

Juni 2023

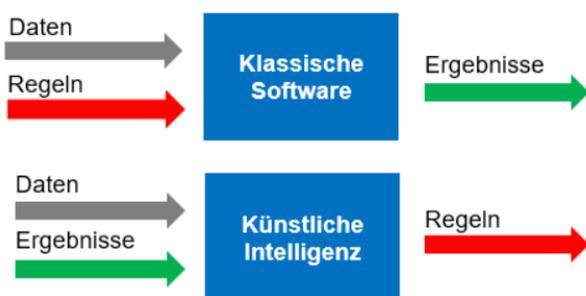
DISKRIMINIERUNG DURCH KI

Künstliche Intelligenz und Personalentwicklung

Künstliche Intelligenz (KI) hält Einzug im Recruiting, in der Personalentwicklung und im Personalmanagement. Vom ihrem Einsatz erhoffen sich Arbeitgeber effizientere Prozesse und geringere Kosten. Die Anbieter werben mit mehr Objektivität, Transparenz und Chancengleichheit bei Personalentscheidungen.

Aber: Auch KI kann diskriminieren. Der Einsatz von KI kann Ungleichbehandlung sogar verstärken: Wird ein diskriminierendes System firmenweit etabliert, sind mehr Menschen davon betroffen als z.B. durch einzelne vorurteilsbehaftete HR-Mitarbeiter*innen.

Man kann KI so gestalten, dass sie die Chancengleichheit vergrößert. Auf jeden Fall sollten sowohl die Software, als auch der Einsatz im Unternehmen von Expert*innen überprüft werden.



Wie kann KI diskriminieren?

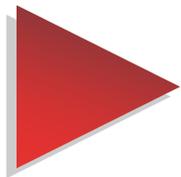
Beispiel Bewerbungsprozess: KI leitet mithilfe von statistischen Verfahren aus Lebensläufen (Daten) und den Lebensläufen von bewährten Mitarbeiter*innen (Ergebnisse) Regeln ab.

Anhand dieser (für uns nicht mehr nachvollziehbaren) Regeln lernt KI, welche (potentiellen) Beschäftigten eine gute Leistung erbringen werden. **Dabei gibt es viele Probleme:**

1. Die Definition von Erfolg. Die Definition dessen, was erfolgreiche Mitarbeiter*innen ausmacht, stammt häufig aus der Vergangenheit und aus einem ungerechten System. Es ist ausreichend nachgewiesen, dass z.B. Frauen anders bewertet werden als Männer.

2. Datenlücken und ungleiche Verteilung von Fehlern. Es gibt sehr viel weniger Daten über Angehörige von Minderheiten als Daten über die Mehrheit unserer Gesellschaft, weniger Daten über Frauen als über Männer. *Und in den meisten Betrieben deutlich weniger Daten über gewerbliche Arbeitnehmer als über Angestellte.* Falls z.B. die Daten von Frauen andere statistische Eigenschaften aufweisen als die von Männern, wird dies sehr häufig von der KI ignoriert. Hinzu kommt, dass Daten von Minderheiten (und Frauen) häufiger Fehler aufweisen. Sie alle haben in solchen Auswahlverfahren viel schlechtere Chancen, wenn man nicht aktiv gegensteuert.

3. Veränderung und ungenaue Konzepte. Ein zusätzliches Problem bei der Definition von Erfolg ist, dass es sehr viele sehr unterschiedliche Vorstellungen davon gibt. Dies kann zwischen Branchen variieren aber auch innerhalb eines Unternehmens zu unterschiedlichen Zeitpunkten. **Wenn ein KI-Anbieter das nicht beachtet, sondern das gleiche Modell an alle Kunden verkauft, passt es nur in Ausnahmefällen.** Hinzu kommt, dass sich viele Merkmale wie z.B. Gewissenhaftigkeit oder



Faktenblatt Diskriminierung durch KI

Kommunikationsstärke nur sehr schlecht in maschinenlesbare Sprache (Zahlen) übersetzen lassen. Aus diesem Grund arbeiten alle KI-Systeme mit sogenannten Stellvertretervariablen. Dies sind messbare Größen, von denen die KI-Entwickler*innen hoffen, dass sie irgendwas mit den gesuchten Variablen (Pragmatismus, Gewissenhaftigkeit) zu tun haben.

