.Rules: Orientierungshilfen für Entwickler:innen und ihre Führungskräfte. Gütersloh: Bertelsmann Stiftung. DOI: 10.11586/2020029.

- Reisman, Dillon/Schultz, Jason/Crawford, Kate/Whittaker, Meredith (2018). *Algorithm Impact Assessment: A Practical Framework for Public Agency Accountability*. New York, NY: AI Now Institute. URL: <https://ainowinstitute.org/aiareport2018.pdf>.
- Salganik, Matthew J. et al. (2020). Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration, in *Proceedings of the National Academy of Sciences of the United States of America* 117(15), S. 8398–8403. DOI: 10.1073/PNAS.1915006117.
- Schäufele, Fabia (2017). *Profiling zwischen sozialer Praxis und technischer Prägung: Ein Vergleich von Flughafensicherheit und Credit-Scoring*. Wiesbaden: Springer Fachmedien Wiesbaden. ISBN: 978-3-658-16968-8. DOI: 10.1007/978-3-658-16969-5.
- Skitka, Linda J./Mosier, Kathleen L. / Burdick, Mark D. (1999). Does Automation Bias Decision-Making?, in *International Journal of Human-Computer Studies* 51(5), S. 991–1006. DOI: 10.1006/IJHC.1999.0252.
- Smith, Aaron (2018). *Public Attitudes Toward Computer Algorithms*. Washington, DC. URL: <https://www.pewresearch.org/internet/2018/11/16/attitudes-toward-algorithmic-decision-making/>.
- Vieth, Kilian/Wagner, Ben (2017). *Teilhabe, ausgerechnet. Wie algorithmische Prozesse Teilhabechancen beeinflussen können. Volume #2. Impuls Algorithmenethik*. Gütersloh: Bertelsmann Stiftung. DOI: 10.11586/2017027.
- Wagner, Ben (2019). *Liabe, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems*, in *Policy & Internet* 11(1), S. 104–122. DOI: 10.1002/POI3.198.
- Wei, Yanhao/Yildirim, Pinar/Van den Bulte, Christophe/Dellarocas, Chrysanthos (2016). *Credit Scoring with Social Network Data*, in *Marketing Science* 35(2), S. 234–258. DOI: 10.1287/MKSC.2015.0949.
- Wieringa, Maranke (2020). *What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability*, in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. FAT* '20*. Barcelona: ACM, S. 1–18. DOI: 10.1145/3351095.3372833.
- Yeomans, Michael/Shah, Anuj/Mullainathan, Sendhil/Kleinberg, Jon (2019). *Making Sense of Recommendations*, in *Journal of Behavioral Decision Making* 32, S. 403–414. DOI: 10.1002/BDM.2118.
- Zerilli, John/Knott, Alistair/Maclaurin, James/Gavaghan, Colin (2019). *Algorithmic Decision-Making and the Control Problem*, in *Minds and Machines* 29(4), S. 555–578. DOI: 10.1007/S11023-019-09513-7.
- Zweig, Katharina Anna/Fischer, Sarah/Lischka, Konrad (2018). *Wo Maschinen irren können: Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung. Volume #4. Impuls Algorithmenethik*. Gütersloh: Bertelsmann Stiftung. DOI: 10.11586/2018006.



Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen

Einleitung	77
Die Bedeutung von AI und der Versuch einer Definition	78
Die richtigen Daten für den richtigen Zweck	82
Gefahr der fehlerhaften Verarbeitung von Daten (Bias)	82
Umgang mit Bias-Problemen	83
Vor- und Nachteile von AI-Systemen	83
AI-Auslegung der Plattform Industrie 4.0	84
Empfehlungen bei der Implementierung einer AI-Lösung	85
Vor der Einführungsphase	86
Während der Einführungsphase	91
Nach der Einführungsphase	94
Der Mensch als wichtigster Faktor	96
Positive Akzeptanzfaktoren	96
Negative Akzeptanzfaktoren	98
AI-Beispiele aus der Praxis	100
Effektives Zeit- und Selbstmanagement	100
Produktionsüberwachung	102
Qualitätsbestimmung Baumstamm	104
Danksagung	106
Über die Autorin	107

Einleitung

Künstliche Intelligenz (KI) oder Artificial Intelligence (AI) wird in den kommenden Jahren eine zunehmend weite Verbreitung erfahren und beschäftigt schon heute Unternehmen in strategischen und operativen Belangen. Neben technischen Herausforderungen und Aspekten bei der Einführung neuer Technologien gibt es eine Reihe organisatorischer Aspekte, die berücksichtigt werden sollten, damit eine erfolgreiche Einführung eines neuen Systems gewährleistet ist.

Dem folgenden Leitfaden liegt die These zu Grunde, dass die Berücksichtigung menschenzentrierter Faktoren AI-Systeme erfolgreicher macht, die Akzeptanz bei Nutzer:innen steigt und damit die Wirksamkeit dieser Systeme deutlich erhöht wird. Dies führt in weiterer Folge dazu, dass der Wirtschaftsstandort Österreich wettbewerbsfähiger wird.

Erfahrungen aus der Praxis, Erkenntnisse aus dem Projekt „AI for GOOD“ und eine Vielzahl von Projektaktivitäten sind in diesen Leitfaden eingearbeitet. Dieser richtet sich an Unternehmen und Mitarbeiter:innen gleichermaßen, die erstmalig mit dieser Thematik durch den innerbetrieblichen Einsatz an fortschrittlichen Technologien in Berührung kommen oder nach Möglichkeiten suchen, um weitere Potenziale auszuschöpfen.

Der vorliegende Leitfaden bezieht sich auf AI hinsichtlich innerbetrieblicher Anwendungen. Daneben gibt es auch weitere Arten: im persönlichen Umfeld, etwa auf Smartphones oder andere für ganz spezielle Bereiche wie automatisiertes Fahren.

Verein Industrie 4.0 – Die Plattform für intelligente Produktion

Der Verein „Industrie 4.0 Österreich“ wurde 2015 als Initiative des österreichischen Bundesministeriums für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie sowie von Arbeitgeber- und Arbeitnehmerverbänden gegründet. Diese erarbeiten gemeinsam mit Mitgliedern aus Wirtschaft, Wissenschaft und Interessenvertretungen in spezifischen Expert:innengruppen Strategien zur nachhaltigen und erfolgreichen Umsetzung der digitalen Transformation im Kontext von Industrie 4.0. Ziel ist es, die technologischen Entwicklungen und Innovationen durch Digitalisierung bestmöglich und sozialverträglich für Unternehmen, Beschäftigte und die Gesellschaft in Österreich zu nutzen und verantwortungsvoll umzusetzen. Der Verein Industrie 4.0 Österreich nimmt dabei eine wichtige Rolle in der nationalen und internationalen Koordinierung, Strategiefindung und Informationsbereitstellung ein.

Die Berücksichtigung menschenzentrierter Faktoren macht AI-Systeme erfolgreicher, steigert die Akzeptanz und damit auch die Wirksamkeit.

AI for GOOD - Menschenzentrierter Einsatz von AI-Lösungen

Ziel von „AI for Good“ ist, den Weg für einen menschenzentrierten Ansatz bei Artificial Intelligence (AI) -Anwendungen zu ebnet. Die Konsortialpartner Software Competence Center Hagenberg (SCCH) und Know-Center Graz begleiteten das Projekt wissenschaftlich. Durch eine Befragung von Betriebsrät:innen und Produktionsleiter:innen und eine eingehende Analyse erfolgreicher Use Cases, die von Axians ICT, TietoEVERY und dem Know-Center eingebracht wurden, wurden Erfolgsfaktoren für AI-Anwendungen abgeleitet. Im Rahmen mehrerer Workshops wurden konkrete Erfahrungen präsentiert, denn es ging darum, anhand einfacher Darstellung von Unternehmen zu Unternehmen zu erklären, was die Schwierigkeiten für Mitarbeiter:innen und Unternehmen gleichermaßen waren und welche Ratschläge für eine menschenzentrierte Implementierung gegeben werden können. Der Fokus des Projekts „AI for Good“, welches vom AK-Digifonds gefördert wird, liegt darin, Informationen und Impulse für die Menschenzentrierung von AI-Lösungen zu erarbeiten und zu kommunizieren.

Die Bedeutung von AI und der Versuch einer Definition

Eine zentrale Zielsetzung bzw. Empfehlung dieses Leitfadens ist den „AI“-Begriff zu entmystifizieren und Missverständnissen und Fehlannahmen vorzubeugen, indem unternehmensintern klar und transparent definiert und kommuniziert wird, welche Art von AI angestrebt wird.

Als AI werden sowohl äußerst komplexe Algorithmen und Systeme als auch recht einfache Verfahren bezeichnet. Innerhalb dieser Bandbreite sind AI-Systeme teils nichts weiter als schrittweise Entwicklungen im Zuge von Automatisierung und Digitalisierung, so wie sie bereits seit Jahrzehnten voranschreiten. Daneben gibt es aber auch technologische Sprünge, die den Stand der Technik hinsichtlich einzelner, spezifischer Problemstellungen revolutionieren. Meist kommen dabei neue Ansätze der Daten-Analyse und des Maschinellen Lernens („Machine Learning“) (1) zum Einsatz.

Eine Gemeinsamkeit vieler AI-Systeme ist, dass es zum einen eine große Menge an gesammelten Daten gibt, und zum anderen Modelle, die (laufend) mit diesen Daten gefüttert werden. Die Modelle werden von Systemen benutzt, um kontextabhängige Antworten bereitzustellen. Sehr oft sind das Informationen, Entscheidungs-

gen oder Entscheidungsvorschläge, die von Anwender:innen weiterverarbeitet werden.

Als „intelligent“ werden Software-Systeme meist dann bezeichnet, wenn sie herausfordernde Aufgaben lösen können, die bisher nur Menschen erledigen konnten, und selbstständig dazulernen. Manche technische Systeme versuchen dabei, Entscheidungsstrukturen des Menschen nachzubilden.

Die Anwendungsbereiche von AI-Technologien sind äußerst breit gestreut. Viele Wertschöpfungsprozesse können grundsätzlich unterstützt werden. Abbildung 1 zeigt diese vielseitige Einsetzbarkeit von AI in verschiedenen Anwendungsbereichen in Betrieben. So können etwa die Kommunikation mit Kunden, innerbetriebliche Abläufe, Produktionsprozesse oder die Logistik durch den Einsatz von AI-Technologie unterstützt werden. Durch diese Vielfältigkeit ist AI für einen Großteil der österreichischen Unternehmen ein Thema. In vielen Bereichen werden neue Technologien in Kombination mit AI zu Veränderungen am Markt führen, sodass für die Verbesserung der Wettbewerbsfähigkeit deren Einsatz rechtzeitig geplant werden muss.

„Intelligente“ Software-Systeme können herausfordernde Aufgaben lösen und selbstständig dazulernen. Dabei versuchen sie die Entscheidungsstrukturen des Menschen nachzubilden.

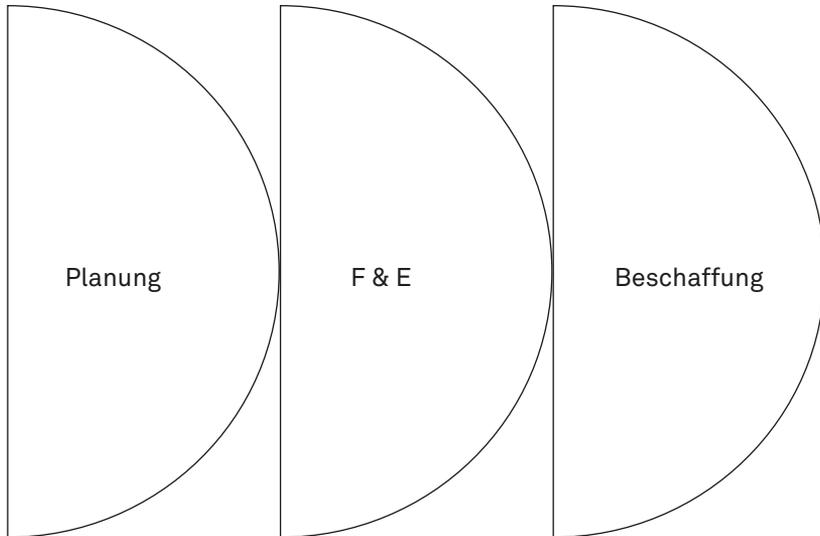
(1) Maschinelles Lernen („Machine Learning“) ist ein Teilgebiet der künstlichen Intelligenz. Durch das Erkennen von Mustern in vorliegenden Datenbeständen sind IT-Systeme in der Lage, eigenständig Lösungen für Probleme zu finden.

Einsatz neuer AI-basierter Technologien in Wertschöpfungsprozessen

Der Einsatz von AI verbindet zunehmend komplexe und heterogene Systeme und sorgt für eine langfristig bessere Integration entlang der Wertschöpfungskette.

Unternehmensorganisation

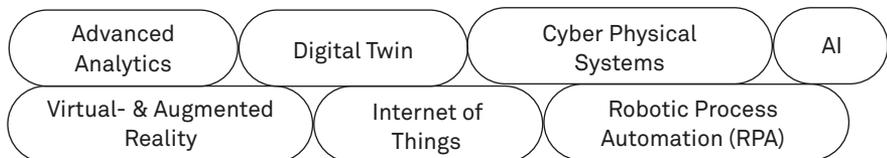
Funktionen



Anwendungen

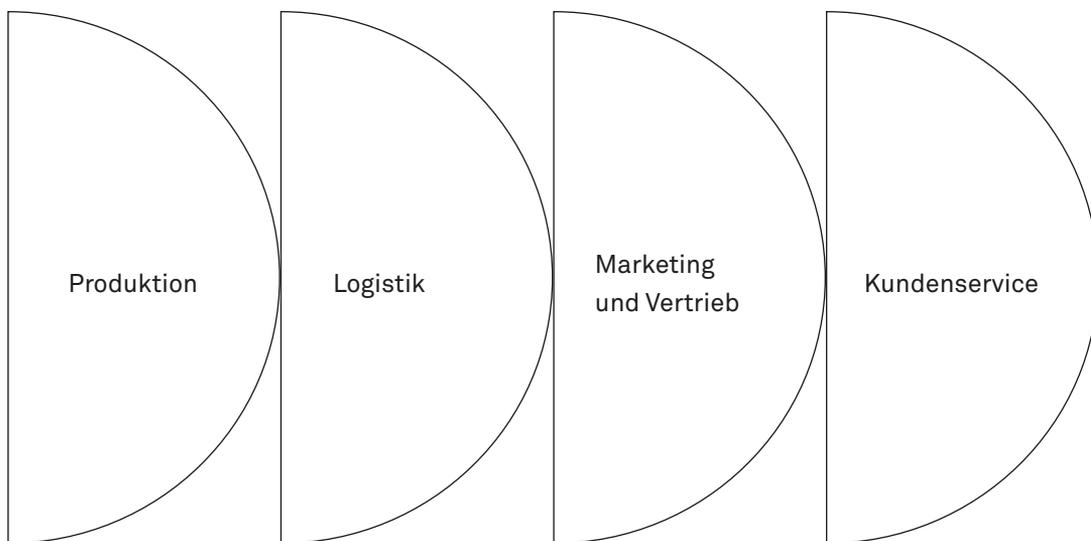
- Bedarfs- und Verkaufsanalysen
- Kapazitätsplanung
- Lieferant:innen-integration
- Generatives Produktdesign
- Kognitive Assistenten im Innovationsprozess
- Markt- und Nutzeranalysen
- Priorisierung von Projekten
- Saisonalitäts-simulationen
- Lieferant:innen-auswahl
- Ausschreibungen
- Produktplanung

Technologien



- IT- und Cyber-Security
- Kognitive Expert:innensysteme
- Intelligente Automatisierung von Unternehmensprozessen
- Unterstützung am Arbeitsplatz
- Unterstützung von Strategieprozessen

und Strategie



- Intelligente Produktionsplanung
- Predictive Maintenance
- Flusssimulationen
- Materialwirtschaftsplanung
- Qualitätskontrolle
- Selbstoptimierende Maschinen und Prozesse

- Speditionsauswahl
- Warenlagermanagement
- End-zu-End Track & Trace

- Verkaufsplanung
- Produktauswahl
- Produktüberwachung
- Bedarfabstimmung

- Produktrücknahmen
- Lagertransfers
- Beschwerdemanagement

Abbildung 1:
AI-Anwendung in unternehmerischen Wertschöpfungsprozessen
(Quelle: Eigene Darstellung)

Die richtigen Daten für den richtigen Zweck

Viele Unternehmen haben bereits Geschäftsprozesse digitalisiert und speichern die dabei anfallenden Daten. Üblicherweise wird aber nur ein Bruchteil dieser Daten genutzt, um etwa neue Erkenntnisse daraus zu ziehen, Trends zu verstehen, Optimierungspotenziale abzuleiten, neue Geschäftsmodelle zu entwickeln oder die Daten über einen Marktplatz zu verkaufen. In Datensets steckt oft außergewöhnlich viel Wissen und es wird eine Sammlung von Fällen und Sonderfällen abgebildet, mit denen Mitarbeiter:innen im betrieblichen Alltag konfrontiert sind. Ein gemeinsamer Nenner vieler AI-Systeme ist es nun, dass aus solchen Datensets mit Machine Learning-Verfahren Modelle errechnet werden, die selbstständig Entscheidungen treffen können oder einen Menschen in der Zusammenarbeit mit Information und Hinweisen unterstützen können.

