

Kersting, M. & Strohmeier, S. (2020). Künstliche Intelligenz, Kommentare zu verbreiteten Vorbehalten. Personalwirtschaft, Heft 10 (2020), S. 24-29.

Der Golem **wird kein Personaler**

Fünf Klischees, zwei Experten: Was Martin Kersting und Stefan Strohmeier, Universitätsprofessoren in Gießen respektive Saarbrücken, von verbreiteten Vorbehalten des Menschen gegenüber Künstlicher Intelligenz halten.

INTERVIEW: NICOLAS RICHTER

► So viel wird über Künstliche Intelligenz geredet, so selten präzise und kenntnisreich. Wen wundert's: Der Gegenstand der Diskussion ist komplex und unscharf. Kann sich die Wissenschaft kaum darauf einigen, was unter „Intelligenz“ zu verstehen ist, tut sich die Tech-Fachwelt bei der „künstlichen“ Variante nicht leichter. Gewiss aber scheint: KI wird enormen Einfluss auf die Arbeit der Zukunft nehmen, auch in HR.

Aber dafür braucht es eben Kenntnis: von den technisch-organisatorischen Chancen, aber auch den ethischen Risiken, über deren Ausmaß viele Anwender bis dato nur spekulieren können.

Wir haben zwei Experten mit sehr verschiedenem Zugang zum Thema um Statements zu einigen von uns wahrgenommenen Vorbehalten gegenüber KI gebeten: Martin Kersting ist Professor für Psychologische Diagnostik an der Uni Gießen, Mitglied im Ethikbeirat HR Tech und Verfechter eines menschenfreundlichen KI-Einsatzes. Stefan Strohmeier setzt sich als Professor für Management-Informationssysteme in Saarbrücken seit Jahren für die Nutzung von KI in HR ein. Beide äußern sich zunächst jeweils unabhängig voneinander zu den fünf Statements, bevor sie sich direkt aufeinander beziehen. ■



Martin Kersting leitet seit 2011 den Arbeitsbereich Psychologische Diagnostik an der Justus-Liebig-Universität Gießen. Früher war der Diplom-Psychologe unter anderem als Berater für Eignungsdiagnostik tätig, später verantwortete er mehrere Jahre lang die Führungskräfteentwicklung bei der Zollverwaltung in Münster.



Stefan Strohmeier führt den Lehrstuhl für Management-Informationssysteme an der Universität des Saarlandes. Der BWL-Professor ist Spezialist für digitales Personalmanagement und hat mehrere Bücher zum Thema verfasst. Seine nächste, für Mitte 2021 geplante Veröffentlichung beschäftigt sich mit KI-Technologien für HR.

„Im Kern steht ein moralisches Dilemma.“

Martin Kersting

STATEMENT



„Die Entscheidungen von Algorithmen sind nicht nachvollziehbar.“

→ Martin Kersting: Häufig liegen Algorithmen einfache Regeln zugrunde, die – anders als einige menschliche Entscheidungen – transparent sind. Es gibt aber auch komplexe und dynamische Entscheidungssysteme, zum Beispiel auf Basis maschinellen Lernens, die eine Art „Black Box“ darstellen. Solche Systeme kann man für Personalentscheidungen nur anwenden, wenn es gelingt, sich durch systematische **Input-Output-Analysen** ein Bild von der Wirkungsweise des Algorithmus zu verschaffen.

Die **Konsequenzen des Praxiseinsatzes** solcher Systeme müssen ebenso vorab untersucht werden wie die Frage, **wie stabil sie bleiben**, wenn es verschiedene Datensätze (etwa von Personen mit unterschiedlichen Merkmalen wie Alter, Geschlecht, Religion et cetera) zu verarbeiten gilt. Ein weiterführendes Stichwort ist hier „**explainable artificial intelligence**“.

→ Stefan Strohmeier: Die Forschung unterscheidet drei Gründe, warum es Algorithmen und ihren Ergebnissen an Nachvollziehbarkeit fehlen kann. Erstens: **Algorithmen werden mangels informatischer Kompetenz nicht verstanden** („illiterate opacity“). Zweitens: **Algorithmen werden vom Softwareanbieter etwa aus Wettbewerbsgründen geheim gehalten** („intentional opacity“). Drittens: **Sie sind tatsächlich sehr komplex und lassen sich objektiv schwer nachvollziehen** („intrinsic opacity“). Das unbeliebte, aber probate Gegenmittel ist, sich **informatische Kompetenz** anzueignen und so den ersten Grund zu beseitigen. Auf dieser Basis kann man die beiden weiteren Gründe angehen: Nämlich **Softwareanbieter bitten, ihre Algorithmen offenzulegen** – oder andernfalls deren Software nicht anzuwenden –, und komplexe Algorithmen wie etwa neuronale Netze Schritt für Schritt nachvollziehen.



