

# Was leistet KI für die Eignungs- diagnostik?



Der Einfluss von künstlicher Intelligenz (KI) nimmt heute in vielen Lebensbereichen spürbar zu. So verwundert es kaum, dass auch das Recruiting von dieser Entwicklung betroffen ist. Aber welche Auswirkungen sind durch den Einsatz von KI-Systemen zu erwarten? Werden sie die Eignungsdiagnostik schneller, treffsicherer und fairer machen?

**W**elchen Nutzen hat KI für die Eignungsdiagnostik? Um diese Frage zu beantworten, lohnt es sich, Bewerber und Unternehmen separat zu betrachten. Diese verfolgen nämlich unterschiedliche Ziele: Unternehmen möchten herausfinden, welcher Bewerber oder welche Bewerberin die Anforderungen einer Stelle am besten erfüllt. Hierzu bedienen sie sich verschiedener diagnostischer Möglichkeiten, wie zum Beispiel der Analyse von Bewerbungsunterlagen, Leistungs- und Persönlichkeitstests sowie Interviews und Arbeitsproben. Bewerber wiederum möchten attraktive Stellen besetzen und sich dabei gegen die Konkurrenz durchsetzen. Während leistungsstarke Bewerber dafür möglichst zutreffende Informationen über ihre wahren Kompetenzen vermitteln möchten, haben leistungsschwächere Kandidaten einen Anreiz, ihre Kompetenzen zu beschönigen (sogenanntes „Faking“). Beide Seiten – Unternehmen und Bewerber – werden sich in naher Zukunft immer mehr durch künstliche Intelligenz unterstützen lassen. Dies soll nicht nur die Zielerreichung verbessern, sondern auch den zeitlichen Aufwand und die Kosten senken.

## Wie Unternehmen KI nutzen

Beginnen wir mit der Arbeitgeberseite. Schon heute sind viele KI-Tools verfügbar, um HR-Prozesse zu beschleunigen und zu automatisieren: CV-Parser lesen automatisch Daten aus Lebensläufen aus und importieren sie in Bewerbermanagementsysteme. Chatbots liefern schnelle Antworten auf Fragen

von Bewerbern, ohne ihnen das Gefühl zu geben, mit einer Maschine zu sprechen.

KI-Systeme können allerdings auch in Auswahlentscheidungen einbezogen werden, zum Beispiel durch die *automatisierte Auswertung* von Videointerviews, Bewerbungsunterlagen und Social-Media-Profilen. Die Software zieht so auf Basis von Sprache, Texten, Fotos, Gestik und Mimik Rückschlüsse auf verschiedene Persönlichkeitsmerkmale. Beispielsweise wirbt der kanadische Anbieter Receptiviti damit, mittels KI nicht weniger als 200 unterschiedliche Persönlichkeitsmerkmale zu erfassen, und dies schneller als jeder Persönlichkeitstest. Die Anbieter von KI-Systemen versprechen Wettbewerbsvorteile durch schnellere Auswahlprozesse, weniger Diskriminierung und bessere Candidate Experience.

## Prognostische Validität fraglich

Sollte man also zuschlagen? Besser nicht, denn die Liste von Problemen ist lang. Es beginnt mit der *unklaren Qualität*. Die Qualität von Auswahlinstrumenten lässt sich mithilfe von Gütekriterien wie Objektivität, Reliabilität und Validität beschreiben. Vor allem die prognostische Validität, also die Vorhersagekraft des Verfahrens, ist wichtig und sollte möglichst unabhängig überprüfbar sein. Entsprechende Studien, Manuale und Gütekriterien findet man bei Anbietern von KI-Tools mit Verweis auf Geschäftsgeheimnisse jedoch selten. Stattdessen muss man sich mit euphorischen Erfahrungsberichten anderer Kunden begnügen. Für den Anwender ist das riskant, weil er im Vorfeld nicht beurteilen

kann, ob das gekaufte Verfahren die Qualitätsversprechen auch tatsächlich einlöst.

Vor allem bei KI-basierter Messung von *Persönlichkeitseigenschaften* ist Vorsicht geboten. Zwar ist es richtig, dass sprachliche und visuelle Parameter mit Persönlichkeitseigenschaften zusammenhängen, die Zusammenhänge sind aber erstens geringer, als wenn man sie mit klassischen Persönlichkeitstests erfassen würde. Zweitens ist meist nicht nachvollziehbar, mit welchen Daten das KI-System trainiert wurde. Reproduziert es valide gemessene Persönlichkeitsmerkmale oder Bauchentscheidungen menschlicher Beurteiler? Drittens wirken sich veränderliche Merkmale wie Tagesform (Müdigkeit, Krankheit) und Technik (Kamera, Mikrofon) des Kandidaten auf die untersuchten sprachlichen und visuellen Parameter aus und damit auch auf die Persönlichkeitsmessung.

