

Spielkamp, Matthias; Gießler, Sebastian

**Working Paper**

## Automatisiertes Personalmanagement und Mitbestimmung: KI-basierte Systeme für das Personalmanagement - was ist fair, was ist erlaubt?

Working Paper Forschungsförderung, No. 191

**Provided in Cooperation with:**

The Hans Böckler Foundation

*Suggested Citation:* Spielkamp, Matthias; Gießler, Sebastian (2020) : Automatisiertes Personalmanagement und Mitbestimmung: KI-basierte Systeme für das Personalmanagement - was ist fair, was ist erlaubt?, Working Paper Forschungsförderung, No. 191, Hans-Böckler-Stiftung, Düsseldorf

This Version is available at:

<https://hdl.handle.net/10419/224667>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/de/legalcode>

# WORKING PAPER FORSCHUNGSFÖRDERUNG

---

Nummer 191, September 2020

## Automatisiertes Personalmanagement und Mitbestimmung

KI-basierte Systeme für das  
Personalmanagement –  
was ist fair, was ist erlaubt?

Matthias Spielkamp und Sebastian Gießler

---

© 2020 by Hans-Böckler-Stiftung  
Georg-Glock-Straße 18, 40474 Düsseldorf  
[www.boeckler.de](http://www.boeckler.de)



„Automatisiertes Personalmanagement und Mitbestimmung“ von  
Matthias Spielkamp und Sebastian Gießler ist lizenziert unter

**Creative Commons Attribution 4.0 (BY).**

Diese Lizenz erlaubt unter Voraussetzung der Namensnennung des Urhebers die Bearbeitung, Vervielfältigung und Verbreitung des Materials in jedem Format oder Medium für beliebige Zwecke, auch kommerziell. (Lizenztext: <https://creativecommons.org/licenses/by/4.0/de/legalcode>)

Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z. B. von Schaubildern, Abbildungen, Fotos und Textauszügen erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

**ISSN 2509-2359**

# Inhalt

Zusammenfassung.....	4
Einleitung .....	8
Rechtliche und ethische Analysen .....	9
Rechtliche Rahmenbedingungen des Einsatzes von Algorithmen im Personalmanagement .....	9
Arbeitsrechtliche Aspekte und Beschäftigtendatenschutz .....	12
Ethische Aspekte .....	16
Arbeitshilfen für die Praxis.....	21
Ein Leitfaden für Betriebsräte und andere Personalvertretungen .....	21
Ein Puzzle zur Veranschaulichung von Machine Learning im Personalmanagement.....	23
Literatur.....	25
Autoren .....	26

## Zusammenfassung

Unternehmen, die im Personalmanagement Systeme mit so genannter Künstlicher Intelligenz verwenden, handeln möglicherweise rechtswidrig – oft, ohne sich darüber im Klaren zu sein. Denn „People-Analytics“-Verfahren dürfen nicht eingesetzt werden, ohne dass Beschäftigte individuell eingewilligt haben oder eine Betriebsvereinbarung vorliegt. Das wird in der Praxis oft nicht der Fall sein. Zudem erfüllen „Black-Box“-Systeme, deren Funktionsweise dem Betriebsrat auf Anfrage nicht erläutert werden kann, nicht die Auskunftsanforderungen des Betriebsverfassungsgesetzes. Autonomie und Mitsprache der Beschäftigten müssen auch dann gewährleistet sein, wenn algorithmische Systeme zum Personalmanagement eingesetzt werden; es ist ethisch nicht zu rechtfertigen, Arbeitnehmer\*innen zu reinen Objekten derartiger Verfahren zu machen. Doch sie müssen sich auch selber angemessene Kompetenzen erarbeiten, um diese Verfahren zu verstehen. Das ist das Ergebnis des Forschungsprojekts „Automatisiertes Personalmanagement und Mitbestimmung“, das AlgorithmWatch in den vergangenen zwei Jahren mit Mitteln der Hans-Böckler-Stiftung durchgeführt hat.

Beschäftigte müssen einzeln und freiwillig zustimmen, wenn sogenannte People-Analytics-Verfahren eingesetzt werden. Es ist jedoch davon auszugehen, dass diese Einwilligung sehr oft nicht vorliegt. Alternativ könnte eine Verarbeitung von Beschäftigtendaten im Rahmen von People Analytics auch durch Betriebsvereinbarungen legitimiert werden. Sie sind allerdings nicht erzwingbar, weil ein einschlägiges Mitbestimmungsrecht fehlt.

Betriebsräte und Unternehmen müssen prüfen, ob der Einsatz von KI-Systemen rechtskonform erfolgt. Es ist zu erwarten, dass Arbeitgeber in vielen Fällen den Beschäftigten und ihren gewählten Interessenvertretungen nicht in dem Maß Informationen zur Verfügung stellen können, wie es das Gesetz verlangt, weil die KI-Anbieter diese Informationen nicht herausgeben.

Die Bundesregierung muss gesetzlich klarstellen, dass Arbeitgeber auch dann Transparenz über die verwendeten Methoden gewährleisten müssen, wenn die Hersteller der Software keine Auskünfte erteilen wollen.

Hersteller müssen aktiv Möglichkeiten dafür entwickeln und anbieten, um sowohl den Beschäftigten als auch ihren Betriebsräten angemessen Auskunft über die Funktionsweise KI-basierter Systeme geben zu können.

Betriebsräte müssen ihre Auskunftsrechte gegenüber Arbeitgebern bereits in der Planungsphase von KI-Systemen aktiv durchsetzen, da

sich die praktische Wirksamkeit und Durchsetzbarkeit ihrer Mitbestimmungsrechte verringert, sobald Systeme eingeführt sind.

Unternehmen und Betriebsräte müssen dafür sorgen, dass auf Seiten der Beschäftigten und des Personalmanagements angemessene Kompetenz vorhanden ist, aus diesen Informationen die richtigen Schlüsse zu ziehen. Gewerkschaften sollten hierbei sowohl auf der betrieblichen, als auch der politischen Ebene unterstützen.

„Um Beschäftigte vor den Möglichkeiten und insbesondere vor dem immanenten hohen Kontrollpotenzial zu schützen, das sich mit KI-Software und mit selbstlernenden Algorithmen verbindet, ist ein Ausbau bestehender Mitwirkungs- und Mitbestimmungsrechte durch den Gesetzgeber unumgänglich.“ So fasst der Spezialist für Arbeitsrecht und Digitalisierung, Peter Wedde, die aktuelle Situation in seinem Gutachten „Automatisierung im Personalmanagement – arbeitsrechtliche Aspekte und Beschäftigtendatenschutz“ zusammen, das im Rahmen des Projekts entstanden ist. Von solchen wünschenswerten Änderungen abgesehen, stellt sich aber bereits die geltende Rechtslage so dar, dass Arbeitgeber schon mit Beginn der Planung von KI-basierter Personalsoftware Betriebsräten weitreichende Auskünfte über deren Funktionsweise geben müssen. In vielen Fällen verfügen sie aber gar nicht über diese Informationen, da die Anbieterfirmen sie nicht preis geben. Zudem müssen Beschäftigte einzeln und freiwillig zustimmen, dass ihre personenbezogenen Daten mit People-Analytics-Verfahren verarbeitet werden, es sei denn, diese Verarbeitung ist durch einschlägige Betriebsvereinbarungen datenschutzrechtlich legitimiert. Es ist jedoch davon auszugehen, dass sehr oft weder individuelle Einwilligungen vorliegen, noch Betriebsräte bereit dazu sind, entsprechende Betriebsvereinbarungen abzuschließen. Damit befinden sich Firmen, die derartige Systeme einsetzen, zumindest in einer rechtlichen Grauzone; oft wird der Einsatz der Systeme sogar rechtswidrig sein.