Oft werden die Vollständigkeit und der Nutzen dieser Daten falsch eingeschätzt und es stellt sich heraus, dass beispielsweise zuerst weitere Sensoren integriert werden müssen, um ein AI-System mit den notwendigen Daten füttern zu können. Eine schrittweise Planung und entsprechende Vorprojekte sind empfehlenswert, um die Qualität der verfügbaren Daten bereits früh zu evaluieren. Unnötige Speicherung von Daten sollte vermieden werden, da dies zu zusätzlichen Kosten und datenschutzrechtlichen Problemen führen kann.

Gefahr der fehlerhaften Verarbeitung von Daten (Bias)

Wird ein AI-System durch fehlerhafte Daten trainiert oder werden diese fehlerhaft verarbeitet, kann es zur Diskriminierung bestimmter Personengruppen oder Minderheiten durch die AI-Entscheidung kommen. Diese Datenverzerrung nennt man Bias. Diese kann auch durch menschliche Fehlentscheidungen aus der Vergangenheit entstehen. Eine weitere Ursache für Bias ist der fehlende ethische Aspekt bei der Entwicklung von KI-Systemen. Erfolgreich ist ein KI-System erst dann, wenn es auch im Sinne des Gemeinwohls funktioniert.

Umgang mit Bias-Problemen

Beispiel aus
der Praxis

Ein holzverarbeitendes Unternehmen entwickelt in Kooperation mit Data Science Spezialist:innen ein AI-System, das bei der Klassifizierung von Holzbrettern unterstützen soll. Auf Basis von optischen Kamera-Daten soll vorgeschlagen werden, für welchen Zweck sich ein vorliegendes Stück Holz am besten eignet. Bisher wurde diese Klassifizierung rein manuell durchgeführt. Das Unternehmen hat dazu über mehrere Jahre zigtausende Beispieldaten gesammelt, die jeweils aus einer Kombination von Bild-Daten und der manuell getroffenen Entscheidung bestehen. Die Beispieldaten umfassen Entscheidungen von ca. 20 Mitarbeiter:innen.

Im ersten Schritt wird aus der Gesamtheit der Beispieldaten mittels Machine Learning-Verfahren ein Modell „trainiert“, das für beliebige neue Bretter einen Vorschlag für die Eignung liefern und die Mitarbeiter:innen dadurch unterstützen soll. Eine Vielzahl an Merkmalen spielt dabei eine Rolle, z. B. die Gleichmäßigkeit von Maserungen oder spezifische Varianten und Größen von verwachsenen Ästen. Das Modell wird von allen Mitarbeiter:innen getestet und die Qualität der Vorschläge evaluiert. Schnell wird klar, dass viele Vorschläge nicht dem Wunsch des Unternehmens entsprechen. Die Data Science Spezialist:innen können die Ursache dafür aufklären: In den Trainings-Daten enthalten ist ein gewisser Anteil an Beispielen mit Zirben-Hölzern, deren Verarbeitung sich stark von den anderen Holz-Arten unterscheidet. Die Berücksichtigung beim Trainieren des Modells hat einen Bias (Datenverzerrung) mit sich gebracht. Einige unbedeutende Merkmale fließen zu stark in den Vorschlag für die Verwendung des Holz-Stückes ein. Durch eine entsprechende Änderung der Gewichtung bzw. einer Aufteilung in zwei getrennte Modelle kann das Bias-Problem behoben werden. Eine Wiederholung der Tests verläuft erfolgreich.

Vor- und Nachteile von AI-Systemen

Der Einsatz von AI-Systemen kann vielfältige positive Auswirkungen haben. In vielen Bereichen wären Unternehmen ohne führende Technologien nicht länger konkurrenzfähig. Viele bestehende Jobs können sich im Zuge der Einführung von AI-Systemen auch verbessern, insbesondere dann, wenn Mitarbeiter:innen von Routine-Tätigkeiten entbunden werden, um sich auf herausfordernde Wissensarbeit konzentrieren zu können, die mit einer persönlichen Weiterentwicklung und besserer Entlohnung einhergeht. AI hat somit nicht nur wirtschaftliches Potenzial für Betriebe, son-

dern auch Potenzial für Unterstützung und Entlastung der Beschäftigten. Gleichzeitig gibt es Funktionalitäten/Aufgaben, die von einem AI-System nicht beherrscht werden können, wie z. B. motivieren oder verhandeln.

Auf der anderen Seite steht aber die Gefahr der Entwertung von (menschlicher) Arbeit, Kontrolle, Überwachung, Diskriminierung, Nicht-Nachvollziehbarkeit und Intransparenz der Entscheidungen und Aktionen der AI. Daher braucht es jedenfalls klare Regeln und geeignete Schutzmaßnahmen für die Beschäftigten. Auf EU-Ebene bestehen im Weißbuch zur Künstlichen Intelligenz der Europäischen Kommission Ansätze, gute und vertrauenswürdige AI zu definieren (2):

- Vorrang menschlichen Handelns und menschlicher Aufsicht
- Technische Robustheit und Sicherheit
- Privatsphäre und Datenqualitätsmanagement
- Transparenz
- Vielfalt, Nichtdiskriminierung und Fairness
- Gesellschaftliches und ökologisches Wohlergehen und
- Rechenschaftspflicht

AI-Auslegung der Plattform Industrie 4.0

Die Plattform Industrie 4.0 versteht unter dem Begriff AI den Einsatz von Daten zur effektiveren Entscheidungsfindung und autonomen Steuerung als Unterstützung für Mitarbeiter:innen im Produktionsbetrieb. AI-Systeme weisen in vielen Fällen selbstlernende Funktionen auf, um daraus Entscheidungsvorschläge abzuleiten, wobei der Mensch die Letztentscheidungsmöglichkeit haben sollte.

Die Abkürzung AI steht für „Artificial Intelligence“ und wird von der Plattform Industrie 4.0 als „Assistive Intelligence“ verstanden. Die AI-Systeme dienen dem Menschen als Unterstützung und helfen diesem bei der Entscheidungsfindung.

Unter AI kann auch „Assistive Intelligence“ verstanden werden, denn AI-Systeme dienen dem Menschen als Unterstützung und helfen diesem bei der Entscheidungsfindung.

Empfehlungen bei der Implementierung einer AI-Lösung

Eine Reihe an Faktoren ist relevant, um AI-Anwendungen erfolgreich in der Praxis einzuführen. Diese lassen sich grob in drei Phasen einteilen: vor, während und nach der Einführungsphase. Im Folgenden werden Empfehlungen anhand dieser Abschnitte dargestellt. (3)

Ganz wesentlich und über alle drei Phasen hinweg erfolgsentscheidend ist die Einbindung der Anwender:innen für das Gelingen des AI-Projektes. Werden durch AI-Anwendungen personenbezogene Daten der Beschäftigten verarbeitet, ist deren Einbindung bzw. die des Betriebsrats rechtlich notwendig. Der Betriebsrat fungiert als Interessensvertretung und Sprachrohr der Mitarbeiter:innen, ist Verhandler für Betriebsvereinbarungen, Informationsdrehscheibe uvm. und kann dadurch eine wichtige Rolle bei der erfolgreichen Implementierung von AI-Lösungen spielen.

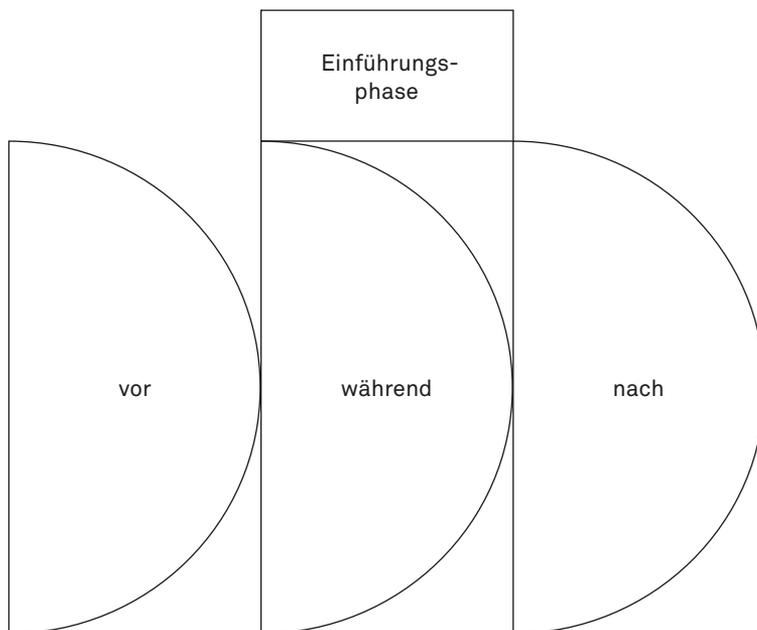


Abbildung 2:
Betrachtung von drei
Phasen innerhalb
der Einführungsphase
(Quelle: eigene
Darstellung)

(2) Diese Erkenntnisse basieren auf Interviews im Rahmen des Projekts „AI for GOOD“, die von April bis Juli 2020 durchgeführt wurden, finanziert durch den AK Digifonds.

(3) WEISSBUCH Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, KOM (2020) 65 endg.

Vor der Einführungsphase

Mit der Einführung neuer AI-Systeme gehen viele Fragen einher, die am besten schon früh in der Entwicklungsphase geklärt werden, unter Einbeziehung aller Betroffenen, besonders den damit direkt betrauten Mitarbeiter:innen. Folgende Punkte sollen schon im Vorfeld abgeklärt bzw. definiert werden:

Zusammensetzung des Projektteams

- Hier entscheidet sich schon der Erfolg des Projektes. Es wird empfohlen alle betroffenen Parteien im Unternehmen mit einzubinden, vor allem frühzeitig. Vom IT-Spezialist:innen bis zu der Person, die die Daten generiert, sollen alle involviert werden.
- Die Etablierung einer Arbeitsgruppe für die Auswahl eines AI-Systems wird empfohlen.
- Auch ein frühzeitiges Einbinden des Betriebsrates (sofern vorhanden) wird empfohlen. Nur wenn das Verständnis für das Projekt breit im Unternehmen verankert ist, kann man vorab Ängste und Bedenken verhindern.
- Falls das AI-System zum Ziel hat, den/die Mitarbeiter:in zu unterstützen, ist es zweckmäßig, dass diesen erläutert wird, was im Hintergrund des Programmes während der Nutzung passiert und welche Daten verarbeitet und gespeichert werden.

Es ist wichtig, frühzeitig alle betroffenen Parteien im Unternehmen mit einzubinden.

Festlegung eines konkreten Einsatzgebietes

inklusive Nutzendefinition und Budget

- Die Frage, welche Zielsetzung durch den Einsatz der AI-Anwendung erreicht werden soll, soll beantwortet werden.
 - / Je genauer und konkreter die Zieldefinition vorgenommen wird, d. h. für welche Anwendung die AI-Lösung vorgesehen wird, umso leichter wird es für die kommenden Implementierungsphasen und für die Wahl einer technischen Lösung. AI ist nicht immer zwingend die optimale Anwendung und dasselbe Ergebnis kann mit anderen Maßnahmen einfacher, besser oder kostengünstiger erreicht werden.
 - / Es gibt vielfältige AI-Systeme, die verschiedene Zwecke erfüllen. Die Zusammenarbeit eines AI-Systems mit dem/der Anwender:in kann unterschiedlich gestaltet sein. Ein AI-System kann beispielsweise Vorschläge liefern oder Entscheidungen autonom treffen, die ein/eine Mitarbeiter:in gegebenenfalls abweisen kann. Diese Art der Zusammenarbeit gilt es, vorab klar zu definieren. Dazu gehört die Frage, wer die Verantwortung für finale Entscheidungen trägt.
 - / Je klarer der erwartete Nutzen beschrieben wird, desto einfacher ist die konkrete Umsetzung. Dabei kann der Nutzen unterschiedliche Dimensionen umfassen, z. B. Reduktion der Fehlerquote, Entlastung von Mitarbeiter:innen, bessere Planung der Logistik.
 - / Grundsätzlich sollten AI-Anwendungen zweckmäßig, zielführend und datenschonend sein.
- Bei der Auswahl von Trainingsdaten für Algorithmen basierend auf Machine Learning muss sorgfältig analysiert werden, ob die Daten unerwünschte Verzerrungen beinhalten, die zu Bias-Problemen führen würden, beispielsweise Gruppen von Anwender:innen benachteiligen würden.
- Eine Erhebung, welche Daten im Unternehmen für den spezifischen Use Case bereits vorhanden sind und welche zusätzlich erhoben werden müssen, ist notwendig.

- Eine zumindest grobe, aber realistische Einschätzung des Budgetbedarfs soll durchgeführt werden.