■ **Anmerkung Kersting:** Im Kern steht ein „moralisches Dilemma“: Fundamentale Erkenntnisfortschritte sind insbesondere von hochkomplexen, letztlich nicht nachvollziehbaren Algorithmen zu erwarten. Unsere Verantwortung für Menschen spricht sowohl dafür, diese Erkenntnisse zugunsten der Menschen zu nutzen, als auch dafür, auf intransparente Systeme zu verzichten. Wir dürfen dieses Dilemma nicht vertuschen.



■ **Replik Strohmeier:** Nochmals: Jenseits absichtlicher Verschleierung ist jeder Algorithmus in Prozess und Ergebnis nachvollziehbar – zumindest dem Prinzip nach, oft sogar in allen Details. Statt sich einzureden, dass dem nicht so sei, ist es viel sinnvoller, wenn Anbieter weiter an „explainable AI“ arbeiten und Anwender ihre „illiterate opacity“ abbauen. Mit der KI-Kompetenz kommt dann auch die Transparenz.

STATEMENT



„Am Ende ist nicht der Algorithmus entscheidend, sondern der Mensch, der ihn programmiert. Solange in der Gesellschaft diskriminiert wird, gibt es keine diskriminierungsfreien Algorithmen.“

→ Strohmeier: Die gegenwärtig diskutierten Algorithmen stammen regelmäßig aus dem KI-Kernbereich des maschinellen Lernens. Ihre „Entscheidungen“ sind nicht in technologischem Sinn „programmiert“, sondern basieren auf Trainingsdaten zu „guten“ und „schlechten“ Entscheidungen. Damit ist nicht der Algorithmus entscheidend, sondern das menschliche Handeln zuvor und die diese Handlungen abbildenden Daten. Solche reinen Maschinenlern-Algorithmen sind an sich diskriminierungsfrei. Enthalten die eingegebenen Daten allerdings Diskriminierungen, können diese vom Algorithmus repliziert werden. Deshalb gibt es unter der Bezeichnung „Fairness by Design“ Bemühungen, auch und gerade dann technische Lösungen zu finden, wenn diskriminierende Trainingsdaten verwendet werden. Solche Algorithmen sind nicht nur diskriminierungsfrei, sondern sogar entdiskriminierend.

→ Kersting: Das Gute an der Debatte ist, dass sie zur Begriffsklärung beiträgt. Mit „Diskriminierung“ meint man zumeist die Benachteiligung von Personen oder Gruppen ohne sachlichen Grund. Computeranalysen unterscheiden zwei oder mehrere Personen oder Gruppen anhand von Merkmalen respektive Kriterien. Diese Kriterien können unsachgemäß sein. Der Vorteil des datengestützten Vorgehens ist, dass die Kriterien, transparente Algorithmen vorausgesetzt, expliziert und somit zur Diskussion gestellt werden. So – und nur so – kann man beurteilen, ob die Kriterien sachlich begründet und seitens der ganzen Organisation – statt nur seitens einzelner Mitglieder – gewünscht und rechtlich zulässig sind. Natürlich entscheidet der Mensch, aber beim datengestützten Vorgehen muss er seine Entscheidungskriterien offenlegen und rechtfertigen.



■ Anmerkung Strohmeier: Ich stimme zu: Konträr zu den Vermutungen der öffentlichen Diskussion haben Algorithmen das Potenzial, die Gleichbehandlung zu verbessern. Viele Maschinenlern-Algorithmen explizieren entscheidungsrelevante Kriterien, die bisher nur „in den Köpfen“ existierten. Das ermöglicht die ethische Überprüfung und Verbesserung von Entscheidungen.



■ Replik Kersting: Von den Maschinen werden vorurteilsfreie Entscheidungsvorschläge verlangt. Dabei legt die Arbeit mit den Maschinen offen, dass wir Menschen diese Maschinen mit Daten gefüttert haben, die aufgrund von Diskriminierung zustande kamen, und/oder dass wesentliche Daten fehlen. Was soll ich sagen? – Warum siehst Du den Splitter im Display Deines Computers, aber den Balken in Deinem Auge bemerkst Du nicht?