Um zu analysieren und vorherzusagen, wie Individuen oder Teams arbeiten oder ob sie ihre Ziele erreichen werden, setzen die Anbieter der Systeme statistische Methoden der Mustererkennung und Wahrscheinlichkeitsrechnung ein, etwa Maschinelles Lernen oder „Deep Neural Networks“. Auf Basis welcher Modelle und Annahmen diese Prognosen getroffen werden und mit welchen Daten die Systeme trainiert wurden, halten die Anbieter üblicherweise geheim – meist mit der Begründung, dass sie sich zum einen vor Nachahmern schützen müssen, zum anderen Beschäftigte Aussagen der Systeme manipulieren könnten, wenn ihnen ihre Funktionsweise bekannt wäre.

Aus ethischer Perspektive sind diese Bedenken zwar durchaus begründet, lassen in dieser Einseitigkeit jedoch gerechtfertigte Ansprüche

der Beschäftigten außer Acht und sind daher so nicht akzeptabel. Wie Michele Loi, Experte für ethische Aspekte des Einsatzes von Methoden Künstlicher Intelligenz der Universität Zürich, in seiner Studie „People Analytics must benefit the people. An ethical analysis of data-driven algorithmic systems in human resources management“ darlegt, haben Beschäftigte einen Anspruch darauf, die Logik derartiger Systeme verstehen zu können. Denn das ist eine notwendige – wenn auch keine hinreichende – Voraussetzung dafür, dass die Autonomie der Beschäftigten gewahrt bleibt; sie dürfen keine passiven Objekte einer von Algorithmen getriebenen Steuerung („algorithmic governance“) werden. Vielmehr müssen sie aktiv dazu beitragen können, die Leistung der Organisation zu verbessern, indem sie Erkenntnisse nutzen, die durch den Einsatz dieser algorithmischen Systeme entstehen. Letztendlich müssen die Systeme so gestaltet und eingesetzt werden, dass immer auch die Beschäftigten davon profitieren, nicht ausschließlich der Arbeitgeber.

Es kann also nicht davon ausgegangen werden, dass Anbieter in der Lage sind, „ethische Algorithmen“ zu entwickeln, die Arbeitgeber dann per se unproblematisch einsetzen können. Im Gegenteil, es müssen alle beteiligten Akteure – die Anbieter algorithmischer Systeme, die Unternehmen, die sie einsetzen, und die Beschäftigten, die davon betroffen sind – bereit sein, gemeinsam Lösungen zu erarbeiten, die die verschiedenen Interessen berücksichtigen. Zu diesem Zweck müssen die Anbieter Informationen zur Verfügung stellen, die es Arbeitgebern und Beschäftigten erlauben zu verstehen, wie die Systeme funktionieren. Zugleich kann es in vielen Fällen gerechtfertigt sein, dass Firmen ihre Investitionen auch mithilfe von Betriebs- und Geschäftsgeheimnissen schützen.

Um hier eine angemessene Balance zu erreichen, hat Algorithm-Watch in Zusammenarbeit mit Sebastian Stiller, Professor für Mathematische Optimierung der Technischen Universität Braunschweig, einen Prüfkatalog entwickelt: „Automatisierte Entscheidungen und Künstliche Intelligenz im Personalmanagement. Ein Leitfaden zur Überprüfung essenzieller Eigenschaften KI-basierter Systeme für Betriebsräte und andere Personalvertretungen“. Er ermöglicht es Betriebsräten und Personalmanager\*innen, von Anbieterfirmen Informationen zu erfragen, die essentiell dafür sind, die Funktionsweise, Qualitäts- und Zuverlässigkeitskriterien algorithmischer Systeme zu verstehen, ohne aber beispielsweise zu verlangen, dass Programmiercode oder die Datenbasis offengelegt werden müssen.

Angesichts der Komplexität der Verfahren wäre allerdings die Forderung verfehlt, sie Laien gegenüber auf einfache Art erklär- oder nachvollziehbar zu machen. Stattdessen müssen alle Akteure – Anbieter von

Personalmanagementprodukten, Anwender (Personalabteilungen) und Betroffene (Beschäftigtenvertreter\*innen) – aktiv daran mitarbeiten, Systeme angemessen zu entwickeln und einzusetzen. Das bedeutet für die Vertreter\*innen der Beschäftigten, dass sie sich ein Grundverständnis von Methoden und Verfahren der KI-basierten Systeme aneignen oder Möglichkeiten schaffen müssen, auf diese Expertise zuzugreifen, etwa durch externe Sachverständige.



## Einleitung

Anlass für das Forschungsprojekt „Automatisiertes Personalmanagement und Mitbestimmung“ ist der zunehmende Einsatz von Personalmanagement-Software, die sogenannte „People-Analytics“-Fähigkeiten bietet – also die Möglichkeit, Leistungen und Potenzial von Mitarbeiter\*innen oder Teams zu analysieren, vorauszusagen oder zu steuern. Bekannte Produkte, die mit derartigen Funktionen werben, sind Office 365 Workplace Analytics von Microsoft, Watson Talent Insights von IBM, Success Factors People Analytics von SAP und People Analytics von Workday. Auf dem Markt sind darüber hinaus etliche Produkte kleinerer Firmen.

Um zu analysieren und vorherzusagen, wie Individuen oder Teams arbeiten oder ob sie ihre Ziele erreichen werden, setzen die Anbieter der Systeme statistische Methoden der Mustererkennung und Wahrscheinlichkeitsrechnung ein, etwa Maschinelles Lernen oder „Deep Neural Networks“. Auf Basis welcher Modelle und Annahmen diese Prognosen getroffen werden und mit welchen Daten die Systeme trainiert wurden, halten die Anbieter üblicherweise geheim – meist mit der Begründung, dass sie sich zum einen vor Nachahmern schützen müssen, zum anderen Beschäftigte Aussagen der Systeme manipulieren könnten, wenn ihnen ihre Funktionsweise bekannt wäre. Das macht einen wissenschaftlichen und rechtlichen Zugriff auf diese Systeme schwierig.

Die Kombination aus der Geheimhaltung der Arbeitsweise dieser Systeme und der hohen Wissenshürde, die ein Verständnis von People Analytics voraussetzt, erschwert ein intuitives Verständnis dieser Produkte. Darüber hinaus wurden viele Enterprise-Softwareprodukte als People-Analytics-Systeme bezeichnet und mit Künstlicher Intelligenz in Verbindung gebracht, die derartige Kapazitäten nicht haben. So ergibt sich eine Mischung aus Technikoptimismus, Marketingversprechen und technisch bedingter Undurchsichtigkeit. Die erste Herausforderung war daher herauszuarbeiten, was People Analytics ist und über welche Fähigkeiten diese Softwaresysteme verfügen. Diese Grundlagenarbeit ist relevant, da sie den Kern unserer Argumentation und Basis für unsere Empfehlungen darstellt.

Die disziplinären Perspektiven reichen von Ethik, über Informatik bis zu Datenschutz- und Arbeitsrecht. Die Neuorganisation der Steuerung und Kontrolle von Arbeit macht eine derlei breite fachliche Perspektive nötig. Ziel war es, empirisch und fachlich fundierte Empfehlungen über den Umgang mit People Analytics für Arbeitgeber und Arbeitnehmer zur Verfügung zu stellen.

# Rechtliche und ethische Analysen

## Rechtliche Rahmenbedingungen des Einsatzes von Algorithmen im Personalmanagement

Lewinski, Kai v./Barros Fritz, Raphael de/Biermeier, Katrin (2019): Bestehende und künftige Regelungen des Einsatzes von Algorithmen im HR-Bereich. Berlin: AlgorithmWatch.