Vermeidung von Biases und Risiko- bzw. Folgenabschätzung

- Nach dem risikobasierten Ansatz ist eine grobe Abschätzung der Kosten, des Nutzens und des Risikos ratsam, beispielsweise welche Konsequenzen es gibt, wenn das System Fehlentscheidungen trifft.
- Durchführung einer Technikfolgenabschätzung: AI-Anwendungen sind immer mit Anwendungen weiterer Technologien gekoppelt (z. B. Sensoren wie Kameras, Maschinen, die miteinander kommunizieren) Entsprechend können bei der Analyse von Technikfolgen frühzeitig Probleme identifiziert und Fragen rund um diese Technologien und deren Zusammenwirken beantwortet werden.

Mit einer Risiko- bzw. Technikfolgenabschätzung können frühzeitig Probleme identifiziert werden.

Beachtung der gesetzlichen Rahmenbedingungen

- Bei der Implementierung eines AI-Systems sind gesetzliche Vorgaben zu beachten, insbesondere dann, wenn Rechte von Menschen berührt werden. Beispielsweise muss über bestimmte Maßnahmen/Systeme vorab der Betriebsrat informiert (§ 91 Abs 1 und 2 ArbVG) und mitunter eine Betriebsvereinbarung abgeschlossen werden (§§ 96 Abs 1 Z 3, 96a Abs 1 Z 1 ArbVG).
- Neben dem ArbVG ist auch die Beachtung der Datenschutzvorschriften (DSGVO) wichtig. Um weitere Informationen zu erhalten, wird die Kontaktaufnahme mit der jeweiligen Interessensvertretung oder sonstigen Expert:innen empfohlen. (4)
- Aktuell gibt es verschiedene Überlegungen zur Regulierung des AI-Einsatzes auf europäischer Ebene. Während derzeit noch zahlreiche Diskussionen zum Vorschlag der Europäischen Kommission (2021/0106(COD)) stattfinden und noch keine konkrete Verordnung existiert, gilt es beim zukünftigen Einsatz von AI-Systemen diese Entwicklungen im Auge zu behalten.

Rund um die Verwendung von AI-Systemen müssen gesetzliche Vorgaben wie das Arbeitsverfassungsgesetz und die Datenschutzgrundverordnung, beachtet werden.

Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen

Klärung, ob personenbezogene Daten betroffen sind

- Kann ein AI-System so ausgestaltet werden, dass es keine personenbezogenen Daten verarbeitet, z. B. weil diese nicht benötigt oder ausreichend anonymisiert/pseudonymisiert werden, entfallen Datenschutz-Rechtspflichten.
- Zunächst sollte eine Erfassung/Beschreibung, welche Daten erhoben werden (müssen) und was mit diesen Daten passiert, im AI-Projekt erfolgen. Kann ein Personenbezug eliminiert werden, ist diese Variante zu präferieren.
- Kann ein Personenbezug nicht zur Gänze ausgeschlossen werden, ist dieser auf das notwendige Mindestmaß zu reduzieren.
- Zugleich sollte eine Evaluierung erfolgen, welche Zwecke/Ziele mit der beabsichtigten Verarbeitung personenbezogener Daten verfolgt werden. Welche Auswirkungen, Risiken etc. ergeben sich daraus für die Mitarbeiter:innen? Wer benötigt Zugang zu Daten?
- Nach den Datenschutzvorschriften sind auch die Aspekte der Richtigkeit und Sicherheit der personenbezogenen Daten vor Verlust und Zugang Dritter zu beachten. Neben technischen kommen dafür auch organisatorische Maßnahmen in Betracht. Die Schaffung eines klaren und transparenten Konzepts für Zugangsberechtigungen ist notwendig.
- Löschkonzepte müssen erarbeitet werden. Wer ist für die Löschung faktisch verantwortlich? Wie lange müssen Daten für den Zweck/das Ziel unbedingt erhoben/gespeichert werden. Grundsätzlich dürfen personenbezogene Daten nur so lange verarbeitet und gespeichert werden, wie dies unbedingt erforderlich ist.

Grundsätzlich soll die Aufzeichnung personenbezogener Daten durch die Verwendung von AI-Systemen vermieden werden. Kann ein Personenbezug nicht zur Gänze ausgeschlossen werden, ist dieser auf das notwendige Mindestmaß zu reduzieren.

Strukturierung der Daten, Prüfung, ob sie konsistent,

durchgängig, fehlerfrei etc. sind

- Der gesamte Lebenszyklus von der Erhebung bis zur Löschung von Daten soll durchdacht sein. Ist es möglich bzw. nötig, die Daten zu pseudonymisieren und anonymisieren? Wurden die Daten auf Schiefen (Bias) geprüft? Welche Verwendungszwecke sind vorgesehen und welche klar ausgeschlossen? Welche Daten mit Personenbezug sind zu löschen, wenn jemand das Unternehmen verlässt? Für alle Phasen sind klare Prozessschritte und vor allem verantwortliche Personen festzulegen. Unternehmen wären schlecht beraten, die Zuständigkeiten dafür nur schwammig zu regeln oder zuzulassen, dass Regelungen missachtet werden.
- Viele Daten sind unstrukturiert, in nur schwer oder nicht verarbeitbaren Formatierungen vorliegend, nicht konsistent, mit Fehlern behaftet, etc. Bevor das System startet, sollte eine Datenbereinigung, die diese Punkte umfasst, durchgeführt werden.
- Für die Datenbereinigung ist es wichtig, ausreichend Zeit und Ressourcen einzuplanen, um die adäquate Datenbasis für die langfristige Funktionstüchtigkeit des AI-Systems sicherzustellen.

Durchführung differenzierter Schulungsmaßnahmen

- Abhängig von der Qualifikation und vom Anwendungsgrad (Programmierer:in, Anwender:in, Projektleiter:in, Führungsebene ...) ist es wichtig, individuell angepasste Schulungen und Weiterbildungen zu ermöglichen
- Schulungen zur besseren Bedienbarkeit sind wichtig und können wertvolle Rückschlüsse auf die Usability des Systems ermöglichen.

Schulungen und Weiterbildungen helfen Mitarbeiter:innen die AI-Systeme und ihre Anwendung besser zu verstehen und mehr Transparenz in die Entscheidungsprozesse zu bringen.

Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen

Während der Einführungsphase

Initiierung der Zusammenarbeit zwischen IT-Spezialist:innen, die das System entwickeln, und den Personen, die mit dem System arbeiten werden

Entscheidend ist das Wissen, um die Algorithmen und das Domänenwissen (Fachwissen) in Einklang zu bringen.

Wenn Machine Learning eingesetzt wird, um aus historischen Beispiel-Daten zu lernen

- Ausreichend Zeit einplanen, um das System lernen zu lassen. Allen Beteiligten muss bewusst sein, dass es eine „Kaltstart“-Phase gibt, in der noch zu wenige Daten verfügbar sind, und daher die Qualität der Modelle noch vermindert ist.
- Sollten Trainingsdaten extern bezogen werden, sind diese vorab auf Schiefen (Bias) zu prüfen oder prüfen zu lassen, um das System nicht fehlerhaft lernen zu lassen.

Ausreichend lange Testphase, um das System validieren zu können

- Für die Einführung neuer Systeme ist dies essenziell, um die Qualität der Ergebnisse bestimmen zu können.
- Eine realistische Erwartungshaltung aller Beteiligten verhindert vielschichtige Probleme.
- Referenzprojekte anderer Unternehmen/Branchen können eine gute Grundlage für die Abschätzung des Potenzials von AI-Systemen sein.

Ausreichend lange Testphasen und eine realistische Erwartungshaltung der Beteiligten erhöhen die Erfolgchancen.

(Unterstützung) Die Universität Helsinki hat mit einem Unternehmen eine Plattform mit kostenlosen Onlinekursen entwickelt. Das Ziel von „The Elements of AI“ ist, das Thema AI möglichst vielen Menschen niederschwellig und leicht verständlich zugänglich zu machen. Näheres unter: www.elementsofai.de

Ausreichend Zeit einplanen, damit das System an Anforderungen der Anwender:innen angepasst werden kann bzw. sich die Anwender:innen an das System gewöhnen können

- Die Einführung großer neuer Systeme geht mit der Veränderung von Arbeitsprozessen, Zuständigkeiten, Verantwortlichkeiten und letztlich Job-Profilen einher. Dies muss im Sinne eines umfassenden Change Management-Prozesses durchdacht und individuell kommuniziert werden. Es empfiehlt sich eine „sozio-technische“ Betrachtung, bei der das Zusammenspiel zwischen Menschen und AI-System als Einheit verstanden wird und die menschlichen Bedürfnisse und Faktoren dabei nicht ausgeblendet werden.
- Thema Verantwortung: Es muss die Frage geklärt werden, wer entscheidet, ob dieser Vorschlag/das Ergebnis akzeptiert wird oder nicht.
- Wenn beispielsweise Vorschläge nach Wahrscheinlichkeiten geordnet angeboten werden, müssen dabei Spielregeln in Form von Untergrenzen definiert werden.

Mit Hilfe einer AI-Funktion in einem Kundenmanagementtool wird angezeigt, dass ein bestimmter Kunde mit 66 % Wahrscheinlichkeit das Kundenverhältnis aufkündigen wird und schlägt Maßnahmen zur Vermeidung der Kündigung vor. Eine Wahrscheinlichkeitsrate von 66 % kann die Entscheidung des/der Mitarbeiter:in beeinflussen und zur Setzung von Maßnahmen (z. B. Zusendung eines Angebots) verleiten. Wie sollen Mitarbeitende mit dieser Information umgehen? Eine einheitliche Lösung gibt es hier nicht. Auch die Frage, ob 66 % ausreichend sind, kann nur mit viel Erfahrung in genau diesem Arbeitskontext beantwortet werden. Jedenfalls aber hilft die AI-Lösung, keine Kunden zu übersehen, die eventuell wechseln möchten. Was man dann mit dieser Information tut, liegt bei dem/der Mitarbeiter:in, genau wie bisher.

Beispiel aus der Praxis

(Unterstützung) Um Verantwortlichkeiten strukturiert durchzudenken, empfiehlt sich das kostenfreie Online-Tool „VERA“. Dieses hilft dabei, Fragen der Verantwortlichkeit bei der Einführung von AI zu beantworten und Verantwortungslücken zu identifizieren. Näheres unter: www.vera.arbeiterkammer.at

Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen

Interpretierbarkeit & Transparenz: Der/die Mitarbeiter:in soll das

Ergebnis selbst bestmöglich verstehen bzw. nachvollziehen können.

Für die Anwender:innen des spezifischen AI-Systems sind sowohl die Art der Information (z. B. einzelne Werte oder Zahlenbereiche) als auch die Aufbereitung dieser (z. B. über ein Dashboard) relevant. Ausführliche Überlegungen diesbezüglich werden empfohlen.

Sicherstellung der Qualität der Daten

- Es ist immer Sorge zu tragen, dass Datenverzerrungen (Bias) korrigiert, oder zumindest hinreichend transparent kommuniziert werden. Ein Modell von Daten bildet immer jene Aussagen nach, die in diesen Daten stecken. Sollten diese Daten einen Schiefstand aufweisen, dann wird auch das Modell verzerrt sein.
- In der Informatik wird oft die Phrase „Garbage In, Garbage Out“ (GIGO) verwendet, die auf die Notwendigkeit der qualitativ hochwertigen Datenbasis hinweist. Insbesondere bei AI-Systemen ist dieses „Prinzip“ ebenfalls zu beachten.

Berücksichtigung organisatorischer Veränderungen, die sich

durch den Einsatz eines AI-Systems ergeben könnten

(Verantwortung, Zusammenarbeit, Arbeitsorganisation, ...)

- Der Begriff „Verantwortung“ kann in diesem Zusammenhang viele Bedeutungen haben, sich etwa auf Zuständigkeit, Rechenschaft oder Haftung beziehen. Abgesehen von Unternehmen, die AI-Systeme einsetzen, tragen auch die Anbieter:innen und Entwickler:innen solcher Systeme Verantwortung.
- Um den Entscheidungsprozess der Mitarbeiter:innen (ob Maßnahmen zu treffen sind oder nicht) effizienter zu unterstützen, soll das AI-System seine Vorschläge begründen.
- AI-Systeme und ihre Nutzung unterliegen Unternehmenswerten wie zum Beispiel Leitlinien und „Code of Conducts“, ethischen Grundwerten sowie gesellschaftlichen Konventionen. Denn die Technologie ist nicht wertfrei.

- Ein betrieblich vereinbartes Modell für agiles kooperatives Change-Management stärkt sowohl ein gemeinsames Experimentieren als auch eine kooperative Führungskultur. Diese kann zum einen dem Betrieb Kosten sparen und zum anderen Fachkräfte anziehen. (5)

Nach der Einführungsphase

Evaluierung des Systems, ob die gesetzten Ziele erreicht wurden

(Effizienzsteigerung, Reduktion der Arbeitsbelastung etc.)

- Erfüllt das AI-System die Anforderungen und führt so zu einer hohen Akzeptanz der Mitarbeiter:innen, wird eine effektive Nutzung des Systems wahrscheinlicher.
- Blindes Vertrauen in AI-Systeme ist ein Risiko, welches berücksichtigt werden muss. Als Gefahr sei hier z. B. das unberechtigte Vertrauen auf die aktuellen Autopiloten autonomer Fahrzeuge genannt, welches zu Unfällen führte. Ein sanfterer Ansatz wäre es, Anwender:innen öfter die Unsicherheit der jeweiligen AI-Ergebnisse mitzuteilen. Ein solcher Ansatz kann mit der Idee der „erklärbaren AI“ verbunden werden: Unter dem Begriff „Explainable AI“ wird nach technischen Möglichkeiten geforscht, den Anwender:innen eine Erklärung für die Vorschläge und Entscheidungen von AI-Systemen anzubieten. Wie es dazu im konkreten Fall kommt und welche Aspekte wie stark dabei eingeflossen sind, ist bei hochkomplexen Machine Learning-Modellen nicht einfach zu vermitteln.
- Regelmäßige Evaluierungen mit den Anwender:innen führen zu einer nachhaltigen Akzeptanz der AI-Systeme.

Ein neues AI-unterstütztes System sollte zur Erkennung von Qualitätsmängeln in der Produktion eingeführt werden. Allerdings kam die Hälfte der auftretenden Fehler nur sehr selten vor. Daher gab es nicht genügend Fehlerfälle, um daraus typische Eigenheiten zu „lernen“. Für die andere Hälfte der Fehler gab es genügend Fälle. Das System wird im Einsatz folglich alle oft auftretenden Fälle gut erkennen. Dennoch wird es weiterhin Fehler geben, die das System nicht erkennt. Von den Mitarbeiter:innen darf folglich nicht erwar-

Beispiel aus
der Praxis

Praxisleitfaden zur erfolgreichen Einführung von AI-Systemen

tet werden, sämtliche restlichen Fehler, die dem System nicht aufgefallen sind, zu finden. Vielmehr sollte man es so sehen, dass Mitarbeiter:innen durch das semi-automatische Finden von öfter auftretenden Fehlern entlastet werden und sich größtenteils auf die schwierigen Fehler und deren Lösungen fokussieren können.

Aufzeigen von Interventionsmöglichkeiten beim System

Anwender:innen des Systems sollen die Möglichkeit haben, einzugreifen. Die Endentscheidungsmacht haben die Anwender:innen, die Grundlage dazu liefert das System.

Weiterbildungsmaßnahmen

Bei Bedarf sollen laufende Weiterbildungsmaßnahmen, die im Idealfall individualisiert sind, initiiert werden.