## Wann Diskriminierung droht

Ein weiterer problematischer Aspekt KI-gestützter Auswahlentscheidungen ist mögliche *Diskriminierung*. Anbieter berufen sich gern darauf, dass die Software im Gegensatz zu menschlichen Entscheidern keine Vorurteile hat und deshalb fairer auswählen kann. Hier ist allerdings Vorsicht geboten, denn selbst wenn bestimmte Kategorien wie Alter und Geschlecht explizit ignoriert werden, kann es trotzdem zur sogenannten „Proxy-Diskriminierung“ kommen. Dabei diskriminiert die Software sozusagen unabsichtlich, indem sie Merkmale einbezieht, die indirekt mit dem Diskriminierungsmerkmal zusammenhängen.

So lässt sich zum Beispiel aus einem Vornamen auf die Ethnizität schließen oder aus einer Postleitzahl auf den sozioökonomischen Status. Forschungsergebnisse haben sogar gezeigt, dass KI-Software allein von Fotos mit erschreckender Treffsicherheit auf die sexuelle Orientierung (Wang & Kosinski, 2018) und politische Haltung (Kosinski, 2021) der abgebildeten Person schließen kann. Ob solche

Merkmale im Einzelfall berücksichtigt wurden, ist durch den Anwender kaum nachvollziehbar.

## Wenig Rechtssicherheit

Der Mangel an Transparenz führt insbesondere in Deutschland zu einem weiteren Problem, nämlich der *mangelnden Rechtssicherheit*. Voraussetzung für rechtssichere Auswahlentscheidungen sind Transparenz und Dokumentation: Welche Merkmale sind für die Stelle relevant? Mit welchen Verfahren werden sie erhoben? Wie tragen sie zur Entscheidungsfindung bei? Wer diese Fragen nicht beantworten kann, begibt sich auf dünnes Eis. Zum Beispiel müssen Arbeitgeber gemäß § 164 Sozialgesetzbuch IX nachweisen können, dass ein Bewerber nicht aufgrund einer Behinderung abgelehnt wurde. Auch das Allgemeine Gleichbehandlungsgesetz (AGG) verbietet eine Ablehnung aufgrund von Merkmalen wie Alter, Geschlecht, Weltanschauung und sexueller Identität. Im Fall einer Klage liegt die Beweislast meist beim Arbeitgeber, der aber selbst keinen Einblick in das Innenleben der KI hat. So werden abgelehnte Bewerber zur Klage geradezu eingeladen.

Beim Einsatz von KI-Systemen bestehen auch große *datenschutzrechtliche Hürden*. So verlangt die Datenschutzgrundverordnung (DSGVO) in der Regel die freiwillige Einwilligung der Betroffenen in die Verarbeitung (Art. 7, Abs. 4) und transparente Informationen über die involvierte Verarbeitungslogik (Art. 13, Abs. 2 lit. f). Ob eine Einwilligung im Auswahlkontext wirklich freiwillig ist und inwiefern sich die verwendete KI-Logik offenlegen lässt, ist fraglich. Es kommt erschwerend hinzu, dass die KI auch Daten verarbeiten kann, die zu den „besonderen Kategorien“ gemäß Art. 9 DSGVO zählen, was mit zusätzlichen Anforderungen an den Datenschutz einhergeht. Daher verwundert es kaum, dass KI-basierte Personalauswahl im geplanten europäischen Gesetz über künstliche Intelligenz (KI-VO) als „hochriskant“ eingestuft wird.

## „Es besteht die Gefahr, Bewerber mit KI-basierten Auswahlmethoden zu verschrecken. Einer Studie zufolge erleben Bewerber KI-gestützte Telefon- und Videointerviews als ‚creepy‘.“

### Beurteilung durch Menschen bevorzugt

Schlussendlich besteht die Gefahr, Bewerber mit KI-basierten Auswahlmethoden zu verschrecken. Einer Studie von Alina Köchling und Kollegen (2022) zufolge erleben Bewerber KI-gestützte Telefon- und Videointerviews als „creepy“. Auch Agata Mirowska (2020) konnte in einer Studie zeigen, dass Bewerber eine Beurteilung durch Menschen bevorzugen. Unter dieser Bedingung würden sie sich eher bewerben als bei rein KI-basierter Auswahl.

Die möglichen Ursachen sind vielfältig. Zum Beispiel kann bei Bewerbern der Eindruck entstehen, dass sie ihre Stärken und Persönlichkeit nicht zeigen und das Gespräch nicht steuern können. Die Arbeitgeberattraktivität könnte auch darunter leiden, dass eine KI-Auswahl als unpersönlich und kalt empfunden wird und sich dies gegebenenfalls auf das gesamte Unternehmen überträgt. Gerade in Zeiten von Bewerbermangel will eine solche Entscheidung daher wohlüberlegt sein.

### Auch Bewerber nutzen KI

Bewerber haben es dagegen deutlich leichter, künstliche Intelligenz zu ihrem Vorteil zu nutzen. Dies beginnt schon beim *Anschreiben*. Man braucht ChatGPT lediglich mit der Stellenbeschreibung und ein paar persönlichen Informationen zu füt-

tern, und schon erstellt es ein Anschreiben, das die gleichen Floskeln und Gemeinplätze enthält wie ein selbst verfasstes. Viele Unternehmen wissen um den geringen diagnostischen Nutzen des Anschreibens und verzichten daher schon länger darauf. Mit den aktuellen Entwicklungen sollte es endgültig obsolet werden. Ähnliche Instrumente, wie Motivationsschreiben und Fachaufsätze im Assessment-Center, wird dasselbe Schicksal ereilen.