<https://algorithmwatch.org/project/auto-hr/rechtsgutachten-von-lewinski/>

Bei der Diskussion von gesetzlichen Regulierungen algorithmischer Systeme ist es sinnvoll, zu prüfen welche rechtlichen Rahmenbedingungen bereits existieren. Die Spannweite der berührten Rechtsgebiete reichen dabei vom Datenschutzrecht, über das Arbeitsrecht bis zum Diskriminierungsrecht. Ebenso gibt es eine Vielzahl sonstiger Regelungen in angrenzenden Rechtsgebieten die für die Regulierung von algorithmischen Systemen relevant sein können. Das Gutachten von Prof. Dr. Kai von Lewinski, Dipl. Jur. Raphael de Barros Fritz, LL.M. und cand. iur. Katrin Biermeier trägt diese fragmentierten Regulierungsansätze zusammen und bietet eine konzise Analyse der vorherrschenden Rechtsprechung in Bezug algorithmische Systeme.

Laut der Analyse zerfällt der Regulierungsbedarf zwei Felder: Dem innerbetrieblichen Bereich, dem gesamtgesellschaftlichen Bereich sowie dem Technikrecht. Lewinski et al. argumentieren, dass für den Einsatz von algorithmischen Systemen durch die umfassenden Schutzmechanismen im Arbeitsrecht, Antidiskriminierungsgesetzgebung und Arbeitsschutz bereits enge Grenzen gesetzt werden. Auf individueller Ebene müssen sind beispielsweise schon jetzt Regelungen zu Arbeitszeit, Ungleichbehandlung und Datenschutz beachtet werden. Einer maximal effizienten Gestaltung von Personaleinsatz durch ADM stehen so beispielsweise die Regelungen für Ruhezeiten und das Sonn- und Feiertagsarbeitsverbot (§§ 5, 9 ff. ArbZG) sowie Regeln zur Teilzeitarbeit (§ 12 Abs. 2 TzBfG) entgegen.

Auch im Datenschutzrecht sind bereits viele Elemente von ADM im Personalwesen reguliert. Jede personenbezogene Datenverarbeitung muss durch bestimmte Tatbestände im Datenschutzrecht gerechtfertigt sein; das gilt auch für Anwendung von Algorithmen, um über Aufgabenverteilung und Beförderungen zu entscheiden. Personenbezogene Dauerüberwachung von Arbeitnehmern ist nur zulässig, wenn eine freiwillige

Einwilligung der Mitarbeiter vorliegt oder man sie aus guten Gründen einer schwerwiegenden Pflichtverletzung verdächtigt. Das Datenschutzrecht verbietet zudem gemäß Art. 22 DSGVO, dass eine rechnergestützte Entscheidung ausschließlich vom Computer getroffen wird – es muss immer ein (qualifizierter!) Mensch dazu- oder dazwischentreten. Das heißt, ein Mensch darf nicht im Einzelfall einem Algorithmus ausgeliefert sein (Art. 22 DSGVO). Eine Möglichkeit für Unternehmen: Bei algorithmischen Verfahren sollten Betroffene Gelegenheit bekommen, selbst Angaben zu machen, um auf unvollständiger Datenbasis beruhende Entscheidungen zu verhindern; diese Angaben müssen dann berücksichtigt werden. Betroffene müssen zudem ausreichend über Datenverarbeitungen informiert sein (Art. 13 Abs. 2 lit. f, 14 Abs. 2 lit. g, 15 Abs. 1 lit. h DSGVO). Auch in Bezug auf das Diskriminierungspotential von ADM gibt es bereits rechtliche Möglichkeiten. Das Diskriminierungsrecht verlangt von Unternehmen, dass sie Präventivmaßnahmen gegen Benachteiligungen gem. § 1 AGG ergreifen. Grundsätzlich verboten ist zudem die Verarbeitung „sensibler Daten“ gem. Art. 9 DSGVO.

Dem Einsatz von ADM sind auch durch das betriebliche Mitbestimmungsrecht Grenzen gesetzt. Verändern sich Beginn und Ende der täglichen Arbeitszeit oder verlagern sich Arbeitstage innerhalb der Woche (sei es implizit), so entsteht Mitbestimmungsrecht des Betriebsrats (§ 87 Abs. 1 Nr. 2 BetrVG). Auch bei einer vorübergehenden Änderung der regelmäßigen Arbeitszeit muss möglicherweise der Betriebsrat beteiligt werden (§ 87 Abs. 1 Nr. 3 BetrVG). Besonders wichtig für die Industrie 4.0: § 87 Abs. 1 Nr. 6 BetrVG, der technische Einrichtungen betrifft, mit denen Verhalten oder die Leistung von Mitarbeitern überwacht werden kann. Auch wenn Unternehmen für die Umsetzung Dritte einschalten und nicht einmal Zugriff auf die erfassten Daten haben, gilt die Mitbestimmung. Werden dagegen nur anonyme Daten erhoben oder wird nur die Leistung einer ganzen Abteilung oder Gruppe ausgewertet, hat der Betriebsrat in der Regel kein Mitspracherecht. Etwas Anderes gilt wiederum, wenn der Überwachungsdruck auf die Gruppe bis zu einzelnen Kollegen durchschlägt. Die Mitbestimmung gilt nicht nur bei der Datennutzung, sondern schon deren Erhebung. Cyber-physikalische Systeme, also Identifizierungen mittels RFID-Chips oder Fingerprint-Scanner-Systemen, sind meist mitbestimmungspflichtig.

An dieser Stelle zeigt sich die Grenze zwischen rechtlichen Regelungen auf individueller und gesamtgesellschaftlicher Ebene. In Deutschland kann, etwa im Vergleich zu den USA, ADM nicht beliebig eingesetzt werden. Dies ist nach gültiger Rechtslage nicht möglich. Was aber passiert, wenn automatisierte Systeme in Kombination mit der (vermeintlichen) Selbstständigkeit von angestellten Personen einhergeht? Die Mit-

bestimmungsrechte von Betriebsräten greift in diesem Fall nicht mehr. Im rechtlichen Umgang mit diesen Systemen benötigt es, unabhängig der Vielzahl an Einzelregelungen, ebenso eine auf automatisierte Systeme zugeschnittene allgemeine Regulierung. Denkbar sind etwa Diversitätsvorgaben für Entwickler\*innen oder eine Pflicht zum „vorurteilsfreien Programmieren“. Das könnte sogar in eine Pflicht zum Einsatz von ADM münden. Ebenso gibt es in anderen Rechtsgebieten Vorbilder wie automatisierte Systeme reguliert werden könnten:

- Das Wertpapierrecht kennt bereits Notfallvorkehrungen für unvorhergesehene Störungen (§ 80 Abs. 2 S. 4 WpHG).
- Das Verwaltungsrecht kennt die Beschränkung vollautomatischer Entscheidungen auf bestimmte Einsatzfelder (§ 35a VwVfG) – eine solche Eingrenzung auf Verfahren, die keine irreversiblen Folgen für den Menschen haben, ist zumindest bei selbstlernenden Algorithmen angezeigt.
- Produkthaftungsregeln könnten auf Hersteller von Algorithmen ausgedehnt werden. Zu erwägen ist außerdem eine Zertifizierungspflicht („Algorithmen-TÜV“). Unabhängige Fachleute könnten eingesetzte Algorithmen überprüfen und zertifizieren. Selbstlernende Systeme sollten fortlaufend geprüft werden. Allerdings sind solche Prüfungen praktisch schwer umsetzbar, da die Arbeit von Algorithmen stark vom Input abhängt und der sich laufend ändert.
- Möglicherweise braucht es auch umfassende Überwachungs- und Beobachtungspflichten, ähnlich wie es schon jetzt bei automatisiertem Fahren geregelt ist (§ 1b StVG).
- Der blinde Fleck bei der bestehenden Regulierung von ADM liegt eher im Technikrecht und in den gesamtgesellschaftlichen Auswirkungen. Personalabteilungen und Betriebsratsbüros können mit den bestehenden Regelungen erst einmal (weiter-)arbeiten. Arbeitnehmer und Arbeitgeber auf Verbandsebene und die Gesellschaft insgesamt müssen erst noch passende Normen für den Umgang mit ADM finden.