Nutzung der generierten Daten

- Laufende Kontrolle ist unerlässlich, um auch im Betrieb Schieflagen (Bias) zu verhindern.
- Es muss sichergestellt werden, dass Daten nicht für weitere, nicht genehmigte Zwecke, verwendet werden.

Bei der Bestimmung der Qualitätsklasse von Baumstämmen wird ein AI-System eingesetzt. Dieses hilft Mitarbeiter:innen bei der Entscheidung, indem es sie bei der Argumentation der Qualitätsbestimmung unterstützt und so die ermüdende und monotone Arbeit erleichtert. In diesem Fall kann die Arbeit nur dann erfolgreich durchgeführt werden, wenn die Mitarbeiter:innen dem System vertrauen und die Ergebnisse akzeptieren.

Beispiel aus der Praxis

(5) Zum Thema kooperatives Changemanagement siehe auch: https://www.offensive-mittelstand.de/fileadmin/user_upload/pdf/uh40_2019/2_1_4_kooperatives_changemanagement.pdf

Der Mensch als wichtigster Faktor

Der Einsatz von AI-Systemen dient dazu, den Menschen in seinen Tätigkeiten zu unterstützen. Entsprechend müssen Bedingungen geschaffen werden, unter denen sich der Mensch im Umgang mit dem AI-System wohlfühlt und dem System vertraut. Daher gibt es eine Reihe an positiven und negativen Akzeptanzfaktoren im betrieblichen Umfeld.

Positive Akzeptanzfaktoren

Viele erfolgreiche AI-Systeme arbeiten nicht völlig autonom, sondern sind konzipiert, um den Menschen bestmöglich zu unterstützen. Die Mitarbeiter:innen stehen im Mittelpunkt, können frei entscheiden und bleiben damit in der Verantwortung. AI sollte nur Vorschläge liefern, diese im besten Fall auch mit einer Wahrscheinlichkeit oder mit einer Begründung untermauern. Die Entscheidung, ob der Vorschlag angenommen wird oder nicht, treffen weiterhin die Mitarbeiter:innen. Deshalb ist es wichtig, dass Mitarbeiter:innen auch über die Funktionsweise und technische Grenzen von solchen AI-Systemen Bescheid wissen.

Es empfiehlt sich meist nicht, AI-Systeme so zu gestalten, dass sie die Mitarbeiter:innen ersetzen, indem sie deren Aufgaben möglichst vollautomatisch übernehmen. Stattdessen sollte versucht werden, die Mitarbeiter:innen in ihren Tätigkeiten so zu unterstützen, dass sie bessere und fundiertere Entscheidungen treffen können. Ein Degradieren zum monotonen Monitoring von hochgradig automatisierten Systemen kann verhindert werden.

Mitarbeiter:innen sollen die Chance erkennen und nützen, ihren Aufgabenbereich positiv zu verändern. Dafür braucht es eine offene und transparente Kommunikation. Während die AI bei Routineaufgaben entlastet, können sich Mitarbeiter:innen auf herausfordernde Aufgaben konzentrieren. Um bestmöglich und effizient mit AI-Systemen umgehen zu können, sind Weiterbildungsmaßnahmen nötig. (6)

AI-Systeme dienen lediglich als Unterstützung, die Letztentscheidung bleibt bei der/dem Mitarbeiter:in. Die Lösungsvorschläge der AI-Systeme müssen für den/die Mitarbeiter:in nachvollziehbar sein.

- AI-Grundlagen: Für ein solches Verständnis wird oft der Begriff „AI Literacy“ verwendet. Dabei soll erklärt werden, was AI ist, wie AI-Algorithmen funktionieren, was AI-Systeme können und nicht können und was die Anwendung für Mitarbeiter:innen bedeutet.
- Soziale Aspekte: AI hat Auswirkungen auf Arbeit und Menschen. Damit soll bewusst umgegangen werden.
- Wirtschaftliche Implikationen des Einsatzes von AI-Technologien
- Rechtliche Rahmenbedingungen und Gesetze, die etwa Arbeitsrecht und den Datenschutz betreffen

Negative Akzeptanzfaktoren

Allzu oft werden den menschlichen Aspekten abseits der technischen Problemstellungen zu wenig Aufmerksamkeit geschenkt. Im Folgenden sind einige Aspekte aufgezählt, die von Unternehmen beachtet werden sollten.

- Verfehlen der Ziele, da es seitens der Mitarbeiter:innen Bedenken und Unklarheiten gegenüber der neuen Technologie gibt, und daher die Akzeptanz nicht gegeben ist.
- Bias unterschiedlicher Art, etwa aus einer ungünstigen Gewichtung von Eigenschaften in Datensätzen, die zu Trainieren von Machine Learning-Modellen verwendet werden. Dies kann Auswirkungen wie Diskriminierung oder einen Mangel an Fairness bei vom System getroffenen Entscheidungen haben.
- Intransparenz von AI-Algorithmen und dadurch schlechtere Zusammenarbeit zwischen Menschen und AI
- Interessen und Rechte von Mitarbeiter:innen werden verletzt, z. B. durch Überwachung mittels Datensammlung, Kameras.
- Konflikt mit gesetzlichen Rahmenbedingungen
- Ängste von Mitarbeiter:innen, die durch nicht ausreichende Kommunikation hervorgerufen werden.
- Ängste von Mitarbeiter:innen durch neue technische Systeme den Job zu verlieren oder sich gravierend umorientieren zu müssen.

Folgende Punkte zeigen negative Einflussfaktoren auf

- Arbeitnehmer:innen haben ein Gefühl der Überwachung.
- Umgang mit personenbezogenen Daten ist nicht klar
 - / Direkte personenbezogene Daten (z. B. Analyse von Arbeitsgewohnheiten): Diese erfordern eine freiwillige Einverständniserklärung durch die Arbeitnehmer:innen, wobei keine „gezwungene Freiwilligkeit“ entstehen darf.
 - / Indirekt personenbezogene Daten (z. B. können über die Maschineninteraktion Rückschlüsse auf die Mitarbeiter:innen gezogen werden): Hierbei können Betriebsvereinbarungen und Schulungen zur Vertrauensgewinnung zum Einsatz kommen.
- Verlust von Arbeitsplätzen wird angenommen.

Vertrauen in die AI ist ein wesentlicher Akzeptanzfaktor

Die Funktionsweise der AI ist verständlich/einfach erklärt. Fortschritte in der AI-Forschung können dazu führen, dass AI-Modelle schwerer erklärbar werden, wenn beispielsweise AI-Systeme durch andere AI-Systeme oder Algorithmen trainiert werden. Ansätze aus dem Gebiet der erklärbaren AI (Explainable AI) wirken hier entgegen: Man versucht Systeme zu bauen, die eine Begründung für ihre Vorschläge und Entscheidungen liefern können. Das erlaubt Mitarbeiter:innen, solche Systeme besser nachzuvollziehen und über Interventionen zu entscheiden.

Beispiel aus der Praxis

AI-Beispiele aus der Praxis

Im Folgenden veranschaulicht eine Sammlung von Anwendungsfällen die Vielfalt von AI-Technologien. Es werden Erfolgsfaktoren basierend auf empirischen Erfahrungen angeführt. Die Anregungen können dabei helfen, weitere Potenziale für den Einsatz von AI-Technologien im eigenen Unternehmen zu entdecken.

Effektives Zeit- und Selbstmanagement

Eine Applikation zur Förderung des effektiven Zeit- und Selbstmanagements (mittels Self-Coachings) wurde im Rahmen eines EU-Projektes entwickelt. Ziel war es, effektives Selbst- und Zeitmanagement in Hinblick auf die Bildschirmarbeit zu erforschen und zu unterstützen. Die große Klammer war „Reflektion am Arbeitsplatz“, also die Selbstreflektion über die eigene Arbeitssituation. Die Applikation protokolliert Aktivitäten auf dem eigenen PC, bietet eine Ist-Analyse davon, welche Tätigkeiten im Arbeitsalltag durchgeführt werden und hilft dabei, über vergangene Arbeitstage und ihre Arbeitsinhalte über längere Zeiträume zu reflektieren. Das Programm zeichnet auf, welche Applikationen, Dokumente und Webseiten bei der Bearbeitung von Aufgaben verwendet werden. Wenn die Applikation dazu verwendet wird diese Aufgaben zu erfassen, können später auch alle damit verbundenen Ressourcen schnell und einfach wiedergefunden werden. Die persönlichen und aufgezeichneten Aktivitätsdaten sind nur den Anwender:innen selbst zugänglich und bleiben auf deren Computern. Sie können bei Bedarf händisch bearbeitet und bereinigt werden. Das Programm erhebt verwendete Ressourcen und Applikation auf Rechnern mit Windows-Anwendungen, d.h. wie oft, wie lange und wann diese verwendet wurden. Dabei wird den Anwender:innen transparent kommuniziert, welche Daten gespeichert und verarbeitet werden.

Der Vorteil der Anwendung ist eine objektive Aufzeichnung der Arbeitsweise (Ressourcen/Applikationen) während der Arbeit, um darüber für sich selbst zu reflektieren und das eigene Zeit- und Selbstmanagement gegebenenfalls zu verbessern. Dadurch kann effizienter gearbeitet werden bzw. auch um Hilfe angesucht werden, falls man aktuell zu viele Tätigkeiten parallel ausüben sollte. Es geht darum, das eigene Zeit und Selbstmanagement zu verbessern, und nicht darum, die Produktivität im Vergleich zu anderen Mitarbeiter:innen, Abteilungen etc. darzustellen.

Erfolgsfaktoren des Beispiels vor, während und nach der Einführungsphase

Vor der Einführungsphase

- Beachtung der menschenzentrierten Sicht, Ziel ist Unterstützung der Mitarbeiter:innen
- Den Mitarbeiter:innen zu vertrauen anstatt auf Überwachung zu setzen, trug zur Akzeptanz und Produktivitätssteigerung bei.
- Transparente Kommunikation wurde eingehalten, speziell bei der Verarbeitung personenbezogener Daten.
- Zweckmäßiges und datenschonendes Vorgehen der Anwendung war im Fokus.

Während der Einführungsphase

- Geeignete Testumgebungen wurden zur Verfügung gestellt, mit der die Anwender:innen die Interaktion mit dem System ohne reale Konsequenzen trainieren und verstehen konnten.
- Feedback der Anwender:innen wurde eingeholt, das System daraufhin angepasst und damit auf möglichst hohe Akzeptanz der Anwender:innen ausgelegt.

Nach der Einführungsphase

- Das neue AI-System wurde als laufendes Projekt verstanden, das immer wieder gewartet werden muss, etwa um Bias-Problemen entgegenzusteuern, die über die Zeit entstehen.
- Laufend wurden Schulungsmaßnahmen wiederholt und die Kommunikation zum Zweck des Systems und zur Verarbeitung von Daten ausgerichtet.

Produktionsüberwachung

Die Verwendung von Metaphern aus der Natur für die Darstellung des Zustandes von Maschinen wurde in einer Forschungseinrichtung wissenschaftlich untersucht und evaluiert. Durch die Visualisierung mit Hilfe von Bäumen und die jeweilige Farbe der Pflanzenblätter werden Maschinenzustände abgebildet. Dadurch wird der aktuelle Zustand komplexer Maschinen oder Anlagen einfacher und menschenfreundlicher vermittelt (z. B. vitaler grüner Baum > alles OK; welker Baum > Probleme in der Anlage).

Ziel dieses AI-Projekts war eine Komplexitätsreduktion durch die Verwendung von Augmented/Virtual Reality-Technologien an der Schnittstelle mit Anwender:innen. Viele Bedienoberflächen für professionelle Systeme sind durch eine hohe Komplexität der dargestellten Informationen, der angebotenen Bedienfunktionen und der darauf bezogenen Arbeitssituationen gekennzeichnet. Daraus entstehen meist hohe Anforderungen an die kognitive Leistungsfähigkeit von Personen und die Erlernbarkeit dieser Systeme. Es wurde eingangs auch untersucht, ob Endbenutzer:innen signifikante Unterschiede in der Visualisierung durch „traditionelle“ Industrie 4.0-Dashboards (Charts, Zahlenreihen, Symbole etc.) und eben Bäumen als Metapher sehen. Dabei wurde gezeigt, dass sich Anwender:innen sogar besser über den aktuellen/sich gerade ändernden Status von Maschinen informiert fühlen, wenn dafür diese natürlichere Repräsentation verwendet wird. Personenbezogene Daten wären durch die Schichtpläne zur Verfügung gestanden, jedoch wurden diese nicht benötigt und daher nicht integriert. Lediglich Maschinen- und Sensordaten wurden erhoben. Zugriff auf die vorhandenen Daten hatten Anwender:innen der Produktionssysteme und Forscher:innen. Die Anwendung unterstützt im aktuellen Prozess die Überwachung der Maschinen/Geräte und soll die Komplexität reduzieren und das Treffen schnellerer Entscheidungen fördern.

- Bei AI-Systemen wurde nicht nur an intelligente Algorithmen gedacht, sondern auch auf die Zusammenarbeit und Interaktion mit den Mitarbeiter:innen fokussiert. Besonders erfolgreich war, unvoreingenommen mehrere Interaktionsmöglichkeiten in Betracht zu ziehen und jene zu wählen, die Mitarbeiter:innen am besten unterstützte. Beachtung der menschenzentrierten Sicht, Ziel ist Unterstützung der Mitarbeiter:innen.
- Wenn Menschen die Zustände von Maschinen monitoren sollen, sind menschliche Aspekte besonders stark zu berücksichtigen. Die Konzentration von Menschen ist nicht immer gleich. In gewissen Situationen ist eine der Situation entsprechende Intervention äußerst rasch nötig. Entsprechend wurden Tätigkeiten des stundenlangen passiven Monitorings vermieden. Stattdessen wurde der Fokus darauf gelegt, wie im Anlassfall Mitarbeiter:innen rasch und richtig reagieren können.
- Unabhängige Forscher:innen wurden eingebunden, um die Akzeptanz neuer Systeme zu evaluieren.

Qualitätsbestimmung Baumstamm

In einem Sägewerk werden pro Minute mehrere Baumstämme verarbeitet. Die Mitarbeiter:innen entscheiden, welcher Qualitätsklasse der Baumstamm entspricht und bestimmen damit den Preis, welchen das Unternehmen den LieferantInnen zahlt. Dabei besteht ein Interessenskonflikt, Meinungsverschiedenheiten können entstehen. Dies kann durch eine standardisierte und nachvollziehbare Bewertung vermieden werden. Die Entscheidung, die immer wieder gefällt werden muss, ist für die Mitarbeiter:innen sehr ermüdend und daher hat sich das Unternehmen für eine AI-basierte Unterstützung zur objektiven Beurteilung der Qualität der Baumstämme auf Basis automatisch erstellter Bilder entschieden. Die Technologie soll den Mitarbeiter:innen bei der Argumentation der Qualitätsbestimmung helfen, mehr Transparenz der Bewertung schaffen und das Aufgabenfeld der Mitarbeiter:innen positiv verändern. Mühsame und monotone Aufgaben werden durch die AI übernommen. Die Mitarbeiter:innen behalten den Überblick, können andere Aufgaben übernehmen bzw. sollen ihre Expertise in unklaren Bewertungen einbringen. Diese lernt dadurch ständig weiter und passt sich an Veränderungen an. Die Änderungen der AI-Lösung können nach Bedarf von beiden Seiten (Lieferant:innen/Sägewerk) evaluiert und als neue Version produktiv geschaltet werden. Die AI-basierte Lösung stellt dabei ein unterstützendes Werkzeug für die Mitarbeiter:innen dar.

