

Künstliche Intelligenz (KI) für Gute Arbeit

März 2020

1. Das Ziel: Umsetzung von „Gute Arbeit by design“

Der **Deutsche Gewerkschaftsbund** (DGB) hat im Januar 2019 ein erstes **Impulspapier** „Künstliche Intelligenz und die Arbeit von morgen“ veröffentlicht, mit dem ein Modell zur betrieblichen Gestaltung von KI-Anwendungen skizziert wird. Mit „Gute Arbeit by design“ wird ein vorausschauender Ansatz zur Arbeitsgestaltung im Kontext von autonomen Softwaresystemen vorgeschlagen: „Die Gestaltung muss bereits bei der Konzeptionierung und in der Entwicklungsphase von KI-Systemen beginnen. Voraussetzung für eine gute Gestaltung ist ein breiter Beteiligungsprozess, der bereits bei der Definition der Zielsetzung für die KI und deren Anwendung beginnt und eine Folgenabschätzung einschließt“.

Dieser Ansatz hat eine breite Resonanz gefunden und wird zum Beispiel von der „Plattform Lernende Systeme“: KI-Plattform der Bundesregierung oder dem „Ethikbeirat HR-Tech“ grundsätzlich geteilt. Auch im gesellschaftlichen Kontext (zum Beispiel Bertelsmann-Stiftung „Algorules“: Regeln für die Gestaltung algorithmischer Systeme) oder als Ergebnisse von Arbeitsforschungsprojekten (Umsetzungshilfen Arbeit 4.0) zeigen sich ähnliche Ansätze einer prozessorientierten präventiven Gestaltung von KI.

Mit diesem **Konzeptpapier** soll der DGB-Ansatz „Gute Arbeit by design“ konkretisiert werden. Der Blick soll auf **Fragestellungen und Sollbruchstellen** für die Gestaltung von Veränderungsprozessen (Change-Prozesse) im Kontext der KI-Nutzung gerichtet werden. In Anlehnung an die „Bewertungsliste“ für Leitlinien der HLEG für künstliche Intelligenz der EU-Kommission, die für den allgemeinen Einsatz von KI entwickelt wurde, soll der Fokus an dieser Stelle auf die Umsetzung und deren Rahmenbedingungen gelegt werden.

Das Konzeptpapier soll damit einen **Orientierungsrahmen** für die gewerkschaftliche, politische und betriebliche Diskussion darstellen. Gleichzeitig geht es auch darum, die **rechtlichen Lücken in den Rahmenbedingungen für innovative Aushandlungsprozesse** zu schließen.

2. Der politische Debattenstand

Die Debatte um „Künstliche Intelligenz“ (KI) für die Arbeitswelt hat in den letzten zwei Jahren rasant an Fahrt aufgenommen. Auch wenn im öffentlichen Diskurs „KI“ oft nur den Begriff „Digitalisierung“ ersetzt, werden KI-Anwendungen – also lernenden Software- und Entscheidungssystemen – tatsächlich große Potenziale zugeschrieben. Dies gilt sowohl für die Automatisierung als auch für intelligente Assistenz menschlicher Arbeit. Der internationale Konkurrenzdruck für die Entwicklung und Anwendung von KI in Wirtschaft und Arbeitsmarkt hat spürbar zugenommen. Dennoch stehen viele Unternehmen – auch in Deutschland – noch am Anfang, KI-basierte Möglichkeiten zu nutzen. Ein Grund ist eine hohe Unsicherheit im Umgang mit KI-Anwendungen. Gleichzeitig bestehen in der Gesellschaft, nicht zuletzt bei Arbeitnehmerinnen und Arbeitnehmern, Sorgen, die von der Angst um Arbeitsplatzverluste über Kontrollverlust bis zur „Entmündigung“ durch KI-Systeme reichen.

Die politischen Ansätze sind stark auf **Vertrauen und Akzeptanz** für KI ausgerichtet, bleiben aber bislang allgemein. Auf internationaler wie auch auf europäischer Ebene steht die Zielsetzung einer „vertrauenswürdigen KI“ im Mittelpunkt. So hat die **OECD** im Mai 2019 Prinzipien vereinbart, um „Menschenrechte und demokratische Werte zu respektieren“. Der Fokus liegt hier auf Grundwerten wie gesellschaftlichem Wohlergehen, Nachhaltigkeit, Fairness, Sicherheit oder Rechenschaftspflichten. KI-Nutzung soll darüber hinaus auch auf die stärkere Einbindung unterrepräsentierter Bevölkerungsgruppen und die Verringerung der wirtschaftlichen, sozialen, geschlechtsspezifischen und sonstigen Ungleichheiten abzielen. Dahinter steht ein Mensch-zentriertes Design für „intelligente“ bzw. selbst lernende Maschinen. Es ist also ein ethischer Ansatz im Umgang mit KI-Systemen, der jedoch sehr grundsätzlich formuliert ist und rechtliche, ökonomische und auch kulturelle Rahmenbedingungen vernachlässigt. Damit werden insbesondere Fragen hinsichtlich der Veränderung von Wertschöpfung, Wirtschaftsstrukturen, Arbeitsmärkten und Beschäftigung nicht thematisiert.