## Arbeitsrechtliche Aspekte und Beschäftigtendatenschutz

Wedde, Peter (2020): Automatisierung im Personalmanagement – arbeitsrechtliche Aspekte und Beschäftigtendatenschutz. Berlin: AlgorithmWatch.

<https://algorithmwatch.org/project/auto-hr/gutachten-arbeitsrecht-datenschutz-wedde/>

„Um Beschäftigte vor den Möglichkeiten und insbesondere vor dem immanenten hohen Kontrollpotenzial zu schützen, das sich mit KI-Software und mit selbstlernenden Algorithmen verbindet, ist ein Ausbau bestehender Mitwirkungs- und Mitbestimmungsrechte durch den Gesetzgeber unumgänglich.“ So fasst der Spezialist für Arbeitsrecht und Digitalisierung, Peter Wedde, die aktuelle Situation in seinem Gutachten Automatisierung im Personalmanagement – arbeitsrechtliche Aspekte und Beschäftigtendatenschutz zusammen, das im Rahmen des Projekts entstanden ist. Von solchen wünschenswerten Änderungen abgesehen, stellt sich aber bereits die geltende Rechtslage so dar, dass Arbeitgeber schon mit Beginn der Planung von KI-basierter Personalsoftware Betriebsräten weitreichende Auskünfte über deren Funktionsweise geben müssen. In vielen Fällen verfügen sie aber gar nicht über diese Informationen, da die Anbieterfirmen sie nicht Preis geben. Zudem müssen Beschäftigte einzeln und freiwillig zustimmen, dass ihre personenbezogenen Daten mit People-Analytics-Verfahren verarbeitet werden, es sei denn, diese Verarbeitung ist durch einschlägige Betriebsvereinbarungen datenschutzrechtlich legitimiert. Es ist jedoch davon auszugehen, dass sehr oft weder individuelle Einwilligungen vorliegen, noch Betriebsräte bereit dazu sind, entsprechende Betriebsvereinbarungen abzuschließen. Damit befinden sich Firmen, die derartige Systeme einsetzen, zumindest in einer rechtlichen Grauzone; oft wird der Einsatz der Systeme sogar rechtswidrig sein.

Betriebsräte haben bezogen auf KI-Systeme und die hier verwendeten Algorithmen sowohl umfangreiche Informations- und Beratungsrechte als auch wirksame Mitbestimmungsrechte. Sie leiten sich aus dem allgemeinen Informationsanspruch des § 80 Abs. 2 BetrVG ab. Dieser ist weit gefasst und gilt uneingeschränkt auch für die Automatisierung des Personalmanagements mittels KI-Systemen. Über die entsprechenden Planungen müssen Arbeitgeber die Betriebsräte rechtzeitig und umfassend informieren. Gleiches gilt für den Informationsanspruch nach § 90 Abs. 1 Nr. 2 und 3 BetrVG bezüglich der Planung von technischen Anlagen oder von Arbeitsverfahren und Arbeitsabläufen. Die Informationen,

die der Arbeitgeber mitzuteilen hat, müssen beispielsweise Ausführungen dazu beinhalten, wie verwendete technische Anlagen funktionieren oder wie Arbeitsverfahren verändert werden sollen. Bezogen auf KI-Systeme gehören hierzu auch Aussagen, ob und wenn ja welche Software bzw. welche KI-Systeme eingesetzt werden sollen und welche Effekte oder Veränderungen Arbeitgeber sich davon versprechen.

Arbeitgeber müssen bezüglich verwendeter KI-Systeme und der Algorithmen, die darin zum Einsatz kommen, sicherstellen, dass sie ihre gesetzlichen Informationsverpflichtungen gegenüber Betriebsräten erfüllen können. Über die entsprechenden Kenntnisse müssen sie verfügen, um ihre Verpflichtungen nach der Datenschutzgrundverordnung und dem Bundesdatenschutzgesetz erfüllen zu können. Hierzu gehören beispielsweise die in Art. 5 Abs. 2 DSGVO verankerten Rechenschaftspflichten, die Arbeitgeber als datenschutzrechtliche Verantwortliche haben. Sie müssen dafür sorgen, dass im Betrieb nur solche KI-Systeme eingesetzt werden, die den datenschutzrechtlichen Anforderungen an die Transparenz der Verarbeitung oder eine klare Zweckbindung gerecht werden. Entsprechendes gilt auf der kollektivrechtlichen Ebene, auf der es dem Arbeitgeber obliegt, dem Betriebsrat die Informationen zur Verfügung zu stellen, die dieser für die Wahrnehmung seiner gesetzlichen Aufgaben und Mitbestimmungsrechte benötigt.

### **Betriebsräte müssen feststellen können, ob KI-basierte Systeme diskriminieren**

Das Zustimmungsrecht des Betriebsrats zu Auswahlrichtlinien in § 95 BetrVG eröffnet den Weg zur Mitbestimmung der Ausgestaltung interner Personalprozesse. Dieses Zustimmungsrecht erfasst einen wesentlichen Teil des operativen Personalmanagements und die dort stattfindenden automatisierten Prozesse. Es besteht beispielsweise auch bezogen auf KI-Systeme, die vorliegende personenbezogene Informationen über einzelne Arbeitnehmer eigenständig analysieren und daraus allgemeine Entscheidungsvorschläge oder Richtlinien für die Personalauswahl ableiten. Damit unterliegen auch Software-Anwendungen, die die Personalauswahl teil- oder vollautomatisiert vornehmen, bezüglich der fachlichen und persönlichen Voraussetzungen und der sozialen Gesichtspunkte dem Mitbestimmungsrecht nach § 95 BetrVG.

Kollektivrechtliche Informationsansprüche der Betriebsräte müssen präzisiert und abgesichert werden, indem Arbeitgeber gesetzlich verpflichtet werden, nur solche technischen Einrichtungen einzusetzen und zu verwenden, deren Funktionsweisen sie detailliert kennen und die sie

so anpassen (lassen) können, dass sie gesetzlichen Vorgaben ebenso gerecht werden wie Anforderungen, die sich aus durchgeführten Mitbestimmungsprozessen ableiten.

Dieses Ziel könnte durch eine gesetzliche Verpflichtung der Arbeitgeber erreicht werden, Betriebsräten bezogen auf KI-Systeme und die hier verwendeten Algorithmen, die eingesetzt werden sollen, strukturierte Informationen zur Verfügung zu stellen. Dies würde einerseits die Verpflichtung beinhalten, nur solche Systeme einzuführen und anzuwenden, über die entsprechende Informationen verfügbar sind. Andererseits kommen auch neue Wege in Betracht, wie etwa der Einsatz von Fragenkatalogen, aus deren qualifizierter Beantwortung sich notwendige Informationen zu den Möglichkeiten und zur Funktionsweise der eingesetzten IT-Systeme ableiten lassen.