Erfolgsfaktoren des Beispiels vor, während und nach der Einführungsphase:

Vor der Einführungsphase

- Interpretierbarkeit & Transparenz: Die Mitarbeiter:innen sollten das Ergebnis selbst bestmöglich verstehen bzw. nachvollziehen können. Für eine produktive Zusammenarbeit war es wichtig, dass die Vorschläge und Entscheidungen des AI-Systems nachvollziehbar und „erklärbar“ waren („Explainable AI“). Mit diesem System wird den Mitarbeiter:innen geholfen, mehr Transparenz in die Bewertung zu bringen.
- Es wurde darauf geachtet, dass AI-Anwendungen zweckmäßig, zielführend und datenschonend sind.

Während der Einführungsphase

- Ausreichend Zeit wurde eingeplant, damit das System an Anforderungen der Anwender:innen angepasst werden kann bzw. sich die Anwender:innen an das System gewöhnen konnten. Die Einführung dieses neuen Systems ging mit der Veränderung von Arbeitsprozessen, Zuständigkeiten, Verantwortlichkeiten und letztlich Job-Profilen einher. Dies wurde im Sinne eines umfassenden Change Management Prozesses durchdacht und individuell kommuniziert. Das System übernimmt mühsame und monotone Aufgaben. Mitarbeiter:innen behalten den Überblick, können andere Aufgaben übernehmen bzw. sollen ihre Expertise bei unklaren Bewertungen durch die AI einbringen. Es wird davon abgeraten, die Zusammenarbeit zwischen Menschen und AI-System so zu gestalten, dass die Anwender:innen für längere Zeiträume nur noch zum passiven Überwachen vollautomatischer Systeme herangezogen werden. Aspekte wie Angst vor Langeweile, das Halten von Konzentration und gegebenenfalls eine schnelle und richtige Reaktion sowie das Verständnis, für eine sinnvolle Aufgabe zuständig zu sein, spielten für die Einführung dieses Systems eine essenzielle Rolle.
- Wenn Machine Learning eingesetzt wird, um aus historischen Beispiel-Daten zu lernen, muss ausreichend Zeit eingeplant werden, um das System lernen zu lassen, wie es in diesem Use Case der Fall war. Allen Beteiligten war klar, dass es eine „Kaltstart“-Phase gibt, in der noch zu wenige Daten verfügbar sind, und daher die Qualität der Modelle noch vermindert ist. Das System lernte ständig weiter und passte sich an Veränderungen an.

Nach der Einführungsphase

- Das System wurde evaluiert und untersucht, ob die gesetzten Ziele erreicht wurden (Effizienzsteigerung, Reduktion der Arbeitsbelastung etc.). Erfüllt das AI-System die Anforderungen und führt so zu einer hohen Akzeptanz der Mitarbeiter:innen, wird eine effektive Nutzung des Systems wahrscheinlicher. Die Änderungen der AI-Lösung können nach Bedarf von beiden Seiten (Lieferant:innen/Sägwerk) evaluiert und als neue Version eingesetzt werden.

Danksagung

Wir danken der Arbeiterkammer als Initiator des Digifonds, unseren Konsortialpartnern, Mitgliedern der Expert:innengruppe „Mensch in der digitalen Fabrik“ und allen mitwirkenden Unternehmen, Forschungseinrichtungen und Interessensvertretungen für die Teilnahme und Arbeit am Projekt AI for GOOD.

Besonderer Dank gebührt dabei dem Redaktionsteam bestehend aus Vertreter:innen folgender Organisationen (in alphabetischer Reihenfolge):

- Arbeiterkammer Wien
- Axians ICT Austria
- FEEI Fachverband der Elektro- und Elektronikindustrie
- FMTI Fachverband Metalltechnische Industrie
- Gewerkschaft GPA
- Industriellenvereinigung
- KNOW-CENTER Graz
- Plattform Industrie 4.0 Österreich
- Produktionsgewerkschaft PRO-GE
- Software Competence Center Hagenberg
- TietoEVRY

Über die Autorin

Nikolina Grgic

ist Senior Referentin der Plattform Industrie 4.0 und betreut die Expert:innengruppen Mensch in der digitalen Fabrik, Security und Safety und Normen und Standards. Zudem leitet Nikolina Grgic weitere Projekte zu menschenzentrierter KI und KI-Themen in der Standardisierung.



SensiTrack – Privacy By Design Konzept zur Gestaltung eines Tracking- und Tracing- Anwendungsfalls im Arbeitsumfeld

1	Einleitung	111
2	Theorie	113
2.1	Tracking- und Tracing-Systeme (TATS) in der Arbeitswelt	113
2.2	Mögliche Folgen beim Einsatz von TATS für die Mitarbeiter:innen	115
3	Methode	115
3.1	Gründe für die Implementierung von Datenschutzfunktionen	116
3.2	Durchführung von Expert:innen-Interviews	117
4	Mitarbeiter:innenzentrierte Gestaltung von Tracking- und Tracing-Anwendungen in Organisationen	120
4.1	Fallbeispiele für den Einsatz von Tracking-Technologien im Arbeitsumfeld	121
4.1.1	Risikoanalyse	121
4.1.2	Datenschutz sicherstellen	122
4.1.2.1	Datenorientierte Strategien:	123
4.1.2.2	Prozessorientierte Strategien	125
4.1.3	Richtlinien zur mitarbeiter:innenzentrierten Gestaltung von TATS	127
5	Fazit	131
6	Acknowledgement	131
	Über den Autor	132
	Quellenverzeichnis	133

1 Einleitung

Organisationen können durch den Einsatz digitaler Technologien in vielfacher Weise, etwa durch schnellere und effizientere Arbeitsprozesse oder Kosteneinsparungen, profitieren. Jedoch steigt durch die Nutzung neuer Technologien auch die Menge an erfassten Daten im Organisationskontext stetig an. Dadurch entstehen Vorteile für Unternehmen, denn mehr Daten ermöglichen grundsätzlich auch genauere Prognosen oder eine bessere Echtzeitübersicht, beispielsweise über Produktionsergebnisse. Durch die Zunahme an verfügbaren Daten steigt auch der Anteil an personenbezogenen Daten über Mitarbeiter:innen wie z. B. deren Arbeitsleistung oder Aufenthaltsort. Für das Wohl der Arbeitnehmer:innen kann diese Erfassung durchaus von Vorteil sein. Erste Anwendungen zur Analyse von Gesundheitsdaten von Mitarbeiter:innen helfen frühzeitig gesundheitliche Probleme zu erkennen und können dadurch präventive Maßnahmen vorschlagen (Austin 2021). Dadurch lässt sich der Krankenstand von Mitarbeiter:innen verringern, was sowohl für diese als auch der Organisation von Vorteil ist. Weiters kann eine Zunahme an verfügbaren Daten über das Verhalten der Mitarbeiter:innen eine objektivere Bewertung der individuellen Leistung ermöglichen, was zu einem höheren Grad an Fairness durch gerechtere Entlohnung führen kann (Hancock, Hioe und Schaninger 2018).

Durch die Nutzung digitaler Technologien steigt jedoch auch die Transparenz in Bezug auf das Verhalten der Arbeitnehmer:innen und dadurch entstehen auch Auswirkungen auf deren Privatheit. Heutzutage können Unternehmen umfangreiche Daten präzise erfassen und automatisiert auswerten – sei es die Anwesenheit durch Zutrittskontrollen oder Aufenthaltsorte der Mitarbeiter:innen mittels Tracking- und Tracing-Systemen. Während dies für Vorgesetzte mehr Kontrolle über ihre Mitarbeiter:innen und Informationen über deren Verhalten bedeutet, kann dies auf Seiten der Mitarbeiter:innen zu einem Gefühl der permanenten Überwachung und des Leistungsdrucks führen. Folgen davon können ein höheres Stresslevel und ein Verlust der intrinsischen Motivation der Arbeitnehmer:innen sein (Aiello und Kolb 1995). Negative Auswirkungen auf die Arbeitsleistung und somit auch Nachteile für die Organisation können die Folge sein. Darum stellen sich für beide Seiten neue Herausforderungen durch den Einsatz digitaler Technologien, die damit verbundene ständige Erhebung von Daten und den daraus resultierenden Gefahren für die Privatheit der Mitarbeiter:innen.

Die Menge an erfassten Daten im Organisationskontext steigt stetig an.

Der Anteil von personenbezogenen Daten steigt an.

Unternehmen bekommen mehr Kontrolle und Informationen über Mitarbeiter:innen

Die Nutzung von Technologien zur Erhebung von mitarbeiterbezogenen Daten unterliegt einem gesetzlichen Rahmen, der von außen Einfluss auf die technische und individuelle Ebene des Unternehmens nimmt. Gewerkschaften und Vertreter:innen der Industrie vertreten indes ihre Interessen bei der Ausgestaltung des gesetzlichen Rahmens und dessen gerichtliche Interpretation.

Die Verwendung innovativer Tracking- und Tracing Technologien birgt deshalb für Unternehmen reizvolle Potenziale, jedoch auch ernstzunehmende Risiken. Denn erwünschte Effizienzsteigerungen stehen oftmals im Konflikt mit möglichen Verletzungen der Privatheit bzw. unzulässiger und übermäßiger Überwachung von Mitarbeiter:innen. Eine Betrachtung beider Seiten und die Suche nach alternativen Ansätzen ist daher notwendig, um die zunehmende digitale Überwachung am Arbeitsplatz bestmöglich einschätzen zu können. Insbesondere muss untersucht werden, wie diese Daten unter Minimierung von Konflikten in Bezug auf die Privatheit in Organisationen genutzt werden können um neue Vertrauensgrenzen zwischen Arbeitgeber:innen und Arbeitnehmer:innen zu schaffen.

Dieser Beitrag stellt eine Vorgangsweise vor, wie Tracking- und Tracing-Systeme in Organisationen eingesetzt werden können, ohne die Privatheit von Mitarbeiter:innen einzuschränken. Dafür wird einerseits auf bestehende Datenschutzstrategien eingegangen, sowie eigene Richtlinien zur schnellen Bewertung dieser Systeme vorgestellt. Diese Strategien und Richtlinien werden an einem Fallbeispiel im Gesundheitsbereich praktisch angewandt.

Verwendung von Tracking- und Tracing-Technologien birgt für Unternehmen reizvolle Potenziale, aber auch Herausforderungen – erwünschte Effizienzsteigerungen stehen oftmals im Konflikt mit möglichen Verletzungen der Privatheit.

2 Theorie

Um dieses interdisziplinäre Forschungsfeld zu erschließen, ist es wichtig, Wissen über die gesamte Bandbreite von Tracking- und Tracing-Systemen (TATS) und Einflussfaktoren auf das Mitarbeiterverhalten zu gewinnen. Einerseits ist es notwendig, ein gutes Verständnis der Einflussfaktoren auf das Mitarbeiterverhalten in Organisationen zu haben, andererseits ist es wichtig, die Probleme der verwendeten Technologien zu erkennen, da diese oft zu unerwarteten Datenschutzproblemen führen können. Daher werden in Kapitel 2.1 drahtlose Sensor-Netzwerke beschrieben, welche die Grundlage für Tracking- und Tracing-Anwendungen sind.

2.1 Tracking- und Tracing-Systeme (TATS) in der Arbeitswelt

Durch die stetig zunehmende Automatisierung in Organisationen z. B. produzierende Industrie, Gesundheitsbereich oder Logistik wächst der Bedarf an Positions-Informationen über alle Teilnehmer der Wertschöpfungskette (Mitarbeiter:innen, Fahrzeuge, Rohmaterialien, Transportbehälter, Werkzeuge uvm.) am Standort (Zafari, Gkelias und Leung 2019). Die Verfolgung und Nachverfolgung von physischen Objekten wird als Asset Tracking bezeichnet und erfasst alle Aktivitäten und Methoden zur Erfassung und Nutzung von Echtzeit-Positionen und Statusdaten von Objekten, also z. B. Waren, Werkzeuge, Ladungsträger, Rohmaterialien und ist in vielen Branchen bereits Standard (Oztekin u. a. 2010). Aktuelle Asset Tracking-Systeme nutzen sogenannte drahtlose Sensornetzwerke, um die Position eines Objektes ermitteln zu können. Diese Systeme basieren auf standardisierten und auch im Alltag gebräuchlichen Funktechnologien wie z. B. Wifi, Bluetooth, RFID oder ZigBee (Zafari u. a. 2019). Jedes nachzuverfolgende Objekt muss zu diesem Zweck mit einem sogenannten Tag, also einem Funksender, ausgestattet werden. Diese Tags sind aktive oder passive Funksender, die in Intervallen ihre Identifikationsnummer an ein geeignetes System senden. In dem Bereich, in dem die Aufzeichnung von Daten stattfindet, werden Lesegeräte statisch positioniert, um eine optimale Erkennung der Funksignale der Tags zu gewährleisten. In Abbildung 1 wird eine Architektur eines TATS schematisch dargestellt. Die Lesegeräte sammeln die Identifikationsnummer eines Senders und die dazugehörige empfangene Signalstärke, wenn sich dieser in Reichweite befindet. Mit Hilfe der Triangulation können die Werte mehrerer Lesegeräte verwendet werden, um die physische Posi-

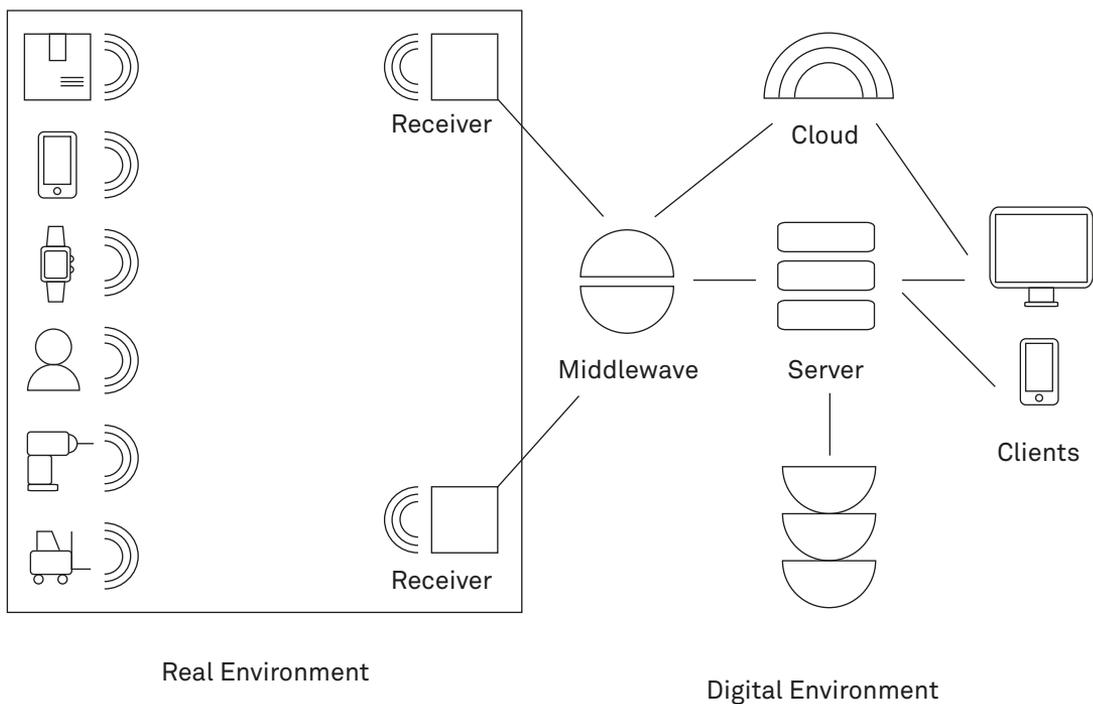
Die Verfolgung und Nachverfolgung von physischen Objekten wird als Asset Tracking bezeichnet.