Im Juni 2019 hat die „Unabhängige **Hochrangige Expertengruppe für Künstliche Intelligenz der EU-Kommission** (HLEG) „Ethik-Leitlinien für eine vertrauenswürdige KI“ zur Diskussion gestellt. Auch hier stehen grundsätzliche Aspekte im Vordergrund: KI soll „rechtmäßig, ethisch und robust“ sein. Als wesentliche Kriterien werden Schadensverhütung, Fairness und Erklärbarkeit genannt, ebenso Schutz der Privatsphäre, Transparenz, Nicht-Diskriminierung und Rechenschaftspflicht. So soll die menschliche Autonomie „geachtet“ bleiben. Nicht

zuletzt wird das „gesellschaftliche wie ökologische Wohlergehen“ als Ziel betont. Immerhin werden hier „Situationen, die sich durch ungleiche Macht- oder Informationsverteilung auszeichnen, etwa zwischen Arbeitgebern und Arbeitnehmern“, als besondere Herausforderung aufgegriffen.

Die **Bundesregierung** hat im November 2018 eine „Strategie Künstliche Intelligenz“ beschlossen, bei der vor allem der Plan im Vordergrund steht, „KI Made in Germany“ zu fördern. Auch hier wird ein allgemeines Ziel für eine „verantwortungsvolle und gemeinwohlorientierte Entwicklung und Nutzung von KI“ ausgegeben. Solche grundsätzlichen Ansätze finden sich auch in Erklärungen von Unternehmen oder Zusammenschlüssen wie „Partnership on AI“ oder anderen.

Die „**Datenethikkommission**“ der Bundesregierung (DEK) hat im Oktober 2019 ähnliche Prinzipien veröffentlicht, die sie allerdings mit Anforderungen im Umgang mit algorithmischen Entscheidungssystemen verknüpft. Dabei ist ein abgestuftes Bewertungsmodell entwickelt worden, das nach Schwere und Wahrscheinlichkeit von Schädigungspotenzialen unterscheidet („Kritikalitätspyramide“). Nach einer Risikofolgenabschätzung werden je nach Einstufung Transparenz- und Offenlegungspflichten, Auskunftsrechte für Betroffene, Zulassungsverfahren und auch Verbote für algorithmische Systeme gefordert. Eine Betrachtung der besonderen Bedingungen in der Arbeitswelt ist von der DEK jedoch nicht vorgenommen worden.

Die **EU-Kommission** hat im aktuell (Februar 2020) ein “White Paper On Artificial Intelligence - A European approach to excellence and trust” veröffentlicht, mit dem erste Regulierungsoptionen hinsichtlich ‚riskanter‘ Einsatzgebiete von KI-Anwendungen auch im Kontext der Arbeitswelt adressiert wurden. Als Beispiele für die Einstufung von Anwendungen mit hohem Risiko werden automatisierte Bewerbungsverfahren oder „Situationen, die sich auf die Rechte von Arbeitnehmern auswirken“, explizit genannt. Ob und welche Regulierungsansätze auf europäischer Ebene umgesetzt werden sollen, wird sich nach der Konsultationsphase ab Juni 2020 zeigen.

Insgesamt zeigt sich ein **breiter politischer Konsens auf einem grundsätzlichen und abstrakten Niveau**. Die Bedeutung von KI für die Zukunft der Arbeitsbeziehungen und deren Rahmenbedingungen wird dabei bislang zu wenig diskutiert. Auch finden sich bislang keine konkreten politischen Regelungsvorschläge für die Arbeitswelt.

3. Ein Orientierungsrahmen für die Einführung von KI im Betrieb

Der Einsatz von KI-Systemen in der Arbeitswelt wirft neue Fragen hinsichtlich der KI-Logiken, des Umgangs mit Daten sowie der Arbeitsgestaltung auf und erfordert deshalb verbindliche Prozesse zur rechtzeitigen Mitbestimmung und Beteiligung der Beschäftigten und ihrer Interessenvertretungen. Es gilt also, die kollektive menschliche Intelligenz zu nutzen, um künstliche Intelligenz in der betrieblichen Arena für eine effiziente und produktive, gesundheits- und lernförderliche sowie gendersensible Arbeitsgestaltung einzusetzen. Dafür sollten auch Beratungs- und Unterstützungsangebote für Unternehmen, insbesondere KMU, sowie für Mitbestimmungsakteure und Beschäftigte öffentlich gefördert werden (vgl. KI-Strategie der Bundesregierung 2018).

In transparenten und beteiligungsorientierten betrieblichen Prozessen sollten die Potenziale der KI-Anwendungen für eine humane Arbeitsgestaltung ausgeleuchtet werden. Alle relevanten Akteure für diese Prozesse sollten bereits vor der Einführung von KI-Systemen im Betrieb zusammenarbeiten, um die Optimierungsziele für den betrieblichen Einsatz gemeinsam zu vereinbaren und signifikante Auswirkungen für die Arbeitsgestaltung frühzeitig zu antizipieren. Dies betrifft insbesondere KI-Anwendungen, bei der Daten der Beschäftigten verarbeitet werden oder dies sogar voraussetzen (Beispiele: Personalisierte Assistenz, Recruiting, Mensch-Roboter-Kollaboration et al.) sowie KI-Anwendungen, die einen Kompetenz- und Autonomieverlust der Beschäftigten – oder gar Arbeitsplatzverluste – zur Folge haben können. Die Anwenderorientierung (Gebrauchstauglichkeit / Usability ISO 9241-110) für die Beschäftigten sollte dabei eine hohe Priorität haben und gendersensibel sein.