Auf der Ebene der Mitbestimmung müssen bestehende Rechte nach dem Betriebsverfassungsgesetz zugunsten der Betriebsräte angepasst werden. Mit Blick auf die Einführung und Änderung technischer Einrichtungen, die zur Verhaltens- oder Leistungskontrolle bestimmt sind, müssen Auslegungsgrundsätze, die von den Gerichten geschaffen wurden, in § 87 Abs. 1 Nr. 6 BetrVG verankert werden. Hierzu gehört insbesondere die Entscheidung des Bundesarbeitsgerichts, dass ein Mitbestimmungsrecht unabhängig davon begründet sein muss, ob der Arbeitgeber Verhaltens- und Leistungskontrollen auch tatsächlich durchführen will; es muss ausreichen, dass ein eingesetztes System es technisch möglich macht.

Denn wenn dieses Mitbestimmungsrecht nur in den Fällen greift, in denen Arbeitgeber eine Verhaltens- und Leistungskontrolle explizit wollen, wäre Betriebsräten die Möglichkeit genommen zu prüfen, ob tatsächlich keine Kontrollen erfolgen werden, bevor ein System eingeführt oder geändert wird. Stellen sie erst nach der Einführung oder Änderung fest, dass Kontrollen stattfinden, wird es für sie faktisch unmöglich, einen Stopp des Betriebs im Wege des einstweiligen Rechtsschutzes durchzusetzen.

Bezogen auf die Einführung und Anwendung von KI-Systemen sollte im Betriebsverfassungsgesetz verankert werden, dass die Hinzuziehung externer Sachverständiger durch Betriebsräte in diesen Fällen grundsätzlich als erforderlich zu qualifizieren ist und dass ein entsprechender Anspruch ab einer bestimmten „KI-Schwelle“ automatisch besteht. Als „Schwellwert“ könnte beispielsweise die Einführung oder Anwendung von KI-Systemen festgelegt werden, die automatisch Entscheidungsvorschläge oder Analysen erstellen, ohne dass die verwendeten Algorithmen für Betriebsräte und Beschäftigte ohne IT-Kenntnisse nachvollziehbar sind. Gleiches könnte für KI-Systeme gelten, die eine „selbstlernen-

de“ Logik bzw. entsprechende Software enthalten. Ein Vorbild für die Vereinfachung der Hinzuziehung von externen Sachverständigen in diesen Fällen könnte die Regelung in § 111 Satz 2 BetrVG für die Hinzuziehung von Beratern sein, allerdings ohne die dort genannte Mindestzahl von Betroffenen. In jedem Fall müssten Betriebsräte fachkompetente Sachverständige ohne Beeinflussung durch den Arbeitgeber eigenständig auswählen können.

Die Bundesregierung hatte in ihrer KI-Strategie vom November 2018 angekündigt zu „ermöglichen, dass sich Arbeitgeber und Betriebsrat in IT-Fragen im Zusammenhang mit dem Mitbestimmungsrecht nach § 87 Abs. 1 Nr. 6 BetrVG, gemeinsam auf die Hinzuziehung eines entsprechenden externen Sachverständigen einigen.“ Dies „soll es insbesondere den Betriebsräten ermöglichen, ihr Mitbestimmungsrecht nach § 87 Abs. 1 Nr. 6 BetrVG bei der Einführung und Anwendung von KI- sowie technischen Anwendungen, die dazu bestimmt sind, das Verhalten oder die Leistungen der Arbeitnehmer zu überwachen, besser auszuüben.“ Die in diesem Vorschlag angedachte Einigung zwischen Arbeitgeber und Betriebsrat würde jedoch hinter dem Standard von § 80 Abs. 3 BetrVG zurückfallen, dem zufolge der Betriebsrat nach eigener Wahl externe Sachverständige hinzuziehen kann, sofern es für die Erfüllung seiner Aufgaben erforderlich ist (vgl. Wedde 2020, S. 44).

Neue Probleme für Beschäftigte können bezogen auf das automatisierte Personalmanagement entstehen, wenn Systeme zur Anwendung kommen, die auf der Grundlage von Data-Mining-Konzepten Vorhersagen zum vermutlichen Verhalten von Beschäftigten machen. Derartige Systeme greifen im Regelfall sehr weit in Persönlichkeitsrechte der Betroffenen ein und werden im Ergebnis einer Verhältnismäßigkeitsprüfung im Regelfall als nicht zulässig zu qualifizieren sein. Dies schützt aber Beschäftigte nicht davor, dass sie dennoch eingesetzt werden und dass dieser Einsatz in betriebsratlosen Betrieben ohne kollektivrechtlich mögliche Begrenzungen erfolgt. Gleiches gilt für sprachanalysierende KI-Systeme.

Der unangemessene Einsatz derartiger Systeme ließe sich nur durch klare gesetzliche Verbote vermeiden, die Bestandteil eines umfassenden Beschäftigtendatenschutzgesetzes sein könnten. Eine solche Regelung lässt aber weiter auf sich warten. Die Notwendigkeit eines speziellen Beschäftigtendatenschutzgesetzes mit Regelungen zu Themen wie dem expliziten Ausschluss heimlicher Überwachungen wird vom deutschen Gesetzgeber gesehen. In der amtlichen Begründung zum Datenschutz-Anpassungs- und -Umsetzungsgesetz EU, das als Artikelgesetz insbesondere das neue BDSG enthält, heißt es hierzu:



„Der Gesetzgeber behält sich vor, Fragen des Datenschutzes im Beschäftigungsverhältnis innerhalb dieser Vorschrift oder im Rahmen eines gesonderten Gesetzes konkretisierend bestimmte Grundsätze, die im Rahmen der Rechtsprechung zum geltenden Recht bereits angelegt sind, zu regeln. Dies gilt insbesondere für das Fragerecht bei der Begründung eines Beschäftigungsverhältnisses, den expliziten Ausschluss von heimlichen Kontrollen im Beschäftigungsverhältnis, die Begrenzung der Lokalisierung von Beschäftigten sowie den Ausschluss von umfassenden Bewegungsprofilen, den Ausschluss von Dauerüberwachungen und die Verwendung biometrischer Daten zu Authentifizierungs- und Autorisierungszwecken.“

In ihrer KI-Strategie hatte die Bundesregierung angekündigt,

„die betrieblichen Mitbestimmungsmöglichkeiten bei der Einführung von KI-Anwendungen [zu] sichern und [zu] prüfen, ob im Rahmen eines eigenständigen Beschäftigtendatenschutzgesetzes mehr Rechtssicherheit bei der Einführung entsprechender Anwendungen im Betrieb geschaffen werden kann.“

Gerade angesichts der hier dargelegten Situation ist es höchste Zeit, diese Ankündigung in die Tat umzusetzen.

Erfolgt eine (datenschutz-)rechtlich unzulässige Verarbeitung von Beschäftigtendaten und erhalten Arbeitgeber als deren Ergebnis Informationen, die sie für arbeitsrechtliche Maßnahmen gegen Beschäftigte nutzen, stellt sich als weiteres das Problem, dass das deutsche Arbeitsrecht kein „Beweisverwertungsverbot“ zugunsten der Beschäftigten kennt (Wedde 2020, S. 41 f.). Nur ein eindeutiges gesetzliches Verbot der prozessualen Verwendung von Beweisen, die unter Verstoß gegen datenschutz- oder arbeitsrechtliche Vorschriften gewonnen wurden, würde Beschäftigte wirksam vor hieraus resultierenden arbeitsrechtlichen Maßnahmen schützen.

## Ethische Aspekte

Loi, Michele (2020): People Analytics must benefit the people. An ethical analysis of data-driven algorithmic systems in human resources management. Berlin: AlgorithmWatch.