Auch der *Lebenslauf* kann von der KI (um)geschrieben und an die Anforderungen der Stelle angepasst werden. Beispielsweise kann die Software bestimmte Erfahrungen oder Eigenschaften des Bewerbers oder der Bewerberin besonders hervorheben, andere hingegen unter den Tisch fallen lassen. Natürlich ist das Ergebnis sprachlich tadellos, sodass keinerlei Rückschlüsse auf die Sprachkompetenz der Person mehr möglich sind.

### Wie KI Tests beeinflussen kann

Ein besonders wichtiges Instrument sind *kognitive Testverfahren*, da diese die mit Abstand besten Vorhersagen beruflicher Leistung ermöglichen. Während solche Tests früher in Papierform vor Ort durchgeführt wurden, hat sich spätestens seit der Corona-Pandemie die unbeaufsichtigte Online-Durchführung etabliert. Damit ist auch das Problem von unerlaubten Hilfsmitteln gestiegen. Bisher waren davon nur bestimmte Aufgabentypen betroffen, bei denen man sich zum Beispiel mithilfe

## „Bewerber werden KI-Systeme zu ihrem Vorteil ausnutzen und damit die Aussagekraft vieler diagnostischer Instrumente verringern.“

von Taschenrechner und Wikipedia einen Vorteil verschaffen konnte.

Durch KI-Systeme wird allerdings eine neue Qualität erreicht, weil diese auch Aufgabentypen wie logische Schlussfolgerungen und Textverständnis bewältigen können. Die Vorhersagekraft unbefaufsichtigter Tests wird hierdurch sinken, sodass Unternehmen entweder das gesamte Testverfahren oder zumindest kurze Kontrolltests unter Aufsicht vor Ort durchführen müssen, um die Validität der Ergebnisse zu gewährleisten.

Neben kognitiven Tests werden auch *Persönlichkeitstests* und *Situational-Judgement-Tests* von der Entwicklung künstlicher Intelligenz beeinflusst. KI-Systeme können die Merkmale verstehen, die durch solche Tests gemessen werden, und lernen, welche Antworten für bestimmte Stellen als optimal gelten. Sie können somit das Antwortverhalten der Teilnehmer beeinflussen und Faking begünstigen. Die Einflussnahme durch die KI kann damit die Validität der Tests verringern, da die Antworten weniger authentisch oder schlicht gelogen sind.

### Fazit

Zusammenfassend fällt die Bestandsaufnahme nüchtern aus. Zwar werden KI-Lösungen im Personalbereich immer mehr Routineaufgaben übernehmen und so die operative Arbeit erleichtern und beschleunigen, der Eignungsdiagnostik könnte diese Entwicklung jedoch eher schaden als nutzen: Bewerber werden KI-Systeme zu ihrem Vorteil ausnutzen und damit die Aussagekraft vieler diagnostischer Instrumente verringern. Instrumente wie Anschrei-

ben beziehungsweise Motivationsschreiben werden überflüssig, Eignungstests müssen unter Aufsicht durchgeführt oder zumindest überprüft werden. KI-gestützte Auswahlentscheidungen sind mit vielen Problemen und Unsicherheiten behaftet: Meist fehlen Validitätsnachweise, die Qualität der Trainingsdaten ist oft nicht überprüfbar, Diskriminierung kann nicht ausgeschlossen werden, Entscheidungen sind intransparent und nicht rechtssicher. Von einem Einsatz kann daher nur abgeraten werden.

#### LITERATUR:

**Kosinski, M.** (2021). Facial Recognition technology can expose political orientation from naturalistic facial images. *Scientific Reports*, (11)1, 1–7.

**Köchling, A., Wehner, M. & Warkocz, J.** (2022). Can I show my skills? Affective responses to Artificial Intelligence in the recruitment process. *Review of Managerial Science*, 1–30.

**Mirowska, A.** (2020). AI evaluation in selection: Effects on application and pursuit Intentions. *Journal of Personnel Psychology*, (19)3, 142–49.

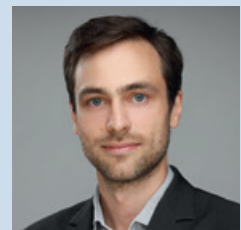
**Wang, Y. & Kosinski, M.** (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, (114)2, 246–257.

#### DER AUTOR:

##### **Dr. Benjamin Haarhaus**

Diplom-Psychologe und Geschäftsführer  
der Perseo GmbH, Düsseldorf.

[haarhaus@perseo.hr](mailto:haarhaus@perseo.hr)



# WIRTSCHAFTS- PSYCHOLOGIE aktuell

Sie wollen von **praktischen, wirtschaftspsychologischen Tipps** für Ihren Berufsalltag profitieren?

Dann melden Sie sich gleich für unseren **kostenlosen Newsletter** an.



QR-Code scannen,  
E-Mail-Adresse eintragen,  
Mailadresse bestätigen,  
fertig.