Auch bei der Auswahl und dem Einkauf externer Systeme muss die aktive Mitarbeit der Beschäftigten und ihrer Interessenvertretungen bei der Formulierung der Anforderungen an unterschiedliche KI-Systeme gewährleistet werden.

Die gemeinsame Verständigung auf die Zielsetzungen ist bei KI-Systemen von besonderer Relevanz, um die nötige Akzeptanz für den betrieblichen Einsatz zu erreichen – auch in Betrieben ohne Betriebsrat. Dies gilt sowohl für Automatisierungsprozesse (Algorithmic oder Automated Decision Making; ADM/AuDM) als auch für Assistenzsysteme (Decision Support Systems; DSS).

Privacy und Datenschutz sind hohe Werte an sich, die es zu schützen gilt. Dies gilt für die Verarbeitung personenbezogener und personenbeziehbarer Daten. Bei KI-Systemen im betrieblichen Kontext geht es aber um weit mehr: KI-Systeme bergen das Potenzial, die Arbeitsbedingungen und sozialen Beziehungen maßgeblich zu verändern.

Ein vorrangiges Ziel sollte sein, KI als Assistenzsysteme zu nutzen, um Arbeitsbelastungen zu reduzieren und Gute Arbeit zu fördern. Im Mittelpunkt muss dabei der Arbeits- und Gesundheitsschutz der Beschäftigten bei gleichzeitiger Wahrung ihrer Persönlichkeitsrechte stehen. Allein das breite Spektrum an Assistenzsystemen von einfacher Unterstützung (z. B. Fahrassistenten) bis hin zu leistungssteuernden oder auch de-qualifizierenden Systemen legt jedoch bereits nahe, die Wirkungsweisen der Anwendungen in verschiedene Kritikalitäts-Stufen zu unterteilen. Denn gleichzeitig bieten KI-Anwendungen neue Gefahren der Vermessung und Bewertung von Beschäftigten und Belegschaften – zum Beispiel durch individuelles Profiling, betriebliche Netzwerkanalysen („people / workplace analytics“) oder Selbststeuerungsmechanismen (von Gamification bis Scoring). Außerdem muss offen gelegt und mitbestimmt werden, wer die Einflussgrößen von Assistenz- und Unterstützungssystemen definiert, da bei plattformbasierten Anwendungen allein die Anbieter Zugriff auf diese Parameter haben. Eine solche Form der Effizienzoptimierung kann zusätzlichen Druck auf Beschäftigte entfachen, neue psychische Belastungen hervorrufen und am Ende neuartige Konflikte auslösen. Da unterschiedliche Stufen der algorithmischen Entscheidungsfindung bzw. Entscheidungsunterstützung möglich sind, müssen die Handlungsspielräume für Beschäftigte sowie die entsprechenden Verantwortlichkeiten geregelt werden.

Außerdem sollten arbeitsrechtliche Konsequenzen für Beschäftigte, die sich theoretisch aus einer „digitalen Führung“ bzw. Überwachung, prädiktiver Analyse und Selbstoptimierungsmechanismen ergeben können, verbindlich ausgeschlossen werden. Anderenfalls könnten Akzeptanzprobleme und Rechtsverstöße selbst bei einer Verbesserung der Ergonomie zu einem schwerwiegenden Hindernis für die betriebliche Umsetzung von KI-Systemen werden.

Es geht deshalb bei der Datennutzung nicht allein darum, rechtliche Grenzen einzuhalten, sondern „Gute Arbeit by design“ zu entwickeln. Dazu gehört – neben der Zielfindung – die Transparenz für nachvollziehbare und überprüfbare Informationen über das System sowie die Erklärbarkeit dessen Wirkungsweise und Wirkungszusammenhängen. Dazu gehört auch die Frage, zu welchem Zweck personenbeziehbare – und selbst anonymisierte – Daten genutzt werden sollen und dürfen. Dafür sollten entsprechende Verantwortlichkeiten festgelegt werden und eine stringente Zweckbindung sowie Folgenabschätzung umgesetzt werden.

Die nun folgenden „Leitfragen“ dieses Konzeptpapiers sind nicht als Checkliste zu verstehen, sondern als **Orientierung für ein neues Verständnis betrieblicher Aushandlungsprozesse**, um KI-Anwendungen für Gute Arbeit zu nutzen.

Leitfragen für den betrieblichen KI-Einsatz:

Angesichts der besonderen Bedingungen des KI-Einsatzes im Betrieb ist ein ganzheitlicher und präventiver Gestaltungsprozess zu wählen, bei dem betriebliche Interessenvertretungen und die beteiligten bzw. betroffenen Beschäftigten von Beginn an aktiv einzubinden sind. Für den „Change Impact Plan“ sind verschiedene Schritte zu empfehlen, die neben der Zieldefinition und der Formulierung von Ansprüchen gegenüber den KI-Systemanbietern insbesondere Regeln zur Nutzung von personenbeziehbaren Daten sowie eine betriebliche Abschätzung qualitativer wie quantitativer Folgen für die Arbeit im Betrieb mit dem Ziel einer humanen Arbeitsgestaltung einschließt.