<https://algorithmwatch.org/project/auto-hr/gutachten-ethik-loi/>

Wie Michele Loi, Experte für ethische Aspekte des Einsatzes von Methoden Künstlicher Intelligenz der Universität Zürich, in seiner Studie darlegt, haben Beschäftigte einen Anspruch darauf, die Logik derartiger Systeme verstehen zu können. Denn das ist eine notwendige – wenn auch keine hinreichende – Voraussetzung dafür, dass die Autonomie der Beschäftigten gewahrt bleibt; sie dürfen keine passiven Objekte einer von Algorithmen getriebenen Steuerung („algorithmic governance“)

werden. Vielmehr müssen sie aktiv dazu beitragen können, die Leistung der Organisation zu verbessern, indem sie Erkenntnisse nutzen, die durch den Einsatz dieser algorithmischen Systeme entstehen. Letztendlich müssen die Systeme so gestaltet und eingesetzt werden, dass immer auch die Beschäftigten davon profitieren, nicht ausschließlich der Arbeitgeber. Es kann also nicht davon ausgegangen werden, dass Anbieter in der Lage sind, „ethische Algorithmen“ zu entwickeln, die Arbeitgeber dann per se unproblematisch einsetzen können. Im Gegenteil, es müssen alle beteiligten Akteure – die Anbieter algorithmischer Systeme, die Unternehmen, die sie einsetzen, und die Beschäftigten, die davon betroffen sind – bereit sein, gemeinsam Lösungen zu erarbeiten, die die verschiedenen Interessen berücksichtigen. Zu diesem Zweck müssen die Anbieter Informationen zur Verfügung stellen, die es Arbeitgebern und Beschäftigten erlauben zu verstehen, wie die Systeme funktionieren. Zugleich kann es in vielen Fällen gerechtfertigt sein, dass Firmen ihre Investitionen auch mithilfe von Betriebs- und Geschäftsgeheimnissen schützen.

Um auf eine ethisch akzeptable Art und Weise geeignete Modelle für „selbstlernende“ Systeme zu entwickeln, muss man sich vor allem mit der Erklärbarkeit und Fairness von algorithmischen Systemen befassen. Nachvollziehbarkeit und Fairness eines Modells können durch technische Methoden in erheblichem Maße verstanden, dokumentiert und verbessert werden. Das erforderliche technische Fachwissen sollte durch weniger formalisierte Erkenntnisse ergänzt werden, z. B. darüber, was im Kontext moralisch angemessen ist. Dieses Wissen muss in Zusammenarbeit mit Stakeholdern entwickelt werden, die möglicherweise eigens zu diesem Zweck eingebunden werden müssen.

Eine Herausforderung für Unternehmen besteht darin, ihre Fähigkeit zu verbessern, die Auswirkungen von KI-Entscheidungen auf bestimmte Gruppen oder die Daten, die indirekt diskriminierend sein können, zu verstehen. Es ist denkbar, dass ethnisch vielfältigere und geschlechtergerechtere Forschungsteams eine höhere Sensibilität für diese Fragen haben. Noch entscheidender ist, dass es kognitive Vielfalt geben sollte, insbesondere Fachleute, die in der Lage sind, „anders zu denken“ als die meisten Fachkundigen aus Informatik, Datenwissenschaft oder Software-Entwicklung (zu denen aber selbstverständlich auch Angehörige dieser Gruppen gehören können, die Ethik-Kompetenzen erworben haben). Die Forschung wird hoffentlich Erkenntnisse über die Wirksamkeit spezifischer Einstellungs- und Diversitätsstrategien für KI-basierte Systeme liefern.

Die wichtigsten ethischen Fragen, die vor der Implementierung der Datenerhebung beantwortet werden müssen, sind:

- Verfügen die Personalverantwortlichen in der Organisation über die technischen Fähigkeiten, die erforderlich sind, um die Nachvollziehbarkeit und Fairness von KI-Tools angemessen zu beurteilen?
- Sammeln Datenwissenschaftler\*innen genügend Informationen über das Training von KI-basierten Systemen, um eingreifen zu können, wenn sie nicht angemessen funktionieren? Werden die potenziellen Mängel und Grenzen der Trainingsdatensätze angemessen verstanden? Sind die technischen Schritte zur Bewertung und Dokumentation von Fairness und Verständlichkeit bereits vorhanden?
- Gibt es genügend kognitive und moralische/politische Vielfalt unter den Bewerter\*innen von KI-basierten Systemen, um Fragen der Fairness und Transparenz KI-basierter Systeme kritisch zu bewerten?
- Ist die Organisation angemessen ausgestattet, um Inputs und Feedback von externen Experten\*innen zu sammeln, die für die Beurteilung der Fairness und Verständlichkeit von KI-Tools relevant sind?

Praktische Empfehlungen zur Umsetzung ethischer Prozessverbesserungen sind:

- Stellen Sie sicher, dass Sie über ausreichende Kompetenzen verfügen, um KI-basierte Systeme ethisch zu entwickeln und umzusetzen. Organisationen, die KI-Tools im Personalmanagement entwickeln und implementieren wollen, müssen Expert\*innen mit der Bandbreite an Fähigkeiten (und informellem Wissen und kognitiven Ausprägungen) rekrutieren, die erforderlich sind, um Verständlichkeit und Fairness eines algorithmischen Systems zu bewerten.
- Beziehen Sie zivilgesellschaftliche Organisationen und Fachwissen aus der Wissenschaft ein. Validieren Sie mit externer Unterstützung, dass die technischen Methoden, die Sie zur Schulung Ihrer Modelle anwenden, wissenschaftlich fundiert und ethisch vertretbar sind. Entwickeln Sie ein Bewusstsein für mögliche ethische Kritik, indem Sie Feedback von außerhalb der Organisation sammeln und einwerben. Stellen Sie nach Möglichkeit (anonymisierte) Datensätze und Ihre Algorithmen zur Verfügung, um Ihr Tool unabhängig bewerten zu lassen.

Engagieren Sie sich in KI- und Ethik-Schulungen:

- Schulen Sie potenzielle Endnutzer\*innen (z. B. Personalmanager\*innen), um sicherzustellen, dass sie über das Know-how und die Fähigkeiten verfügen, die erforderlich sind, um KI-basierte Personalsoftware korrekt zu betreiben. Endnutzer\*innen sollten nicht blindes Vertrauen in KI-Instrumente haben, sondern ein angemessenes Maß an Vertrauen in Verbindung mit kritischen Einstellungen. Binden Sie

die potenziellen Anwender\*innen Ihres Tools ein, um sicherzustellen, dass sie ausreichend über das System, seine Grenzen und die richtige Anwendungsdomäne wissen. Fördern Sie Bildungsressourcen, die Personalmanager\*innen helfen, Missverständnisse über KI-basierte Personalsoftware zu vermeiden.

- Schulen Sie Beschäftigte, die potenziell von algorithmischen Systemen betroffen sind, oder ihre Vertreterinnen und Vertreter (z. B. Betriebsräte), um sicherzustellen, dass sie über das Know-how und die Fähigkeiten verfügen, die notwendig sind, um richtig zu verstehen, wie KI-basierte Personalsoftware eingesetzt wird. Sie sollten nicht blinden Glauben in KI-Instrumente haben, sondern ein angemessenes Maß an Vertrauen in Verbindung mit kritischen Einstellungen. Binden Sie potenziell betroffene Personen Ihres Tools oder deren Vertreterinnen und Vertreter ein, um sicherzustellen, dass sie genug über das Tool, seine Grenzen und seinen richtigen Anwendungsbereich wissen. Fördern Sie Bildungsressourcen für diese Personen, die ihnen helfen, Missverständnisse über KI-basierte Personalsoftware zu vermeiden. Auf welche ethischen Werte sich diese Leitlinie gründet, ist im Gutachten auf S. 48 ausgeführt.