„Vielleicht verlieren Menschen die Kontrolle, aber das bedeutet nicht automatisch, dass Maschinen die Kontrolle übernehmen.“

Martin Kersting

„Diese Angst vor einem Kontrollverlust nährt sich wohl aus den teils erstaunlichen Leistungen der KI.“

Stefan Strohmeier



„Gerade in HR ist der Einsatz von KI bisher problematisch: **Abgesehen von der Diskriminierungsgefahr besitzen viele Personaler (noch) nicht das Know-How?, um KI zielführend und ethisch anzuwenden.**“

→ Kersting: Nicht der Algorithmus diskriminiert, sondern der Mensch. Möglicherweise stützen Menschen ihre Diskriminierung auf mithilfe von Algorithmen ausgewerteten Daten, die historische Diskriminierung widerspiegeln. Mindestens genauso wichtig wie die Algorithmen sind die Daten, auf die sie angewendet werden. Wir dürfen das Data Engineering und die Programmierung der Algorithmen weder den Disruptionspropheten und Nerds noch den Profiteuren überlassen. Vielmehr gilt es, von Anfang an Personalpraktiker, Interessensvertreter sowie Wissenschaftlerinnen und Wissenschaftler verschiedener Disziplinen einzubeziehen.

→ Strohmeier: Natürlich erweitert algorithmische Unterstützung das Anforderungsspektrum an Entscheider in der Personalarbeit. Sie brauchen methodisch-technische Kompetenzen, um algorithmische Modelle erstellen zu können, aber auch Kompetenzen zur Anwendung solcher Modelle. Dies impliziert die angesprochenen ethischen und rechtlichen Qualifikationen. Besonders kritisch ist hierbei die Fähigkeit, **menschliche und algorithmische Beiträge „auszutariieren“** und zu einer sinnvollen Gesamtentscheidung zusammenzuführen.



■ **Anmerkung Kersting:** Damit die Menschen weiterhin ihr Leben selbst bestimmen, bedarf es einer neuen Aufklärung. Wir müssen uns digitale Kompetenz aneignen, aber wir dürfen unser bisheriges Wissen und unsere ethischen Prinzipien nicht vergessen. Notwendiger als je zuvor ist außerdem die Zusammenarbeit verschiedener Professionen.



■ **Replik Strohmeier:** Nochmals nachdrückliche Zustimmung: Die Personalprofession braucht eine „digitale Aufklärung“. Das ist allerdings kein trivialer Vorgang, denn es bedeutet eine erhebliche Erweiterung ihres Qualifikationsspektrums. Wir müssen uns dringend damit auseinandersetzen, welches Wissen, welche Fähigkeiten und welche Kompetenzen HR im digitalen Bereich benötigt und wie dies vermittelt wird.

STATEMENT

4

„Der Datenschutz im Umgang mit KI ist für Unternehmen ein Riesenproblem. **Und zwar auch atmosphärisch – viele Mitarbeiter trauen Algorithmen einfach nicht.**“

→ Strohmeier: In der Tat setzt der Datenschutz Grenzen. Allerdings muss hier nach Phasen des KI-Einsatzes differenziert werden: **Die größeren Datenmengen, die während der Trainingsphase notwendig sind, kann man meist anonymisieren**, da identifizierende Variablen wie Personalnummer oder Name in der Regel keinen prädiktiven Wert aufweisen – sie sind für diese Phase gar nicht notwendig. **In der Anwendungsphase ist das anders.** Hier braucht man in der Regel personenbezogene Daten, um Vorhersagen über die betroffenen Personen machen zu können. **Gefühle oder tatsächliche Datenschutzprobleme sind dabei sicherlich ein Treiber mangelnder Mitarbeiterakzeptanz**; deutlich wichtiger scheint mir aber der insgesamt Mangel an Transparenz und „Erklärbarkeit“ von Algorithmen.

→ Kersting: Die Zweifel am Datenschutz sind berechtigterweise ein Stein des Anstoßes, es besteht aber die Chance, dass diese Zweifel auch zu einem Anstoß des Denkens führen. Es gibt viele gute Ansätze, **Big Data mittels Technikgestaltung und Voreinstellungen in zweckgebundene, datenschutzfreundliche Smart Data zu transformieren.** In jedem Fall müssen informierte, freiwillige Einwilligungen eingeholt werden. **Ohne Transparenz und Verständlichkeit lässt sich kein Vertrauen herstellen.**



■ **Anmerkung Strohmeier:** Die Datenschutzgrundverordnung setzt enge Grenzen, auch was die Anwendung von Algorithmen angeht. Insofern kann man Mitarbeitende durchaus beruhigen: Es besteht ein umfassender Schutz personenbezogener Daten, und wer diesen bei der Anwendung von KI nicht berücksichtigt, hat mit empfindlichen Strafen zu rechnen.