Schritt 1: Klärung von Zielen und Zielkonflikten

Es besteht ein breiter politischer und auch wissenschaftlicher Konsens, dass offene und partizipative Verfahren vorteilhaft sind für die Akzeptanz und damit den erfolgreichen Einsatz von KI-Systemen im Betrieb. Der Prozess beginnt mit einer gemeinsamen Klärung der Ziele, Risiken und Grenzen, die mit der KI-Anwendung verbunden sind. Dazu ist eine Prozessvereinbarung für präventiv ausgerichtete Aushandlungs- und Beteiligungsprozesse anzustreben.

Die entscheidenden Fragen dabei sind:

- Welche Ziele werden mit dem Einsatz von KI-Anwendungen im Betrieb verfolgt?
Sind Zielkonflikte und Risiken absehbar und wie werden diese in den Bewertungsprozess eingebracht?
- Sind die vereinbarten Ziele für die KI-Nutzung verständlich und überprüfbar dokumentiert?
- Welche Beschäftigten oder Beschäftigtengruppen sind davon betroffen?
- Sind die Beschäftigten, Beschäftigtengruppen und ihre gesetzlichen Interessenvertretungen bei der Zielfindung und in die Planungen für den Einsatz, den Zweck und die ggf. zu verwendenden Beschäftigtendaten der KI-Anwendung einbezogen?
- Sind die dafür erforderlichen Kompetenzen zur Nutzung, Einschätzung und Interpretation der Wirkungsweise des KI-Systems vorhanden?
- Wird gewährleistet, dass die Interessenvertretungen bereits vor der Einführung der KI-Anwendung so beraten und qualifiziert werden, dass sie ihre Mitgestaltungs- und Mitwirkungsrechte von Anfang an ausreichend wahrnehmen können?
- Wird die betriebliche Expertise unter Einbeziehung der Beschäftigten genutzt und kann externer Sachverstand einbezogen werden?

Schritt 2: Anforderungen an KI-Anbieter bzw. Entwickler

Die Frage der Vertrauenswürdigkeit bezieht sich nicht allein auf die Funktionsweise, Robustheit oder Sicherheit von KI-Systemen, sondern ebenso auf die Anbieter solcher Systeme für die Arbeitswelt. Neben übergeordneten Grundprinzipien gelten besondere Anforderungen, die sich an KI-Anbieter richten und für die es ein Zertifizierungserfahren braucht. Es muss sichergestellt werden, dass die Systeme durch Menschen beherrschbar und gestaltbar bleiben. Bei der Gestaltung und dem Einsatz autonomer und selbstlernender Algorithmen gilt der Vorrang menschlicher Entscheidungshoheit und menschlicher Aufsicht. Zur Definition von Pflichten zur Transparenz, Nachvollziehbarkeit, Datenschutz für Anbieter haben die Datenschutzbeauftragten von Bund und Ländern mit der „Hambacher Erklärung zu Künstlicher Intelligenz“ (2019) Vorschläge erarbeitet. Dazu zählen:

Grundprinzipien

- Einhaltung rechtlicher Vorgaben, demokratischer Werte und Grundsätze durch KI-Anbieter/Entwickler (Grundgesetz, Grundrechte-Charta; Datenschutz-Grundverordnung, Bundesdatenschutzgesetz, Antidiskriminierungsgesetz u. a.)
- Einhaltung und Förderung der Ergonomie, des gendersensiblen Arbeits- und Gesundheitsschutzes und der Arbeitssicherheit
- Stärkung der Handlungsspielräume der Beschäftigten
- Förderung von Geschlechtergleichstellung und Diversität

Transparenz-Anforderungen für die KI-Anbieter und Entwickler („Beipackzettel“ für betriebliche Anwendung)

- Hohe Datenqualität und Datensicherheit (auch bei Auftragsdatenverarbeitung)
- Diskriminierungsfreiheit
- Gebrauchstauglichkeit / Usability
- Erklärbarkeit
- Transparenz der Einflussgrößen / Kriterien und „Übersetzung“ algorithmische Entscheidungskriterien und Funktionsweise der KI

- Überprüfbarkeit, Rückverfolgbarkeit und Korrigierbarkeit von Entscheidungen der KI-Anwendung
- Rechenschaftspflicht
- Möglichkeiten zur Folgenabschätzung

Nach der Zielvereinbarung (Schritt 1) und der Transparenz über die Wirkungsweise der KI-Anwendung (Schritt 2) sollten **Kritikalitätsstufen** zum Einsatz der KI definiert werden, um mögliche Risiken (und damit ggf. auch Anwendungen) auszuschließen oder die entsprechenden Maßnahmen der präventiven Arbeitsgestaltung im Sinne von „Gute Arbeit by design“ zu vereinbaren.