Diese Systeme basieren auf standardisierten und auch im Alltag gebräuchlichen Funktechnologien.

Die Lesegeräte sammeln die Identifikationsnummer eines Senders und die dazugehörige empfangene Signalstärke, wenn sich dieser in Reichweite befindet.

tion eines Tags zu berechnen (Zafari u. a. 2019). Neben der Objekt-Verfolgung bieten diese Technologien auch neue Möglichkeiten, ein Bewegungsprofil einzelner Mitarbeiter:innen einschließlich der genauen Zeit und Dauer des Aufenthalts aufzuzeichnen. Dafür müssen die Mitarbeiter:innen, ebenso wie bei Asset Tracking-Systemen, nur einen Funksender bei sich tragen.

Abbildung 1:
Architektur eines
Track- und Tracing-
Systems



2.2 Mögliche Folgen beim Einsatz von TATS für die Mitarbeiter:innen

Wie bereits im vorigen Absatz beschrieben, ist es mit TATS möglich, auch Standorte und Positionsverläufe der Mitarbeiter:innen aufzuzeichnen. Diese Systeme können aus Organisationssicht unter anderem dazu beitragen, das Fehlverhalten von Mitarbeiter:innen zu reduzieren, die Effizienz zu erhöhen oder die rasche Auffindbarkeit von Personen zu ermöglichen (Rahmat u. a. 2017). Weiters ist auch eine Steigerung der Arbeitssicherheit beim Einsatz in gefährlichen Umgebungen ein klassischer Einsatzzweck von Personen-Tracking-Systemen (Minch 2005).

Allerdings kann es auch zu einem Verlust der Moral der Mitarbeiter:innen, durch einen Eingriff in die Privatheit, führen. In der Literatur finden sich zahlreiche Belege, dass auch negative Reaktionen aufgrund von Überwachung entstehen können, wie z. B.: A) Gefühl der Verletzung der Privatheit (McNall und Roch 2007), B) Wahrnehmungen von Unfairness (Moorman und Wells 2003), C) Verminderte Arbeitszufriedenheit und organisatorisches Engagement (Wells, Moorman, und Werner 2007), D) Geringere Aufgabenleistung und Produktivität (Aiello und Kolb 1995), E) Größere Wahrnehmungen von Stress bei arbeitsbezogenen Aufgaben (Aiello und Kolb 1995).

Tracking- und Tracing Technologien können Fehlverhalten von Mitarbeiter:innen zu reduzieren, die Effizienz zu erhöhen oder die rasche Auffindbarkeit von Personen zu ermöglichen.

Negative Reaktionen aufgrund von Überwachung

3 Methode

Um Richtlinien zur Bewertung über den Einfluss auf die Privatheit von TATS zu entwickeln, wurden verschiedene Vorarbeiten durchgeführt. Im ersten Schritt wurde eine systematische Literaturrecherche durchgeführt, um auszuwerten, welche Gründe Entwickler von TATS angeben, warum ihre Systeme auf Privatheit fokussiert. Weiters sollte die Frage beantwortet werden, welche technische, sowie organisatorischen Maßnahmen sich hier zur Umsetzung eignen. Im nächsten Schritt wurden Interviews mit Vertreter:innen der Wirtschaft und Organisationen durchgeführt, um den aktuellen Stand zum diesem Thema in Niederösterreich abzubilden.

Interviews mit Vertreter:innen der Wirtschaft und Organisationen

3.1 Gründe für die Implementierung von Datenschutzfunktionen

Um den aktuellen Stand der Forschung von TATS in Bezug auf die Privatheit der Nutzer:innen zu ermitteln, wurde eine systematische Literaturrecherche durchgeführt. Diese dient dazu, die durch die bestehende Forschung abgedeckten Bereiche zu identifizieren oder auch Lücken aufzuzeigen. Darüber hinaus nähert sie sich der vorhandenen Literatur aus verschiedenen Perspektiven und erleichtert die Gewinnung von neuen Erkenntnissen. Unter anderem wurde die Fragestellung, welche die häufigsten Gründe für die Implementierung von Datenschutzfunktionen in TATS sind, in dieser Studie untersucht. Folgende fünf Hauptgründe wurden am häufigsten genannt:

- **Informationsverlust oder -veränderung durch potenzielle Angreifer** bezieht sich darauf, dass es nicht möglich sein sollte, persönliche Daten oder Informationen auszulesen oder zu verändern.
- **Einhaltung der Gesetze:** Ein häufig genannter Grund für die Berücksichtigung des Datenschutzes beim Betrieb von TATS ist das Erreichen von Rechtssicherheit. Unternehmen sind angehalten, die personenbezogenen Daten so zu verarbeiten, dass keine Verletzung der Privatheit erfolgen kann.
- **Verarbeitung sensibler Daten:** Besonders sensible Anwendungen sind z. B. der Einsatz von Tracking-Systemen im Gesundheitswesen, wo Patienten oder Mitarbeiterdaten erhoben werden und diese mit Krankenakten verknüpft werden. Darüber hinaus ermöglichen die erfassten Daten auch die elektronische Überwachung der Leistung und des Verhaltens von Mitarbeiter:innen.
- **Weitergabe der Daten an Dritte:** In manchen Fällen kann es notwendig sein, Daten an Dritte weiterzugeben, z. B. im Supply Chain Management oder im Gesundheitswesen. Dies erfordert sichere Schnittstellen, um den datenschutzkonformen Austausch von Daten zu gewährleisten.

- **Kommerziellen Systeme haben keinen ausreichenden Datenschutz:** Kommerzielle Systeme werden in der Regel für ein breites Spektrum von Anwendungsbereichen entwickelt. Viele Anbieter von Tracking-Systemen erwarten von den Systembetreibern, dass sie die Maßnahmen zum Schutz der Privatsphäre selbst umsetzen. In vielen Fällen ist dies jedoch aufgrund der Architektur der Systeme im Nachhinein nur mehr schwer möglich. Außerdem ist das notwendige Wissen für die Umsetzung in Unternehmen oft nicht vorhanden.

Kommerzielle Systeme haben keinen ausreichenden Datenschutz.

Diese Gründe zeigen gut, dass grundsätzlich die Überlegungen existieren, Datenschutz bzw. Funktionen zum Schutz der Privatheit in diese Systeme zu integrieren, jedoch fehlt es oft an einer durchgängigen Strategie bei der Implementierung in Unternehmen. Kaum eine der untersuchten Arbeiten gab mehr als drei dieser Gründe an. Hier könnte man vermuten, dass eher spezielle Lösungen für Probleme gesucht werden, anstatt eine umfassende Implementierung des Datenschutzes anzustreben (Jandl, Wagner u. a. 2021).

3.2 Durchführung von Expert:innen-Interviews

In einem nächsten Schritt wurden semistrukturierte Expert:innen-Interviews durchgeführt, um der Frage nachzugehen, welche Faktoren die Wahrnehmung von Mitarbeiter:innen beim Einsatz von TATS im Arbeitsumfeld beeinflussen (Jandl, Taurer, u. a. 2021). Um negative Auswirkungen zu vermeiden und die positive Einflüsse beim Verwenden dieser Systeme zu fördern, muss man zuerst erheben, welche Wahrnehmungen Mitarbeiter:innen haben.

Daraus lassen sich Richtlinien ableiten, um einerseits die positiven Möglichkeiten dieser Technologie zu unterstützen, aber auch um die Gelegenheiten für unerwünschte Überwachungen bewusst einzuschränken. Dazu wurde eine qualitative Interviewstudie mit 19 Personen durchgeführt. Interviewpartner:innen waren einerseits Mitarbeiter:innen von produzierenden Unternehmen, sowie öffentliche Organisationen wie z. B. Wirtschaftskammer, Gewerkschaft oder die Allgemeine Unfallversicherungsanstalt. Alle Interviewpartner:innen hatten Erfahrung mit dem Einsatz von TATS und konnten auch verschiedene Problemstellungen bei Einführung und Betrieb. Die Interviews wurden einer qualitativen Inhaltsanalyse nach Mayring (2010) unterzogen und daraus konnten folgende sechs Akzeptanzfaktoren erarbeitet werden, die beitragen, um den Einsatz von TATS für Mitarbeiter:innen besser gestalten zu können:

- **Transparenz:** Es ist wichtig, die Bedenken der Mitarbeiter:innen in einem frühen Stadium der Implementierung anzusprechen. Die meisten Interviewpartner gaben an, dass Transparenz bei der Einführung und dem Betrieb von TATS notwendig ist. Es ist wichtig, dass der erwartete Nutzen und die Auswirkungen auf den Prozess im Voraus kommuniziert werden. In den meisten Situationen sind die Mitarbeiter:innen zunächst gegen die Veränderung oder Umstellung von Prozessen. Wenn jedoch klar kommuniziert wird, welche Daten gesammelt werden und zu welchem Zweck, ist die Akzeptanz höher.
- **Vertrauen:** In Bezug auf das Vertrauen und die Nutzung von TATS wurden zwei wichtige Aspekte genannt. Es ist grundlegend wichtig, dass die Organisation als zuverlässig und vertrauenswürdig bekannt ist, dann sind auch Mitarbeiter:innen eher dazu bereit, personenbezogene Daten zur Verfügung zu stellen. Ist eine Organisation eher dafür bekannt, allgemein unerlaubte Aufzeichnungen oder Auswertungen durchzuführen, ist die Akzeptanz neuer Systeme dementsprechend äußerst niedrig. Ein weiterer wichtiger Faktor in Bezug auf Vertrauen zwischen Organisation und Mitarbeiter:innen ist, wie stark der Einfluss von den Betriebsräten auf die Unternehmensführung ist, ob diese nur geduldet werden oder ob ein gutes Verhältnis und Mitspracherecht bei der Einführung neuer Technologien besteht.
- **Kommunikation** – Einer der am häufigsten genannten Faktor war die organisatorische Kommunikation. Die Befragten gaben an, dass es wichtig ist, mit allen betroffenen Mitarbeiter:innen zu sprechen und nicht nur mit deren Vorgesetzten. Außerdem ist ein offenes Betriebsklima, in dem auch unangenehme Themen angesprochen werden können, akzeptanzfördernd.
- **Datenschutz:** Es sollten nur Daten gespeichert werden, die wirklich für die Erfüllung der beabsichtigten Aufgabe definiert wurden. Außerdem sollen TATS nicht dazu verwendet werden, die Leistung einzelner Mitarbeiter:innen zu ermitteln, sondern die Daten sollten anonymisiert und z. B. über eine ganze Abteilung ausgewertet werden. Die Dauer der Überwachung ist ebenfalls entscheidend. Wenn TATS

Der erwartete Nutzen und die Auswirkungen auf den Prozess im Voraus kommuniziert werden.

für einen begrenzten Zeitraum zur Optimierung von Prozessen eingesetzt werden, ist die Akzeptanz höher als bei permanentem Betrieb.

- **Sicherheitswahrnehmung:** Systeme zum Aufzeichnen von Mitarbeiter:innen-Positionsdaten wurde von den Interviewpartner:innen meist abgelehnt, es sei denn, es kann die Arbeitssicherheit dadurch deutlich erhöht werden, z. B. in gefährlichen Arbeitsumgebungen, oder es besteht dadurch auch ein direkter Nutzen für die Mitarbeiter:innen selbst, wie z. B. bessere Koordination oder kürzere Wege. In dieser besonderen Situation fühlen sich die Mitarbeiter:innen dann in der Regel mehr geschützt als überwacht.
- **Gesundheitsaspekt:** Persönliche Gesundheitsdaten sind ein sehr sensibler Bereich. Wenn eine Organisation anbietet Mitarbeiter:innen in diesem Bereich zu unterstützen (z. B. nach Krankheiten), kann dies auch ein größeres Gesundheitsbewusstsein bei den Mitarbeitern selbst auslösen.

Systeme zum Aufzeichnen von Mitarbeiter:innen-Positionsdaten wurde von den Interviewpartner:innen meist abgelehnt.

Aus diesen Vorarbeiten wurden demnach Richtlinien ausgearbeitet, die es ermöglichen, den Einsatz von TATS datenschutzrechtlich unbedenklich und ebenso für Mitarbeiter:innen so zu gestalten, dass eine hohe Akzeptanz von Systemen zu erwarten ist. Die Richtlinien werden in Kapitel 4 dargestellt und anhand eines Fallbeispiels exemplarisch erklärt.

4 Mitarbeiter:innenzentrierte Gestaltung von Tracking- und Tracing-Anwendungen in Organisationen

Um Tracking-Systeme in Organisationen mitarbeiter:innen-zentriert zu gestalten, wird empfohlen:

1. **Risikoanalysen durchzuführen:** Das Ergebnis dieser Analyse ist für die weitere Vorgangsweise ausschlaggebend. Sind mögliche Probleme bezüglich der Privatheit von Mitarbeiter:innen zu erwarten, sollten unbedingt auch die nächsten beiden Punkte durchgeführt werden.
2. **Datenschutz sicherstellen** – durch die Verpflichtung sich bei der Entwicklung an die Privacy By Design-Strategien der Europäischen Union Agentur für Cybersicherheit (Hoepman 2014) zu halten, kann sichergestellt werden, dass personenbezogene Daten nicht widersprüchlich verarbeitet werden.
3. **Richtlinien zur mitarbeiter:innenzentrierten Gestaltung von TATS** : Diese Richtlinien helfen bei der Bewertung des jeweiligen Falles und helfen den Fokus auf den/die Mitarbeiter:in in der Gestaltung zu lenken.

4.1 Fallbeispiele für den Einsatz von Tracking-Technologien im Arbeitsumfeld

In diesem Abschnitt wird anhand eines Fallbeispiels die Vorgehensweise exemplarisch beschrieben, zuerst allgemein und dann direkt auf das Fallbeispiel bezogen.

Fallbeispiel: Ortungssystem im Krankenhaus

Ein Krankenhaus wird ein Tracking- und Tracing-System zur Ortung von Mitarbeiter:innen installieren. Dabei soll bei jedem/jeder Mitarbeiter:in ein RFID-Sender am Namensschild befestigt werden, welcher Informationen an eine zentrale Recheneinheit sendet. Es soll Krankenhauspersonal dadurch schneller ausfindig gemacht werden können, mit dem Ziel, eine schnellere Kommunikation der Patientenbedürfnisse an das entsprechende Personal zu ermöglichen. Weiters soll das System der Dokumentation der Pflege dienen. Die Administratoren dieses Systems erhalten dadurch Auskunft, wo sich das Personal aufhält, wie viel Zeit an jedem Ort verbracht wird und mit wem man gerade im selben Raum ist.