■ **Replik Kersting:** Menschen wollen informiert werden, welche Daten erhoben werden, wozu Daten verwendet werden und was die Organisation konkret für den Datenschutz tut. Sie möchten aktiv nach ihren Daten gefragt werden und die Kontrolle behalten. Das entspricht alles den Vorschriften, aber mit einem Verweis auf die regelkonforme Umsetzung der Verordnung allein gewinnt man kein Vertrauen. Man sollte so kommunizieren und gestalten, dass der Datenschutz erlebbar wird.

„Die Personalprofession braucht eine digitale Aufklärung.“

Stefan Strohmeier

„Nicht der Algorithmus diskriminiert, sondern der Mensch.“

Martin Kersting

STATEMENT

5

„Wenn wir den Einsatz von KI zu weit treiben und Maschinen immer intelligenter werden, **übernehmen sie irgendwann die Kontrolle über uns und die Welt.**“

→ Kersting: Diese Golem-Geschichte kann ich als Gute-Nacht-Lektüre empfehlen, im Arbeitsalltag bringt sie uns nicht weiter. **Maschinen haben keinen Willen.** Anthropomorphing, also Maschinen menschliche Eigenschaften zuzuschreiben, führt in die Irre. Vielleicht verlieren Menschen die Kontrolle, aber das bedeutet nicht automatisch, dass Maschinen die Kontrolle übernehmen. **Statt den drohenden Untergang zu beschwören, sollten wir für den Aufbruch werben.** Damit wir die Kontrolle behalten, müssen wir uns kümmern, uns (unter anderem in der Digitalkompetenz) weiterbilden, uns einmischen. **Mich stört die Polarisierung,** als ginge es um ein Entweder-oder, entweder der „intuitiv“ entscheidende Mensch oder die „reine“ KI. **Es liegt an uns, KI so zu programmieren, dass sie den Menschen unterstützt.** Unter anderem aus diesem Grund sind Initiativen wie der Ethikbeirat HR Tech bedeutsam.

→ Strohmeier: Diese Angst vor einem Kontrollverlust nährt sich wohl aus den teils erstaunlichen Leistungen der KI, etwa im Bereich autonomen Fahrens. **Sie beruht letztlich auf einer Verwechslung von schwacher mit starker KI.** Während sich schwache KI mit konkreten Anwendungsproblemen beschäftigt, geht es bei der starken darum, eine allgemeine Intelligenz zu schaffen, die dem Menschen gleicht (oder ihn womöglich übertrifft), also selbstständig und zielgerichtet handeln kann. **Es gibt derzeit keine starke KI, und die Diskussion darüber, ob dies jemals der Fall sein wird, hält an.** Aktuelle Anwendungen im Personalbereich sind dezidiert schwache KI. **KI wird daher die Personalarbeit nicht übernehmen.**



■ **Anmerkung Kersting:** Jetzt, zu Beginn einer Entwicklung, deren Verlauf wir nicht absehen können, müssen wir uns auf ethische Spielregeln verständigen. Wir müssen vorgeben, in welchen Bereichen wir uns zu welchem Zweck welche technische Unterstützung wünschen, wir müssen die Vorteile gegenüber den Nachteilen abwägen, wir müssen Grenzen setzen. Wir sollten den digitalen Fortschritt nicht aufhalten, aber gestalten.



■ **Replik Strohmeier:** „Gestalten“ ist die exakt richtige Strategie. Bei den „Spielregeln“ sollten wir allerdings darauf achten, keine doppelten Standards für Menschen hier und Maschinen da zu entwickeln. Wer etwa „vollumfängliche“ Transparenz von KI fordert, muss sich fragen lassen, wie transparent die Alternative menschlicher Entscheidungsfindung ist. Wenn wir bei der künftigen Gestaltung der KI die Latte zu hoch hängen, laufen wir Gefahr, Verbesserungspotenziale herzuschenken.

Personalwirtschaft

Das Magazin für den Job HR

10 2020

INKLUSIVE
SONDERHEFTE:
Arbeitsrecht
Zeitarbeit



KI – die große Unbekannte

WARUM DIE TECHNOLOGIE VIEL DISKUTIERT,
ABER SELTEN GENUTZT WIRD



192121 217507

10

+++ **Im Fokus:** Bürokonzepte der Zukunft +++ **Case Study:** BGM bei der AWO München +++
Events: Der HR Summit 2020 +++ **Special:** SAP & Partner +++ **Stilkritik:** Achtung, Zombies!