Schritt 3: Prozesstransparenz zur Datennutzung

Beim Einsatz von KI-Anwendungen ist neben der Zielsetzung die Frage von entscheidender Bedeutung, ob und zu welchem Zweck personenbeziehbare Beschäftigtendaten genutzt werden. Da KI-Systeme neben Leistungs- und Verhaltenskontrollen auch die Aussicht bieten, prädiktive Analysen der Beschäftigten ermöglichen, muss – neben der Zielsetzung für die KI und der Funktionalitäten der KI – eine frühzeitige, transparente und verbindliche, beteiligungsbasierte Regelungsvereinbarung als Grundlage für die betriebliche Nutzung personenbezogener Beschäftigtendaten getroffen werden. Dabei sollten im Vorfeld transparente Regeln über die Datenverarbeitung und betriebliche Nutzung, Analyse- und Auswertungsmöglichkeiten sowie deren Einsichtsoptionen als auch eine Folgenabschätzung bestimmt werden. Nach der Datenschutzgrundverordnung (DSGVO) gilt das Prinzip des „Privacy by Design“ (Art 25 DSGVO), um den Schutz der Persönlichkeitsrechte bei der Technikgestaltung zu gewährleisten.

Nach Art. 35 DSGVO ist für Systeme, die ein hohes Risiko für die Rechte und Freiheiten betroffener Personen (hier Beschäftigte) darstellen, eine Datenschutzfolgenabschätzung durchzuführen. KI-Systeme dürfen nur zu festgelegten, eindeutigen und legitimierten Zwecken (Art. 5 DSGVO) eingesetzt werden und brauchen eine Rechtsgrundlage (Art. 6 DSGVO) - Tarifverträge, Betriebs- oder Dienstvereinbarungen können in Deutschland nach § 26 BDSG eine Rechtsgrundlage darstellen. Zweckänderungen sind mit Art. 6 Abs. 4 DSGVO klare Grenzen gesetzt. Es ist zu beachten, dass an eine freiwillige Einwilligung als Rechtsgrundlage im Arbeitsverhältnis hohe Anforderungen (BAG 11.12.2014 – 8 AZR 1010/13) gestellt werden. Auch bei KI-Systemen müssen erweiterte Verarbeitungszwecke mit dem ursprünglichen Erhebungszweck vereinbar sein. Der praktische Nutzen (für die Verbesserung der Arbeitsbedingungen, Entwicklungsmöglichkeiten, Entlastungen) für die Beschäftigten muss erkennbar und nutzbar gemacht werden, dies gibt auch die DSGVO vor (beiderseitiger Nutzen als Grundlage zur Einwilligung im Arbeitsverhältnis).

- Werden personenbezogene und personenbeziehbare Daten – also Daten, welche Personen identifizierbar machen – verarbeitet?
- Ist die Nutzung personenbezogener Daten für die geplante Zielerreichung notwendig (Rechtmäßigkeit, Zweckbindung)?
- Sind die Beschäftigten ausreichend über die Datennutzung informiert? Sind die Auskunftsrechte der Betroffenen (DSGVO) bzgl. Verarbeitungszweck, Kategorien der Daten, Empfänger, Speicherdauer, Löschkonzept, Umfang etc. gewahrt?
- Ist transparent und kollektivrechtlich vereinbart, auf welche Weise (Speicherung, Dauer) Beschäftigtendaten genutzt werden sollen und wer wann und wie Zugang zu diesen Daten als auch zu den datenbasierten Analysemöglichkeiten / -ergebnissen hat („beteiligungsbasierte kollektive Nutzungsvereinbarung“ als Grundlage zur Einwilligung)?
- Ist eine Anonymisierung (Aggregation) von Beschäftigtendaten vereinbart?
- Werden arbeitsrechtliche Konsequenzen aufgrund der Nutzung (Auswertung) von Beschäftigtendaten im Betrieb durch entsprechende Vereinbarungen ausgeschlossen?

Schritt 4: Abschätzung der Folgen und humane Arbeitsgestaltung im Betrieb

KI-Anwendungen verändern Arbeitsprozesse, Strukturen und Aufgaben. Damit verbunden sind große Unsicherheiten und Sorgen bei den Beschäftigten hinsichtlich der persönlichen Zukunft (Arbeitsplatzsicherheit) und dem Zusammenspiel mit den „intelligenten“ Systemen (Interaktion von Mensch und KI). Auch sich die Frage, welche Qualifizierungsbedarfe durch KI ausgelöst werden oder wie sich die Belastungsprofile für Beschäftigte verändern. Es gilt, diese Fragen möglichst präventiv abzuschätzen und im weiteren Umsetzungsprozess stetig zu evaluieren. Auch hierfür sind Fragen der Veränderung von Arbeitsplätzen (Beschäftigungssicherung) und Aufgaben, Qualifizierungsbedarfen, Belastungsänderungen und Datennutzung entscheidend. Dazu kann ein Kritikalitätsraster erarbeitet werden, das die unterschiedlichen Aspekte aufnimmt.