4.1.1 Risikoanalyse

Um die möglichen Bedrohungen der Mitarbeiter:innen-Privatheit zu bestimmen, wird im ersten Schritt eine Privatheit-Risikoanalyse durchgeführt (Hong u. a. 2004). Anhand dieser Analyse sind folgende Risiken bezüglich der Privatheit für Mitarbeiter:innen im beschriebenen Fallbeispiel vorhanden:

- Übermäßige Überwachung durch Mitarbeiter:in und Vorgesetzte
- Auffinden durch eine böswillige Person mit Systemzugang z. B. Stalker
- Übermäßige Analyse der ermittelten Daten
- Risikobewertung

Um die gegebenen Risiken bewertbar zu machen, wird in weiterer Folge eine Risikobewertung durchgeführt. Dazu werden die gegebenen Risiken in Bezug auf Eintrittswahrscheinlichkeit und Höhe der Auswirkung bewertet. Das Risiko der übermäßigen Überwachung der Vorgesetzten wurde mit einer hohen Eintrittswahrscheinlichkeit und einem ebenfalls hohen Schaden für den/die Mitarbeiter:in bewertet, da dies direkte Auswirkung auf den Arbeitsplatz haben könnte z. B. Verhältnis zum Vorgesetzten, gegenseitiges Ausspionieren und dadurch Verschlechterung der Mitarbeiter:innen-Zufrie-

denheit. Das Risiko durch die Auffindbarkeit für böswillige Personen wird hier mit einer geringen Eintrittswahrscheinlichkeit, aber einem sehr hohen Schaden für die betreffende Person und die Organisation bewertet.

Die übermäßige Analyse von personenbezogenen Daten wird mit einer mittleren Eintrittswahrscheinlichkeit und einem hohen Schaden bewertet. Hier könnte es passieren, dass der berufliche Aufstieg verwehrt bleibt, weil fehlerhafte Aufzeichnungen dazu führen, dass man insgesamt schlechter bewertet wird. Außerdem ist hier seitens des Krankenhauses ein hoher Reputationsverlust wahrscheinlich, wenn solche Methoden bekannt werden würden. Aufgrund der Risikobewertung, bei zwei der drei gefundenen Gefahren hinsichtlich der Privatsphäre der Nutzer:in entweder mit hoher Eintrittswahrscheinlichkeit oder hohem Schaden eingeschätzt wurde, besteht hier Handlungsbedarf.

4.1.2 Datenschutz sicherstellen

Dazu dienen die acht Strategien der Agentur der Europäischen Union für Cybersicherheit für die Entwicklung im Privacy By Design Ansatz (Hoepman 2014). Strategien geben Handlungsempfehlungen vor, die direkt in Richtlinien übertragen werden können. Die Strategien der ENISA bestehen aus vier daten- und vier prozessorientierten Strategien, wobei datenorientiert bedeutet, dass hier der Fokus bei der Ermittlung, Verarbeitung und Speicherung von Daten liegt, wohingegen prozessorientiert sich auf die Vorgangsweise bezieht.

Acht Strategien der
Agentur der Europäischen Union für
Cybersicherheit

4.1.2.1 Datenorientierte Strategien

Strategie 1: MINIMIZE

Die erste datenorientierte Strategie ist die grundlegendste, aber wertvollste zur Wahrung der Privatheit. Es sollen die verarbeiteten Daten auf den geringstmöglichen Umfang beschränkt werden. Das bedeutet, so wenig personenbezogene Daten wie möglich zu sammeln, um trotzdem die geforderten Aufgaben des Systems zu erfüllen.

Daten auf den geringstmöglichen Umfang beschränken

Anwendung im Fallbeispiel:

Bezugnehmend auf diese Strategie soll beim Fallbeispiel darauf geachtet werden, wirklich nur Daten aufzuzeichnen, die unbedingt für die Erfüllung der gestellten Aufgaben benötigt werden. Folgende Aufgaben soll das System erfüllen:

- *Schnellere Auffindbarkeit von Personal:* Um diese Funktion bereitzustellen, ist es notwendig, dass in der Administration der Name und die ID einer Person hinterlegt werden. Für das Echtzeit-Monitoring ist es jedoch nicht notwendig, sensible Zeit-Weg-Punkte zu speichern. Das bedeutet, dass Datenpunkte, die einer Person eindeutig zuordenbar wären, in regelmäßigen Abständen (jeden Tag) gelöscht werden sollten.
- *Dokumentation der Pflege:* Dafür ist vor allem wichtig, dass eine rasche Kommunikation der Patientenbedürfnisse an das entsprechende Personal erreicht wird. Dazu ist es nicht erforderlich zu wissen, welche Person hier die Pflege durchgeführt hat, das heißt, anstatt den Namen oder ID für diesen Zweck zu speichern, ist es ausreichend, wenn die Position der durchführenden Person (Ärzt:in, Krankenpfleger:in) und die Dauer der Pflege gespeichert wird.

Strategie 2: HIDE

Grundsätzlich beschreibt diese Strategie, dass die Anzeige aller personenbezogenen Daten und deren Zusammenhänge vermieden werden soll. Dies wird oft übersehen, erhöht aber den Datenschutz ungenügend (Hoepman 2014). Wenn es z. B. möglich ist, RFID-Tags oder Netzwerkkennungen wie MAC-Adressen mit einer identifizierbaren Person zu verknüpfen, ist es möglich, diese Person im gesamten

Tracking-Bereich zu verfolgen. Die Kernfunktion der HIDE-Strategie ist es, die Unverbindbarkeit und Unbeobachtbarkeit (Pfitzmann u. a. 2006) der Daten einer Person zu erreichen.

Anwendung im Fallbeispiel:

- *Echtzeit-Analyse:* In der Visualisierung zur Echtzeit-Analyse sollten ebenfalls keine Namen oder eindeutige Kennungen zu Personen angezeigt werden, sondern nur Funktionen oder Rollen. Dadurch lassen sich keine einzelnen Personen mehr direkt überwachen.
- *Datenspeicherung:* Weg-Zeit-Diagramme von Mitarbeiter:innen sind sehr sensibel, da sie die Möglichkeit bieten, die umgebenden Personen und die Aufenthaltsorte einer Person einzusehen. Um die Nachvollziehbarkeit zu reduzieren, sollten Räume, die nicht zur Aufgabenerfüllung wie z. B. Aufenthaltsräume, Toiletten, Gänge als sogenannte gemischte Zonen im System hinterlegt werden. Dadurch kann später ein genauer Positionsverlauf nicht mehr rekonstruiert werden, da diese gemischte Zonen nicht mehr auf einzelne Räume rückführbar sind (Beresford 2003).

Strategie 3: SEPERATE

Durch die Trennung der Verarbeitung und Speicherung verschiedener Quellen personenbezogener Daten, die zu einer Person gehören, ist es nicht mehr so einfach, ein vollständiges Profil dieser Person zu erstellen. Daten aus verschiedenen Quellen sollten auch in verschiedenen Datenbanken gespeichert werden und diese Datenbanken sollten nicht miteinander verknüpft werden. In Sensor-Netzwerken werden Vorverarbeitungen von Sensordaten bereits in den jeweiligen Knoten durchgeführt und nur relevante Ergebnisse in der Datenbank gespeichert. Die verteilte Datenverarbeitung vermeidet zudem eine großflächige zentrale Erfassung. Die verteilte Speicherung personenbezogener Daten begrenzt das Risiko der Verletzung der Privatsphäre durch böswillige Angriffe und unbefugte Zugriffe (Perera u. a. 2020)

Anwendung im Fallbeispiel

Im Fallbeispiel wäre es notwendig sicherzustellen, dass die gesammelten Daten auf einer eigenen Datenbank in der Infrastruktur der

Organisation gespeichert werden. Weiters sollten die Daten anonymisiert abgespeichert werden. Im Sinne der SEPERATE-Strategie wäre es auch wichtig, das Monitoring von Personal nur dort durchzuführen, wo es auch tatsächlich eingesetzt wird. Wird ein gesamtes Gebäude mit einem TATS ausgestattet, sollten Mitarbeiter:innen-Daten nur dort aufgezeichnet werden, wo diese auch tatsächlich eingesetzt werden.

Monitoring von Personal nur dort durchführen, wo es auch tatsächlich eingesetzt wird.

Strategie 4: AGGREGATE

Diese Strategie besagt, dass personenbezogene Daten auf der höchsten Aggregationsstufe und mit dem geringstmöglichen Detaillierungsgrad, in dem sie nutzbar sind, verarbeitet werden sollten. Aggregieren bedeutet in diesem Fall das Gruppieren und Zusammenfassen von Daten zur größten, für die Verarbeitung noch sinnvollen Granularität. Die Aggregation von personenbezogenen Daten in Gruppen mit Attributen schränkt den Detaillierungsgrad der verbleibenden Daten ein. Dies hat den Effekt, dass die Daten in Bezug auf die Privatheit weniger sensibel sind.

Gruppieren und Zusammenfassen von Daten

Anwendung im Fallbeispiel:

Dokumentation: Personenbezogene Daten zu aggregieren bedeutet hier z. B. die Reaktionszeiten auf Patientenanforderung in Kategorien wie z. B., 1-10 Minuten, 11-20 Minuten und länger als 21 Minuten zu speichern. Dadurch kann vermieden werden, dass Auswertungen welches Personal am schnellsten reagiert hat, erstellt werden.

4.1.2.2 Prozessorientierte Strategien

Strategie 5: INFORM

Als erste prozessorientierte Strategie ist INFORM mit dem Begriff der Transparenz verbunden. Wann immer personenbezogene Daten verarbeitet werden, sollten die Betroffenen angemessen informiert werden. Darüber hinaus sollten die Nutzer über ihre Datenzugriffsrechte und darüber, wie sie diese ausüben können, informiert werden (Hoepman 2014). Die Wahl für das Aufzeichnen sollte frei gegeben werden können. Für Tracking-Systeme bedeutet dies, dass Personen überall dort, wo das System Daten erhebt, darüber informiert werden sollten, welche Daten erhoben, verarbeitet und weitergegeben werden. Diese Bereiche sollten mit Warnschildern

Bereiche sollten mit Warnschildern gekennzeichnet werden.

gekennzeichnet werden, wenn diese Bereiche öffentlich sind. Auf diese Weise kann jeder Anwesende frei entscheiden, ob er mit der Verfolgung einverstanden ist.

Anwendung im Fallbeispiel

Im Bezug auf das Fallbeispiel ist hier zu erwähnen, dass das TATS zur Dokumentation und zur Echtzeit-Übersicht im Arbeitskontext eingesetzt wird. Darum ist es hier nicht möglich, bewusst die Aufzeichnungen auszulassen. Jedoch sollten alle Mitarbeiter:innen die Möglichkeiten haben, über sie aufgezeichnete Daten einzusehen.

Strategie 6: CONTROL

Diese Strategie besagt, dass die betroffenen Personen über die Verarbeitung und Speicherung ihrer persönlichen Daten informiert werden sollten. Darüber hinaus sollten die betroffenen Personen über die Verwendung ihrer Daten durch Dritte informiert werden. Zusätzlich ist es wichtig, die Betroffenen über ihre Datenzugriffsrechte und deren Ausübung zu informieren.

Betroffene sollten über ihre Datenzugriffsrechte und deren Ausübung informiert werden.

Anwendung im Fallbeispiel

Ergänzend zur Einsicht über die personenbezogenen Daten der Mitarbeiter:innen sollten diese weiters auch die Möglichkeit haben, Änderungen und Löschung der Daten von den Administrator:innen durchführen zu lassen. Dies ist vor allem wichtig, da es bei TATS auch zu Fehlaufzeichnungen kommen kann. So können durch Reflexionen im Raum oder schlecht eingestellten Empfänger-Antennen eine Person erkannt werden, die sich gerade in einem Nebenraum befindet. Außerdem kann durch das Tragen am Körper der Sender abgeschirmt werden, wodurch eine Aufzeichnung nicht stattfinden würde, was ebenfalls die Dokumentation verfälschen würde.

Strategie 7: ENFORCE

Diese Strategie besagt, dass eine mit den gesetzlichen Anforderungen kompatible Datenschutzrichtlinie vorhanden sein und auch durchgesetzt werden sollte. Konkret bedeutet dies, dass die Erstellung, Pflege und Aufrechterhaltung von Richtlinien und technischen Kontrollen in Bezug auf die Speicherung, Sammlung, Weitergabe, Änderung und Verletzung personenbezogener Daten im Rahmen der vereinbarten Zwecke möglichst umfassend gewährleistet wird.

Anwendung im Fallbeispiel

Diese Strategie kann vor allem für Arbeitnehmer:innen Vertreter:innen in der Sozial-Partnerschaft eingefordert werden. Durch die Durchführung einer Datenschutzfolgeabschätzung und der Umsetzung im Privacy By Design-Ansatz kann dies erreicht werden.

Strategie 8: DEMONSTRATE

Die letzte Strategie verlangt, dass der/die Verantwortliche die Datenschutzrichtlinie und alle anwendbaren rechtlichen Anforderungen einhält und dies gegenüber den betroffenen Personen nachweisen kann. Diese Strategie geht insofern einen Schritt weiter als die ENFORCE-Strategie, als dass der Verantwortliche die Kontrolle über das System nachweisen können muss. Im Fall von Beschwerden oder Problemen sollte der/die Verantwortliche in der Lage sein, das Ausmaß möglicher Datenschutzverletzungen zu bestimmen.

Anwendung im Fallbeispiel

Um die hauptsächlichen Gefahren in Bezug auf die Privatheit bei TATS (unerlaubte Drittnutzung, unerlaubter Zugriff) abzuwenden, ist hier seitens des Systembetreibers sicherzustellen, dass eine genaue Protokollierung über Datenzugriffe und Datenweitergabe an Dritte geführt wird.

4.1.3 Richtlinien zur mitarbeiter:innenzentrierten Gestaltung von TATS

Ist sichergestellt, dass der Datenschutz ausreichend berücksichtigt worden ist, können die folgenden Handlungsempfehlungen im Fallbeispiel angewendet werden. Die Richtlinien sollen helfen, einen möglichst hohen Mehrwert für Mitarbeiter:innen zu generieren. Diese werden in kurzer Form erklärt und danach an das Fallbeispiel angewandt.

Die Richtlinien sollen helfen, einen möglichst hohen Mehrwert für Mitarbeiter:innen zu generieren.

Unterscheidung zwischen den Anwendungsfällen:

Es kommt immer darauf an, welches Ziel bei der Verwendung von TATS im Vordergrund steht. Sollen Objekte wie Rohwaren, Boxen oder Werkzeuge nachverfolgt werden, dann spricht man von Asset Tracking, hier ist es eher unwahrscheinlich, dass es zu Verletzungen von Privatheit kommt. Trotzdem ist es, gerade bei kommerziellen Lösungen, möglich, dass man durch technische Geräte, die man bei sich trägt, wie z. B. Fitnessarmbänder oder Smartwatches oder

ähnlichen passiv Daten an dieses System sendet (Schauer 2018). Um dies zu vermeiden, sollte in diesem Fall Bluetooth und WLAN am Gerät deaktiviert oder ganz ausgeschaltet werden.