Es sollte möglich sein, die tatsächliche Wirkung der Verwendung von KI-basierten Systemen zur Unterstützung von Personalentscheidungen zu überwachen und zu dokumentieren. Dies kann durch die Implementierung technischer Verfahren, z. B. das automatische Sammeln von Daten über Entscheidungen, die mithilfe der Inputs von KI-Systemen getroffen wurden, und durch die Implementierung sozialer Prozesse erreicht werden, z. B. indem Personalmanager\*innen die Möglichkeit bekommen, Feedback zu geben und die Ergebnisse der KI-Systeme auch mit ihren Entwickler\*innen zu diskutieren.

Die wichtigsten ethischen Fragen, die vor der Implementierung der Datenerhebung beantwortet werden müssen, sind:

- Haben wir Mechanismen um sicherzustellen, dass die Entscheidungen, die algorithmisch getroffen werden, bewertet und korrigiert werden können, wenn etwas schiefgeht?
- Sind unsere Verfahren zur Dokumentation der vom Algorithmus getroffenen Entscheidungen angemessen? Können wir sie so gestalten, dass sie die Privatsphäre der Beschäftigten nicht beeinträchtigen?
- Wie können wir einen Mechanismus für Feedback einrichten, der das Funktionieren des Personalmanagement-Systems nicht behindert, sondern es den Datenwissenschaftler\*innen ermöglicht, es Schritt für Schritt zu verbessern?

- Gibt es einen Mechanismus zur Wiedergutmachung von Personen, die aufgrund einer ungenauen algorithmischen Bewertung oder Vorhersage ungerecht behandelt wurden?

Praktische Empfehlungen zur Umsetzung ethischer Prozessverbesserungen sind:

- Entwickeln Sie einen datenschutzkonformen Mechanismus zur Aufzeichnung von Entscheidungen, die relevante Auswirkungen auf Beschäftigte haben und mithilfe algorithmischer Empfehlungen oder Vorhersagen getroffen werden.
- Entwickeln Sie einen datenschutzkonformen Mechanismus, um festzustellen, ob bestimmte Bevölkerungsgruppen (z. B. Nichtmuttersprachler\*innen, Schwangere, Angehörige von Minderheiten u. a.) durch algorithmische Entscheidungen negativ oder positiv betroffen sind.
- Entwickeln Sie einen Prozess, der es Angestellten möglich macht, Fehler in Daten-Inputs und Daten-Outputs zu korrigieren.
- Entwickeln Sie einen Prozess, der es Beschäftigten ermöglicht, Entscheidungen anzufechten, die vollautomatisiert getroffen wurden (falls vorhanden).
- Wenn eine Untergruppe Ihrer Angestellten (z. B. Nichtmuttersprachler\*innen, Eltern mit Kindern, Angehörige religiöser Gruppen) durch die Einführung algorithmischer Entscheidungen systematisch benachteiligt zu sein scheint, richten Sie ein Verfahren zur Wiedergutmachung für Sie ein und sorgen Sie dafür, dass die Ergebnisse für sie verbessert werden.

# Arbeitshilfen für die Praxis

## Ein Leitfaden für Betriebsräte und andere Personalvertretungen

Stiller, Sebastian/Jäger, Jule/Gießler, Sebastian (2020): Automatisierte Entscheidungen und Künstliche Intelligenz im Personalmanagement. Ein Leitfaden zur Überprüfung essenzieller Eigenschaften KI-basierter Systeme für Betriebsräte und andere Personalvertretungen. Berlin: AlgorithmWatch.

<https://algorithmwatch.org/project/auto-hr/leitfaden/>

Um hier eine angemessene Balance zu erreichen, hat AlgorithmWatch in Zusammenarbeit mit Sebastian Stiller, Professor für Mathematische Optimierung der Technischen Universität Braunschweig, einen Prüfkatalog entwickelt. Er ermöglicht es Betriebsräten und Personalmanager\*innen, von Anbieterfirmen Informationen zu erfragen, die essentiell dafür sind, die Funktionsweise, Qualitäts- und Zuverlässigkeitskriterien algorithmischer Systeme zu verstehen, ohne aber beispielsweise zu verlangen, dass Programmiercode oder die Datenbasis offengelegt werden müssen.

Angesichts der Komplexität der Verfahren wäre allerdings die Forderung verfehlt, sie Laien gegenüber auf einfache Art erklär- oder nachvollziehbar zu machen. Stattdessen müssen alle Akteure – Anbieter von Personalmanagementprodukten, Anwender (Personalabteilungen) und Betroffene (Beschäftigtenvertreter\*innen) – aktiv daran mitarbeiten, Systeme angemessen zu entwickeln und einzusetzen. Das bedeutet für die Vertreter\*innen der Beschäftigten, dass sie sich ein Grundverständnis von Methoden und Verfahren der KI-basierten Systeme aneignen oder Möglichkeiten schaffen müssen, auf diese Expertise zuzugreifen, etwa durch externe Sachverständige.

Um dieses Ziel zu erreichen wurde von Prof. Stiller und AlgorithmWatch ein Leitfaden entwickelt um dieses Grundverständnis herzustellen, offene Fragen zu beantworten und die in den rechtlichen und ethischen aufgeworfenen Forderungen an die Arbeitgeber durchzusetzen. Der Leitfaden ermöglicht es Betriebsräten und Personalmanager\*innen, von Anbieterfirmen Informationen zu erfragen, die essentiell dafür sind, die Funktionsweise algorithmischer Systeme zu verstehen und Qualitäts- und Zuverlässigkeitskriterien zu beurteilen, ohne dass Programmiercode oder die Datenbasis offengelegt werden müssen. Dies vermeidet langwierige Auseinandersetzungen über Geheimhaltungsfragen mit den Anbietern der Systeme. Auf die Auskünfte, die auf Basis des

Leitfadens gefordert werden, haben Beschäftigte einen ethischen und sehr wahrscheinlich auch einen gesetzlichen Anspruch

Die hier in einer Auflistung dargestellten Fragen haben unterschiedlichen Charakter. Sie reichen von allgemeinen Fragen wie „Welche Aussagen trifft und welche Entscheidungen berührt die Software?“ bis hin zu konkreten Fragen wie „Wie ist die Qualität der Implementierung (des Programmcodes) sichergestellt?“ In „Automatisierte Entscheidungen und Künstliche Intelligenz im Personalmanagement – Ein Leitfaden zur Überprüfung essenzieller Eigenschaften KI-basierter Systeme für Betriebsräte andere Personalvertretungen“ (Stiller et al. 2020) ist der Hintergrund der jeweiligen Fragen erläutert.

Wir werden diesen Leitfaden in Zusammenarbeit mit Vertreter\*innen von Beschäftigten und Personalmanagement in der Praxis testen und dabei vor allem Erläuterung dazu erarbeiten, wie die Antworten zu verstehen sind.

## Liste der Leitfragen

- Q1 Welche Aussagen trifft und welche Entscheidungen berührt die Software?
  - Q1.1 Um welche Software und welche Komponenten davon geht es?
  - Q1.2 Welche Aussage trifft die Software und mit welchem Wahrheitsgehalt?
  - Q1.3 In welchem Bereich bereitet das Softwaresystem Entscheidungen vor oder entscheidet autonom (selbstständig)?
- Q2 Wie kommt die Software zu ihren Aussagen?
  - Q2.1 Auf welche Daten hat die Software Zugriff?
  - Q2.2 Nach welchen Kriterien entscheidet die Software?
  - Q2.3 Wenn Maschinelles Lernen verwendet wird, welche Annahmen und wissenschaftlichen Theorien liegen dem Verfahren zugrunde und warum wurde dieses Verfahren gewählt?
  - Q2.4 Wenn Maschinelles Lernen verwendet wird, welche Trainingsdaten wurden verwendet?
  - Q2.5 Wenn Maschinelles Lernen verwendet wird, wie wurde das Verfahren gegen Diskriminierung und andere ungewollte Einflüsse aus den Trainingsdaten gesichert?
  - Q2.6 Wenn Maschinelles Lernen verwendet wird, wie wurde das Verfahren getestet?
- Q3 Wie ist die Qualität des Systems sichergestellt?
  - Q3.1 Setzt der Algorithmus die Kriterien exakt um?