- Welche Abläufe und Arbeitsplätze sind direkt oder indirekt und in welcher Form berührt?
- Welche Beschäftigungsgruppen sind betroffen und wie verändert sich die Beschäftigungssituation im Betrieb?
- Wie verändern sich Anforderungs- und Belastungsprofile der Arbeitsplätze?
- Sind Qualifizierungsmaßnahmen (für und wegen KI-Anwendungen) erforderlich?
- Gibt es ausreichend (zusätzliche) Zeit für die Qualifizierung und Einarbeitung?
- Sind Umschulungsmaßnahmen aufgrund des KI-Einsatzes für Beschäftigte erforderlich und welche Anforderungen ergeben sich an Neuorientierungen (Arbeitsidentität, Eingruppierung u. a.)
- Wie verändern sich Handlungsspielräume für Beschäftigte (Arbeitsteilung/ Mensch-Technik-Schnittstelle)?
- Entstehen neue Gefährdungen (physisch und psychisch) für Beschäftigte im Betrieb?
- Sind Leistungs- und Verhaltensanalysen mit dem KI-Einsatz möglich und/oder geplant?
- Sind prädiktive Persönlichkeits- oder Entwicklungsanalysen von Beschäftigten möglich oder geplant?
- Wie wird Werteneutralität und Transparenz des KI-Einsatzes sichergestellt bzw. Algorithmen-basierte Diskriminierung und deren Intransparenz verhindert?
- Ist die Einrichtung einer KI-Beschwerdestelle, über deren Zusammensetzung und Arbeitsweise die Interessenvertretung mitbestimmt, vorgesehen?

Schritt 5: Autonomie und Verantwortung beim betrieblichen Einsatz von KI

KI-Anwendungen können die Autonomie und Handlungsspielräume vom Beschäftigten am Arbeitsplatz erweitern oder auch einschränken. Deshalb stellt der Einsatz von KI im Betrieb hohe Anforderungen für eine gesundheitsförderliche und menschengerechte Arbeitsgestaltung als auch neue Fragen hinsichtlich der Verantwortung für den Umgang z. B. mit KI-basierten Entscheidungsvorschlägen. So darf der Einsatz von KI insbesondere die menschlichen Entscheidungsprärogativen in arbeitsrechtlichen Fragen nicht vollständig verdrängen. Das gilt insbesondere für (a) die Ausübung des Direktionsrechts durch den Arbeitgeber bzw. Vorgesetzte, (b) die Vorbereitung und Durchführung arbeitsrechtlicher Maßnahmen (von Abmahnung bis zur Kündigung) und (c) die Anbahnung des Arbeitsverhältnisses, Bewerbungsverfahren oder Personalauswahl. Hier muss letztlich immer ein Mensch endgültig entscheiden.

- Nach welchen Kriterien wird die Schnittstelle von KI-Anwendung und Beschäftigten gestaltet (Aufgabenteilung und Handlungsträgerschaft / Entscheidungskompetenz im Umgang mit KI-Systemen, Autonomiestufen menschlichen Handelns)?
- Welche Regelungen zur Kennzeichnungspflicht im Umgang mit KI-Anwendungen gibt es?
- Wie wird der Arbeitssteuerung (Weisung) durch KI Grenzen gesetzt ?
- Was folgt aus dem Einsatz von KI-Systemen für die Haftungsfragen im Arbeitsrecht? Welche bindende Wirkung haben Entscheidungsvorschläge der KI für Beschäftigte?
- Welche Konfliktlösungsmechanismen und Verantwortlichkeiten/Haftung im Umgang mit KI-Systemen gibt es – insbesondere auch für Schwachstellen des Systems?
- Wie wird der Einsatz von KI bei arbeitsrechtlichen Maßnahmen (Vorbereitung von Abmahnungen bis hin zu Kündigungen) begrenzt/geregelt ?
- Wie wird die Diskriminierungsfreiheit der KI-Anwendungen sichergestellt (zum Beispiel bei Bewerbungsmanagement/Personalauswahl)?

Schritt 6: Tests und Kontrollen

Der Einsatz von KI als lernende Systeme erfordert auch „lernende Betriebe“ - also Testphasen, Interventionen, eine Evaluation und ggf. auch Neujustierungen. Dabei sollten die Erfahrungen der Beschäftigten sowie ihrer Interessenvertretungen als Nutzerinnen und Nutzer explizit einbezogen werden und Formate gewählt werden, die eine Bewertung von Zielen und Risiken beinhalten. Die Mitbestimmung spielt dabei eine zentrale Rolle.

- Wie ist die Einführung geplant? Gibt es Experimentier- und Testphasen?
- Wie werden Experimentier- und Testphasen evaluiert und dokumentiert (Veränderungs- und Gestaltungserfordernisse für (a) Kompetenzentwicklung, (b) Belastungsprofile und (c) Handlungsspielräume für Beschäftigte)?
- Wie sind die Mitsprache- und Mitgestaltungsmöglichkeiten der Beschäftigten und deren gesetzlicher Interessensvertretungen für Experimentier- und Testphasen geregelt?
- Wie wird ein transparentes Feedbacksystem unter Einbeziehung der betroffenen Beschäftigten organisiert (Usability, Ziel-Nutzen, Belastungen, Handlungsträgerschaft etc.)?
- Welche Interventions- und Anpassungsmöglichkeiten für die KI-Nutzung gibt es für Beschäftigte und deren gesetzlichen Interessensvertretungen (Einrichtung einer Beschwerdestelle o. ä.)?