Bluetooth und WLAN
am Gerät deaktivieren,
oder ganz ausschalten

Bei TATS, die zum Zweck von Mitarbeiter:innen-Tracking verwendet werden, ist vor allem wichtig zu hinterfragen, welchen Hintergrund für die Einführung dieser Maßnahme besteht. Handelt es sich um Anwendungen zur Erhöhung von Arbeitssicherheit, z. B. bei gefährlichen Arbeitsbedingungen oder zur schnelleren Evakuierung im Alarmfall, dann ist hier grundsätzlich ein Mehrwert für die Mitarbeiter:innen gegeben. Jedoch sollten die Maßnahmen stets ausgewogen sein und dem realen Risiko entsprechen.

Anwendung im Fallbeispiel

Im beschriebenen Fallbeispiel sollen Mitarbeiter:innen schneller auf Patientenanfragen reagieren können und besser auffindbar sein. Wie das System im Notfall zum Vorteil aller eingesetzt werden kann, wird nicht beschrieben. Es würde sich jedoch dazu eignen, im Falle einer Alarmierung schnell eine Übersicht über die sich im Gebäude befindlichen Personen zu bekommen. Dieser Mehrwert rechtfertigt natürlich nicht allein den Einsatz des Systems, jedoch wenn datenschutzrechtlich alle Anforderungen erfüllt werden, ist hier ein Vorteil für Arbeitnehmer:innen argumentierbar.

Privatheit ist ein zwingendes Bedürfnis, keine mögliche Option:

Nicht alles, was man in seinem Privatleben zulässt, ist auch im Arbeitsumfeld akzeptabel. Es ist besonders wichtig zu beachten, dass hier Entscheidungen frei getroffen werden können und auch die Möglichkeit besteht, diese Systeme in bestimmten Fällen nicht zu nutzen. Natürlich sind bei Verwendungszwecken zur Arbeitssicherheit die Vorgaben zu erfüllen, jedoch müssen diese auch einen objektiven Mehrwert erkennen lassen. So ist es durchaus zulässig während der Arbeit in einem sauerstoffarmen Umfeld wie z. B. bei der CO₂-Reinigung, ein System zu nutzen, das den Pulsschlag überprüft. Muss das System jedoch immer verwendet werden und nicht nur in der gefährlichen Arbeitsumgebung, wäre dies abzulehnen. Weiters ist auch darauf zu achten, wie Vorgaben zur Benützung während Pausen gestaltet sind. Hier gibt es keinen offensichtlichen Grund für die Aufzeichnung dieser Daten und wäre deshalb strikt abzulehnen.

Privatheit ist ein
zwingendes Bedürfnis,
keine mögliche Option.

Anwendung im Fallbeispiel

In der Angabe zum Fallbeispiel wird dieses Thema auch nicht näher definiert. Wichtig zu erfahren wäre, wann der Sender getragen werden muss. Jedenfalls sollte in Pausenzeiten und bei körperlicher Hygiene keine Aufzeichnungen erfasst werden. Weiters sollte beim Verlassen der Station (also dem eigentlichen Arbeitsplatz) keine Aufzeichnungen mehr gemacht werden, da hier gezieltes Stalking möglich würde.

Der Stellenwert von Privatheit in der Organisation:

Fast jede:r Mitarbeiter:in erwartet einen gesunden Arbeitsplatz mit einem hohen Maß an Arbeitssicherheit. Auch Unternehmen haben ein Interesse an gesunden und zufriedenen Arbeitnehmer:innen. Der Einsatz von TATS hat Potenzial, die Arbeitszufriedenheit zu senken und sollte deshalb sehr sensibel behandelt werden. Weiters sollte es für Mitarbeiter:innen auch möglich sein, das Tracking vorübergehend zu deaktivieren oder in bestimmten Fällen gar nicht zu nutzen. Besteht hier keine Möglichkeit, wird es zeitnah zur Umgehung des Trackings durch die Mitarbeiter:innen kommen, wie z. B. durch bewusstes Ablegen der Sender. Kann ein Unternehmen deutlich machen, dass auch auf die Privatheit der Mitarbeiter:innen geachtet wird und Maßnahmen dafür geschaffen werden, ist dies ebenfalls positiv zu beurteilen.

Unternehmen haben ein Interesse an gesunden und zufriedenen Arbeitnehmer:innen.

Anwendung im Fallbeispiel

Im Fallbeispiel fehlt die Definition während welcher Tätigkeiten man das System nutzen sollte. So kann es während einer Schicht durchaus möglich sein, dass man zwei Stunden der Arbeitszeit für das Dokumentieren der Pflege aufwendet. In dieser Zeit sollte es möglich sein, das System zu deaktivieren, um in Ruhe arbeiten zu können.

Diskussion auch über unbequeme Themen:

Es ist leicht, über die Vorteile von TATS für die Organisation zu reden. Die Organisation sollte jedoch auch offen für die Befürchtungen von Mitarbeiter:innen sein und diese im Vorfeld aktiv ansprechen. Gerade für ältere Personen sind Technologien oft schwerer zu verstehen und dadurch können Befürchtungen größere Auswirkungen haben. Bieten Unternehmen vorab Informationsveranstaltungen zum Einsatz neuer Technologien an, ist dies ebenfalls positiv

Gerade für ältere Personen sind Technologien oft schwerer zu verstehen.

zu bewerten. Dabei sollte Datenschutz genauso Thema sein wie die möglichen Auswirkungen auf die Privatheit der Mitarbeiter:innen. Auch das Herausarbeiten des Mehrwerts für die Mitarbeiter:innen fällt unter diesem Punkt. So sollten durch den Einsatz von TATS nicht nur Vorteile für das Unternehmen entstehen, sondern auch die Arbeitsbedingungen der einzelnen Mitarbeiter:innen verbessern. Das kann z. B. durch das Automatisieren analoger Arbeiten, Zeitersparnis durch optimierte Wege oder zusätzlichen Informationen und dadurch ein besseres Gefühl der Kontrolle beinhalten.

Anwendung im Fallbeispiel

Im Fallbeispiel sollten ab der ersten Entscheidung, diese Technologie einzuführen Arbeitnehmer:innen-Vertreter:innen in die Planung miteingebunden werden. Weiters sollte man stets die Möglichkeit bereitstellen, wo Arbeitnehmer:innen anonym Feedback geben können. Der Mehrwert für den/die einzelne/n Mitarbeiter:in wäre gegeben, wenn die Dokumentation der Pflege automatisiert wird und sich die Mitarbeiter:innen dadurch Zeit sparen.

Mitarbeiter:innenwahrnehmung ernst nehmen:

Die erarbeiteten Akzeptanzfaktoren aus Kapitel 3.2 sollten ebenfalls in der Vorgangsweise des Unternehmens mitbedacht werden. Werden diese Faktoren in konkrete Maßnahmen umgesetzt, wird auch hier ersichtlich, dass das Unternehmen eine hohe Akzeptanz des Systems erreichen will.

Mitarbeiterwahrnehmung ernst nehmen

- **Transparenz:** Welche Daten werden den Arbeitnehmer:innen zur Verfügung gestellt? Sind Arbeitnehmer:innen-Vertreter:innen in die Planung involviert?
- **Vertrauen:** Ist das Unternehmen eher als vertrauenswürdig bekannt oder gab es schon Datenschutzverletzungen oder ähnliche Vorfälle?
- **Kommunikation:** Wie wird das Projekt intern kommuniziert? Wird offen über die Herausforderung der Privatheit gesprochen?

Speicherung der Daten:

Wird festgelegt, wie lange die Daten gespeichert werden, oder über welchen Zeitraum Daten erfasst werden? Wo werden Daten gespeichert und wer hat alles Zugriff?

- **Sicherheitsempfinden:** Gibt es durch den Einsatz des Systems merkliche Verbesserung in der Arbeitssicherheit? Ist die Argumentation in Bezug auf Arbeitssicherheit im Arbeitsumfeld gerechtfertigt?
- **Gesundheitsaspekt:** Wie wird die Gesundheit der Arbeitnehmer:innen durch das System beeinflusst? Werden sensible Gesundheitsdaten von Mitarbeiter:innen erfasst und verarbeitet und in welcher Weise profitieren davon die Mitarbeiter:innen selbst?

Diese und ähnliche Fragestellungen eignen sich gut, um die Intention von Organisationen beim Implementieren von TATS zu beurteilen. Fallen die Antworten für den/die Mitarbeiter:in durchwegs positiv aus, ist auch zu erwarten, dass Auswirkungen auf die Privatheit der Mitarbeiter:innen eher gering bleiben.

5 Fazit

In diesem Artikel wurde eine Vorgangsweise zur Wahrung der Mitarbeiter:innen-Privatheit bei der Einführung und Implementierung von Tracking- und Tracing-Systemen vorgestellt. Dazu ist im ersten Schritt eine Risikoanalyse bzw. -bewertung durchzuführen. Besteht die Vermutung, dass es Einfluss auf die Privatheit der Mitarbeiter:innen geben kann, wie z. B. durch unbefugte Auswertung durch Dritte, wird empfohlen, den Privacy By Design Ansatz in der Implementierung zu wählen. Anhand der acht Strategien lassen sich Handlungsempfehlungen für die Umsetzung ableiten. Im nächsten Schritt sollte versucht werden, das System möglichst mitarbeiter:innenzentriert zu gestalten, das heißt, dass auf die Bedürfnisse der Nutzer:innen nicht nur Rücksicht genommen wird, sondern diese gezielt in den Entwicklungsprozess miteingebunden werden.

6 Acknowledgement

Diese Arbeit wurde von der Niederösterreichischen Arbeiterkammer „Projektfonds Arbeit 4.0“ im Projekt „Sensitrack – Tracking versus Privacy in der Arbeitswelt 4.0“ unterstützt.

Über den Autor

Dipl.-Ing. Christian Jandl

Jg. 1979, Researcher am Institut für Creative\Media/Technologies;
Dissertant an der TU Wien am Institut für Managementwissenschaften;
Leitung der Forschungsgruppe „Digitale Technologien“ am IC\M/T
St. Pölten Arbeitsschwerpunkte:
Privatsphäre in der digitalen Arbeitswelt, Industrie 4.0.

Quellenverzeichnis

- Aiello, John R., und K. J. Kolb. 1995. „Electronic performance monitoring and social context: Impact on productivity and stress“. *Journal of Applied Psychology* 80:339–53
- Austin, Daniela. 2021. „Delivering Digital Health: The Barriers and Facilitators to University-Industry Collaboration, Elsevier Enhanced Reader
- Beresford, A.; Stajano, F. Location privacy in pervasive computing. *IEEE Pervasive Comput.* 2003, 2, 46–5
- Hancock, Bryan, Elizabeth Hioe, und Bill Schaninger. 2018. „The Fairness Factor in Performance Management“
- Hoepman, Jaap-Henk. 2014. „Privacy Design Strategies“. S. 446–59 in *ICT Systems Security and Privacy Protection*. Bd. 428, herausgegeben von N. Cuppens-Bouahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, und T. Sans. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hong, Jason I., Jennifer D. Ng, Scott Lederer, und James A. Landay. 2004. „Privacy Risk Models for Designing Privacy-Sensitive Ubiquitous Computing Systems“. S. 91 in *Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and techniques - DIS '04*. Cambridge, MA, USA: ACM Press.
- Jandl, Christian, Florian Taurer, Martina Hartner-Tiefenthaler, Markus Wagner, Thomas Moser, und Sebastian Schlund. 2021. „Perceptions of Using Tracking and Tracing Systems in Work Environments“. S. 384–98 in *HCI in Business, Government and Organizations*. Bd. 12783, *Lecture Notes in Computer Science*
- Jandl, Christian, Markus Wagner, Thomas Moser, und Sebastian Schlund. 2021. „Reasons and Strategies for Privacy Features in Tracking and Tracing Systems—A Systematic Literature Review“. *Sensors* 21(13):4501
- Mayring, Philipp. 2010. „Qualitative Inhaltsanalyse. Grundlagen und Techniken“. *Weinheim* 3:58.
- McNall, Laurel A., und Sylvia G. Roch. 2007. „Effects of Electronic Monitoring Types on Perceptions of Procedural Justice, Interpersonal Justice, and Privacy“. *Journal of Applied Social Psychology* 37(3)
- Minch, Robert. 2005. „Legal and Ethical Implications of Employee Location Monitoring“. in *Legal and Ethical Implications of Employee Location Monitoring*.
- Moorman, Robert H., und Deborah L. Wells. 2003. „Can Electronic Performance Monitoring Be Fair? Exploring Relationships Among Monitoring Characteristics, Perceived Fairness, and Job Performance“. *Journal of Leadership & Organizational Studies* 10(2):2–16

Oztekin, Asil, Foad M. Pajouh, Dursun Delen, und Leva K. Swim. 2010. „An RFID Network Design Methodology for Asset Tracking in Healthcare“. *Decision Support Systems* 49(1):100–109.

Perera, Charith, Mahmoud Barhamgi, Arosha K. Bandara, Muhammad Ajmal, Blaine Price, und Bashar Nuseibeh. 2020. „Designing privacy-aware internet of things applications“. *Information Sciences* 512:238–57.

Pfitzmann, Andreas, TU Dresden, Marit Hansen, und ULD Kiel. o. J. „Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology“. 54.

Rahmat, Romi Fadillah, Eka Tama Herly, Baihaqi Siregar, Mohammad Fadly Syahputra, und Opim Salim Sitompul. 2017. „RFID presence monitoring system as an input to measure the workload of employee“. S. 1–6 in *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*.

Schauer, Lorenz. 2018. „Analyzing the Digital Society by Tracking Mobile Customer Devices“. S. 467–78 in *Digital Marketplaces Unleashed*, herausgegeben von C. Linnhoff-Popien, R. Schneider, und M. Zaddach. Berlin, Heidelberg: Springer Berlin Heidelberg.

Wells, Deborah L., Robert H. Moorman, und Jon M. Werner. 2007. „The Impact of the Perceived Purpose of Electronic Performance Monitoring on an Array of Attitudinal Variables“. *Human Resource Development Quarterly* 18(1):121–38.

Zafari, Faheem, Athanasios Gkelias, und Kin K. Leung. 2019. „A Survey of Indoor Localization Systems and Technologies“. *IEEE Communications Surveys & Tutorials*

Impressum

Kontakt

AK Vorarlberg
Schaffarei – Haus für Arbeitskultur
Widnau 10
6800 Feldkirch
+43 (0) 50 258-1625
forschung@schaffarei.at
schaffarei.at

Impressum

Herausgeberin: AK Vorarlberg
Redaktion: Petra Rumpl, Dominic Götz, AK Vorarlberg
Lektorat: Klaudia Häusle, AK Vorarlberg
Grafik: Zeughaus, Feldkirch
Druck: Thurnher Druckerei GmbH, Rankweil

GEFÖRDERT DURCH

Digifonds



Forschung 1

Technikfolgenabschätzung aus
Arbeitnehmer:innenperspektive
1. Auflage: 400 Stück
Eigenverlag der AK Vorarlberg, Feldkirch 2022

Ein Projekt der
Arbeiterkammer Vorarlberg