4. Ursache und Wirkung. KI-Systeme können sehr gut Muster in großen Datenmengen erkennen. Diese Muster stellen aber nicht zwangsläufig auch ein Ursache-Wirkungsverhältnis dar. KI versteht nicht, was diese Muster bedeuten, sie kann relevante nicht von irrelevanten Mustern unterscheiden. Zum Beispiel könnte eine KI erkennen, dass alle Bewerber*innen, die eine besonders gute Jobperformance gezeigt haben, eine Brille tragen, und in der Zukunft Brillenträger*innen besser bewerten

5. Unseriöse Methoden. Aufgrund der oben beschriebenen Probleme kommt es immer wieder zu höchst unwissenschaftlichen Methoden. Besonders KI-Systeme, die im Bereich Personalmanagement entwickelt wurden, fallen immer wieder durch unseriöse Annahmen auf. Ein Beispiel dafür ist der Einsatz von KI, um aus Videopräsentationen Persönlichkeitsmerkmale abzuleiten.

Fazit: Beim Einsatz von KI bei der Auswahl oder Entwicklung von Personal ist höchste Vorsicht geboten, weil es im Hinblick auf Chancengleichheit viele Fehlerquellen und damit eine hohe Wahrscheinlichkeit gibt, dass dieses Systeme Ungerechtigkeiten erzeugen oder verstärken. Potentielle Fehlerquellen liegen sowohl wohl in der Software als auch in den (betrieblichen) Trainingsdaten.

Mitbestimmung des Betriebsrates

Das Mitbestimmungsrecht des Betriebsrates, Einstellungsverfahren und andere personelle Maßnahmen transparent, diskriminierungsfrei und fair zu gestalten, ergibt sich aus § 95 und §§ 99, 102 BetrVG.

- ▶ Die Einbeziehung von externen Sachverständigen (§ 80 Abs. 3 BetrVG) ist bei dem Thema i.d.R. notwendig.
- ▶ KI-Anwendungen zur Bewertung, Auswahl oder Entwicklung von Personal (auch Bewerber*innen), stellen de facto Auswahlrichtlinien dar, die der Zustimmung des Betriebsrats bedürfen. (Auch, wenn die KI nicht die letzte Entscheidung trifft, denn sie trägt zur Entscheidung bei.)
- ▶ In Betrieben mit mehr als 500 Beschäftigten kann der Betriebsrat selbst die Aufstellung einer Auswahlrichtlinie verlangen
- ▶ Der Betriebsrat kann nach § 99 Abs. 2 Nr. 1 die Zustimmung zu einer personellen Einzelmaßnahme verweigern, wenn die personelle Maßnahme gegen die DSGVO verstößt, etwa gegen die umfassende Informationspflicht nach Art. 13 oder Art. 15 DSGVO.
- ▶ Geschäftsgeheimnis des KI-Anbieters ist kein Argument: Auch wenn ein Unternehmen mitbestimmungspflichtige Aufgaben nach außen verlagert, muss sichergestellt werden, dass die Mitbestimmung stattfindet.

Forderungen der IG Metall

Gerade bei Personalauswahl und –entwicklung kommt es auf größtmögliche Transparenz und Vertrauenswürdigkeit an. Die Arbeitgeber sollten ein Eigeninteresse daran haben, den Betriebsrat bei der Einführung von KI-Systemen zu beteiligen.

KI-Systeme die die Einstellung, Entwicklung oder Teilhabe von Beschäftigten gestalten oder zu deren Gestaltung beitragen, sind Hochrisiko-Systeme. Die Einbeziehung von – unabhängigen – Sachverständigen ist deshalb zwingend notwendig.

Chancengleichheit muss als Ziel solcher Systeme definiert und regelmäßig überprüft werden.