4. Beteiligungs- und Mitbestimmungsprozesse weiter denken

Mitbestimmungsprozesse sind für den KI-Einsatz im Betrieb unerlässlich, doch bislang begrenzt und zumeist Anlass bezogen und nicht prozessorientiert. Die betriebliche Einführung und Anwendung von „technischen Einrichtungen, die dazu bestimmt sind, das Verhalten oder die Leistung der Arbeitnehmer zu überwachen“ ist nach dem Betriebsverfassungsgesetz (87 Abs. 1 Nr. 6 BetrVG) mitbestimmungspflichtig. Auch Personalräte haben mitzubestimmen bei der „Einführung und Anwendung technischer Einrichtungen, die dazu bestimmt sind, das Verhalten oder die Leistung der Beschäftigten zu überwachen“ (§ 75 Abs. 3 Nr. 17 BPersVG). Die Betriebs- und Personalräte haben auch die Pflicht darauf zu achten, dass der Schutz der Privatsphäre sowie der informationellen Selbstbestimmung gewährleistet bleibt.

Dafür ist bereits die Möglichkeit der Überwachung durch neue technische Systeme ausschlaggebend. Dies gilt vor allem für die Möglichkeiten einer systematischen, datenbasierten Analyse von persönlichen Merkmalen der Beschäftigten (Profiling). Da der Umgang mit persönlichen und personenbeziehbaren Daten – nicht zuletzt aufgrund der prädiktiven Möglichkeiten durch KI – zu einer Sollbruchstelle für die erfolgreiche Einführung von KI-Systemen werden kann, muss dieser sensible Punkt bei der betrieblichen Umsetzung eine hervorgehobene Beachtung finden. Eine spezifische Herausforderung von KI mit Zugriff auf Massendaten ist die Möglichkeit zur Personalisierung nicht personenbezogener Daten, die erst kontextbezogen zu erkennen sind. Deshalb ist die Mitbestimmung auf die Nutzungsformen von Daten und Techniken zu erweitern, um Risiken auszuschließen und die Daten- und Analyse-Potenziale für Gute Arbeit zu nutzen. Dieser Punkt sollte nicht zuletzt bei Planungen für extern zu entwickelnde Lösungen Beachtung finden. Erforderlich ist ein hohes Maß an Transparenz und verbindliche Vereinbarungen über die analytischen Möglichkeiten und Grenzen. Die Mitbestimmung sollte dafür prozess- und beteiligungsorientierter ausgestaltet werden.

Vorteilhaft sind auch hier Pilot- und Experimentierphasen, um die Schnittstellen der KI-Anwendung hinsichtlich der Zielsetzungen zu überprüfen und unerwünschte Effekte auszuschließen. Wichtig ist, dass auch im weiteren Prozess Interventions- und Korrekturmöglichkeiten des KI-Einsatzes mit verbindlichen Mitbestimmungsmöglichkeiten geschaffen werden. Da die Akzeptanz des KI-Einsatzes dadurch auch bei der Umsetzung und Evaluation erhöht wird, sollte die Mitbestimmung für die gesamte Prozesskette institutionalisiert werden. Denkbar ist – über § 80 BetrVG bzw. § 68 BPersVG hinaus – die Ergänzung der existierenden Mitbestimmungsstrukturen um betriebliche und externe (wissenschaftliche oder gewerkschaftliche) Expertise und Prozessbegleitung. Angesichts der Dynamik von KI-Anwendungen sollte über „lebende Vereinbarungen“ nachgedacht werden, um notwendige Anpassungen im Prozess zu ermöglichen. Dies kann eine Grundlage für ein kooperatives Veränderungsmanagement auf Augenhöhe sein, indem die vereinbarten Prozess-Schritte hinsichtlich ihrer Wirkungen gemeinsam überprüft, ausgewertet und aufgrund der Lernerfahrungen ggf. angepasst werden.

Bereits auf der Grundlage des heute geltenden Rechtsrahmens ergeben sich dazu Handlungsoptionen für die Gestaltung von KI-Systemen:

im Betrieb:

- Vereinbarung zur Einrichtung von prozessualen Beteiligungs- und Mitbestimmungsstrukturen (zur Definition von Zielen, Kritikalitätsstufen, Change-Prozessen, Grenzen und Überprüfung)
- Beratungsangebote und gendersensibler Kompetenzaufbau für Beschäftigte und betriebliche Interessenvertretungen
- Beteiligungs- und prozessorientierte Umsetzung, z. B. durch frühzeitige, transparente und verbindliche Beteiligungsvereinbarungen (auch als Grundlage für eine freiwillige Einwilligung der betroffenen Beschäftigten zur betrieblichen Datennutzung)

auf der Ebene des Sozialpartner/Tarifvertragsparteien:

- Förderung und Vereinbarung prozessualer Beteiligungs- und Mitbestimmungsstrukturen auf betrieblicher Ebene – zum Beispiel durch Rahmentarifverträge
- Förderung der Qualifikations- und Kompetenzentwicklung für KI-Anwendungen (zum Beispiel Tarifverträge zur Weiterbildung in Bereichen wie z. B. Safety und Security, Datafizierung, Datenfolgenabschätzung, Grundlagen-Kenntnisse zur kritisch-konstruktiven Bewertung und Begleitung etc.)
- Beitrag zur Entwicklung von Branchen- oder Sektor spezifischen Standards (Leitlinien, Kritikalitätsmaßstäbe o. a.) für KI-Anwendungen und deren Einsatz
- Förderung von und Einsatz für Best Practice für KI-Anwendungen und des (branchenspezifischen) Transfers aus der KI- und Arbeitsforschung

5. Neue Anforderungen an die Politik

Der Orientierungsrahmen zeigt, dass präventive und nachhaltige Aushandlungs- und Beteiligungsprozesse zum Einsatz von KI im Betrieb anspruchsvoll und aufwändig sind. Gleichwohl sind sie für einen sicheren Umgang, Akzeptanz und einen ausreichenden Schutz der Beschäftigten unerlässlich. Der geltende Rechtsrahmen gibt den Akteuren der Arbeit Gestaltungsmöglichkeiten, die jedoch den Anforderungen zur Nutzung von KI-Anwendungen im Betrieb nur bedingt genügen. Die derzeitige Praxis zeigt Unsicherheiten und eine Mitbestimmung, die nicht ausreichend prozesshaft ausgerichtet ist. Der DGB schlägt deshalb eine Anpassung des Rechtsrahmens vor, die dazu beitragen soll, eine Innovationskultur zu etablieren, die auf Vertrauen und Sicherheit basiert.

10 Punkte-Plan für einen gesetzlichen Ordnungsrahmen für einen verlässlichen KI-Einsatz

1. Schaffung eines gesetzlich verankerten **Zertifizierungsverfahrens** und Aufbau von unabhängigen Prüf- und Beschwerdestellen zur demokratisch legitimierten Aufsicht und Kontrolle
2. Förderung betrieblicher Aushandlungsprozesse durch eine **Stärkung der Mitbestimmungsrechte**
 - a) bei der Einleitung, Planung und Durchführung von Kommunikations- und Beteiligungsprozessen von Beschäftigten sowie zur verbindlichen Prozessgestaltung und deren Kontrolle im Rahmen von mitbestimmungspflichtigen Angelegenheiten („prozedurales Mitbestimmungsrecht“)
 - b) durch Erweiterung des Initiativ- und Mitbestimmungsrechts bei der Durchführung betrieblicher Bildungsmaßnahmen zu einem generellen Mitbestimmungs- und Initiativrecht auch bei der Einführung der betrieblichen Berufsausbildung/Qualifizierung (da zu erwarten ist, dass der Einsatz von KI-Systemen zu – auch längerfristig und grundlegend – veränderten Qualifikationsanforderungen im Betrieb führt)
 - c) durch ein generelles Mitbestimmungsrecht bei der Nutzung personenbezogener und personenbeziehbarer Daten sowie zum Schutz der Persönlichkeitsrechte der Beschäftigten
 - d) verbunden mit der erleichterten Hinzuziehung von externem Sachverstand, die nicht vom vorherigen Einvernehmen mit dem Arbeitgeber abhängig ist.

3. Förderung der **Kompetenzentwicklung** von Betriebs- und Personalräten für den betrieblichen KI-Einsatz (Komplexität von KI-Systemen sowie datafizierungspolitische Sensibilisierung und Qualifizierung), bspw. durch staatlich geförderte Qualifizierungs- und Beratungsangebote
4. Einführung eines eigenständigen **Beschäftigtendatenschutzgesetzes**, um den besonderen Anforderungen bei der Verarbeitung personenbezogener Daten im Betrieb gerecht zu werden.
5. **Sachvortragsverwertungsverbot** und **Beweisverwertungsverbot** für rechtswidrig erlangte Beschäftigtendaten und deren Nutzung
6. Konkretisierung und Verbesserung der bestehenden Regelungen des Allgemeinen Gleichbehandlungsgesetzes (z. B. § 3, 12 und 15 AGG), um Beschäftigte vor Algorithmen-basierter **Diskriminierung zu schützen**.
7. Verbindlichkeit bei der Umsetzung von Prozessen für **Folgenabschätzung und Evaluation** von KI-Anwendungen (analog zur Datenfolgenabschätzung nach DSGVO) hinsichtlich der Veränderung der Belastungssituation im Betrieb
8. Verbindlichkeit für die gesetzlich vorgeschriebene **Gefährdungsbeurteilung** psychische Gesundheit und deren Anpassung an KI-Systeme (insb. Anti-Stress-Verordnung, Stärkung der Aufsicht)
9. Ausbau der **Arbeitsforschung** und kritischen Datafizierungsforschung zur Förderung der sozialpartnerschaftlichen Umsetzung von KI-Projekten (Entwicklung von verbindlichen Standards/Rahmenbedingungen)
10. **Sozialpartnerschaftlich abgestimmte Entwicklung und Einführung ethischer Leitlinien in der Ausbildung** und eines **hippokratischen Eids** für KI-Entwicklung hinsichtlich der arbeits- und gesellschaftspolitischen Implikationen von KI-Systemen sowie Unterstützungsmaßnahmen zur betrieblichen Orientierung an ethischen Leitlinien.