- Q3.2 Wie ist die Qualität der Implementierung (des Programmcodes) sichergestellt?
- Q3.3 Wer hat die Software erstellt und welche Komponenten wurden übernommen?
- Q4 Wie ist das System im Betrieb integriert?
- Q4.1 Welche Fähigkeiten und Kenntnisse werden auf Seiten der Nutzer\*innen der Software benötigt?
- Q4.2 Wo liegt die Verantwortung für das Softwareprodukt im Unternehmen?
- Q4.3 Wie und von wem wird entschieden, welche Funktionalität der Software genutzt wird?
- Q4.4 Wer legt die Kennzahlen fest, anhand derer in der Software Ziele definiert werden?
- Q4.5 Wie transparent ist der Entscheidungsweg?
- Q4.6 Werden subtile Beeinflussungen der Nutzer\*innen durch die Gestaltung der Softwareoberfläche ausgeschlossen?
- Q4.7 Können automatische Entscheidungen korrigiert werden?
- Q4.8 Wurde eine Risikoabschätzung vorgenommen?

## Ein Puzzle zur Veranschaulichung von Machine Learning im Personalmanagement

Lorenz Matzat (2020): HR-Puzzle – ein Erklärspiel zu Machine Learning („Künstliche Intelligenz“) im Personalmanagement (HR/Human Resources). Berlin: AlgorithmWatch.

<https://algorithmwatch.org/hrpuzzle/>

Mit dem HR-Puzzle sollen zwei Dinge verdeutlicht werden: Zum einen das Grundprinzip der Funktionsweise von Maschinellem Lernen und neuronalen Netzen, zum anderen, wie ein Teil von Human Resources Software heutzutage funktioniert. Das HR-Puzzle ist Teil des Dossiers von AlgorithmWatch zur „Automatisierung im Personalmanagement und Mitbestimmung.“ In diesem Spiel muss ein neuronales Netz justiert werden. Innerhalb von drei Runden soll es helfen, unter 100 Mitarbeiter\*innen zu erkennen, wer befördert werden sollte, wer die Absicht hat zu kündigen und wen ein Unternehmen entlassen kann.

Das HR-Puzzle ist notwendigerweise eine Vereinfachung. Die Unternehmensdaten sind fiktiv, aber nicht willkürlich gewählt. Die Demographie der Belegschaft wurde anhand von Daten der Bundesagentur für Arbeit für mittelständische Betriebe des produzierenden Gewerbes ge-



staltet. Die Funktionsweise der Personalmanagement-Anwendung orientiert sich an existierenden Angeboten.

Neuronale Netze und Machine Learning („Künstliche Intelligenz“) sind meist wesentlich komplizierter, als im HR-Puzzle dargestellt. Wenn auch nur schematisch, zeigt das HR Puzzle das Prinzip, wie manche dieser Personalmanagement-Softwaresysteme funktionieren: Mittels besonderer statistischer Verfahren – gerne als „Künstliche Intelligenz“ bezeichnet – werden aus historischen Daten Vorhersagen über die Zukunft getroffen. Die Problematik dabei ist, dass diese Machine-Learning-Verfahren bestehende Ungleichheiten verstärken oder zumindest reproduzieren können: Weil sie auf historischen Daten basieren, „lernen“ sie darin enthaltenen althergebrachte Diskriminierungsmuster.

## Literatur

- Lewinski, Kai v./Barros Fritz, Raphael de/Biermeier, Katrin (2019): Bestehende und künftige Regelungen des Einsatzes von Algorithmen im HR-Bereich. Berlin: AlgorithmWatch. <https://algorithmwatch.org/project/auto-hr/rechtsgutachten-von-lewinski/> (Abruf am 24.8.2020).
- Loi, Michele (2020): People Analytics must benefit the people. An ethical analysis of data-driven algorithmic systems in human resources management. Berlin: AlgorithmWatch. <https://algorithmwatch.org/project/auto-hr/gutachten-ethik-loi/> (Abruf am 24.8.2020).
- Matzat, Lorenz (2020): HR-Puzzle – ein Erklärspiel zu Machine Learning („Künstliche Intelligenz“) im Personalmanagement (HR/Human Resources). Berlin: AlgorithmWatch. <https://algorithmwatch.org/hrpuzzle/> (Abruf am 24.08.2020).
- Stiller, Sebastian/Jäger, Jule/Gießler, Sebastian (2020): Automatisierte Entscheidungen und Künstliche Intelligenz im Personalmanagement. Ein Leitfaden zur Überprüfung essenzieller Eigenschaften KI-basierter Systeme für Betriebsräte und andere Personalvertretungen. Berlin: AlgorithmWatch. <https://algorithmwatch.org/project/auto-hr/leitfaden/> (Abruf am 24.8.2020).
- Wedde, Peter (2020): Automatisierung im Personalmanagement – arbeitsrechtliche Aspekte und Beschäftigtendatenschutz. Berlin: AlgorithmWatch. <https://algorithmwatch.org/project/auto-hr/gutachten-arbeitsrecht-datenschutz-wedde/> (Abruf am 24.8.2020).

## Autoren

**Sebastian Gießler** studierte Kulturanthropologie, Soziologie und Wissenschaftsphilosophie mit den fachlichen Schwerpunkten Philosophie der Sozialwissenschaften und Science and Technology Studies. Sein Forschungsschwerpunkt liegt darauf zu untersuchen, wie Algorithmen und „Big Data“ bei der Analyse und Bewertung sozialer Zusammenhänge eingesetzt werden.

**Matthias Spielkamp** ist Mitgründer und Geschäftsführer von Algorithm-Watch, das mit der Theodor-Heuss-Medaille ausgezeichnet und für einen Grimme Online Award nominiert wurde. Er war mehrfach Sachverständiger in Bundestagsanhörungen, u. a. zu KI in der Arbeitswelt oder KI und Robotik, und ist Mitglied der Expertengruppe zu „Responsible AI“ der Global Partnership on Artificial Intelligence. Er ist Vorstandsmitglied bei Reporter ohne Grenzen, Mitglied des Kuratoriums der Stiftung Warentest und im Beirat des Whistleblower Netzwerks. Spielkamp war Fellow der Zeit-Stiftung, der Stiftung Mercator und des American Council on Germany. Er hat Magister-Abschlüsse in Philosophie an der FU Berlin und in Journalismus an der University of Colorado erworben und ist Autor und Herausgeber von Büchern zu Internet Governance, der Zukunft des Journalismus und des Urheberrechts. Seine journalistischen Beiträge sind in MIT Technology Review, Die Zeit, brand eins und vielen anderen Publikationen erschienen.

---

Unternehmen, die KI-basierte „People-Analytics“-Systeme verwenden, handeln möglicherweise rechtswidrig. Denn solche Systeme dürfen nicht eingesetzt werden, ohne dass Beschäftigte individuell eingewilligt haben oder eine Betriebsvereinbarung vorliegt. Zudem erfüllen „Black-Box“-Systeme nicht die Auskunftsanforderungen des Betriebsverfassungsgesetzes. Autonomie und Mitsprache der Beschäftigten müssen auch dann gewährleistet sein, wenn algorithmische Systeme im Personalmanagement eingesetzt werden. Doch Beschäftigte müssen sich auch selber angemessene Kompetenzen erarbeiten, um diese Verfahren zu verstehen.

---