

Ablehnung durch Selbstüberschätzung

Bei der Personalauswahl wird Algorithmen oft misstraut

Computergestützt berechnen Algorithmen menschliche Präferenzen oder bereiten sogar richterliche Entscheidungen vor. Trotz ihrer Leistungsfähigkeit und rechtlicher Vorgaben misstrauen viele Menschen den Programmen. Verhaltensökonomische Experimente zeigen für den Bereich der Personalauswahl: Neben der Angst von Bewerberinnen und Bewerbern vor Benachteiligung liegt die Ablehnung von Algorithmen oft in der Selbstüberschätzung des Managements begründet.

Dorothea Kübler

Algorithmen sind allgegenwärtig: Sie sagen unseren Buch-, Musik- und Filmgeschmack vorher, verteilen junge Menschen auf Schulen und Universitäten und schätzen ein, wie leistungsfähig Bewerberinnen und Bewerber für eine bestimmte Stelle sind. Auch im Bereich der öffentlichen Ordnung werden Algorithmen immer öfter eingesetzt: bei Entscheidungen, wer aus dem Gefängnis entlassen wird, wo und wann Polizeikontrollen stattfinden oder bei der Einschätzung der Kreditwürdigkeit. Algorithmen können sehr leistungsfähig sein, sodass ihre Prognosen der Wahrheit nahekommen. Gleichzeitig misstrauen ihnen viele Menschen zutiefst.

Und das, obwohl in Deutschland und international eine Reihe von rechtlichen Anforderungen an algorithmische Entscheidungen gestellt werden. Oft verwenden Algorithmen Geschlecht, Rasse und andere persönliche Informationen, um Verhalten vorherzusagen, was als „Profiling“ bezeichnet wird. Die Frage, ob die Profilerstellung akzeptabel ist und ob sie als Diskriminierung betrachtet

werden kann, ist eine offene Frage, da sich die Qualität der Vorhersagen des Algorithmus verschlechtern kann, wenn weniger Informationen verwendet werden. So verlangt Erwägungsgrund 71 der deutschen Datenschutz-Grundverordnung (DSGVO) von Datenverantwortlichen, diskriminierende Auswirkungen von Algorithmen zu verhindern, wenn sensible personenbezogene Daten verarbeitet werden. Artikel 13 und 14 der DSGVO besagen, dass Personen, deren Profil erstellt wird, das Recht auf aussagekräftige Informationen über die Logik der Informationsverarbeitung haben. Und eben hat das Europäische Parlament das Gesetz über digitale Dienste (Digital Services Act DSA) verabschiedet, das eine weitere Verschärfung der Anforderungen an die Offenlegung von Algorithmen vorsieht und ausdrücklich fordert, den Benutzern von digitalen Diensten eine Option ohne Profilerstellung anzubieten, sowie die Profilerstellung von Minderjährigen für Werbung vollständig verbietet.

Wodurch speist sich also das öffentliche Misstrauen gegenüber Algorithmen? Welche uner-

wünschten Nebeneffekte haben sie? Und wann werden sie akzeptiert? Diese Fragen haben wir im Zusammenhang mit der Bewertung von Bewerberinnen und Bewerbern untersucht. Viele Unternehmen greifen hier auf Algorithmen zurück, um eine erste Auswahl zu treffen. In die Studie haben wir die Perspektiven von BewerberInnen (im Folgenden männlich) und die des Managements (im Folgenden weiblich) einbezogen. Bewerber sind direkt von den Einstellungsentscheidungen betroffen. Wir erwarten daher, dass sie denjenigen Einstellungsprozess wählen, von dem sie glauben, dass er sie begünstigt. Von Managerinnen kann man dagegen erwarten, dass sie den Einstellungsprozess wählen, von dem sie glauben, dass er mit höherer Wahrscheinlichkeit zur Auswahl der besten Bewerber führt.

Für unsere Untersuchung haben wir eine Reihe von Laborexperimenten durchgeführt. Diese Methode ermöglicht es, die Entscheidungsumgebung genau zu kontrollieren; die Leistung der Bewerber lässt sich durch Aufgaben am Computer auf einfache Weise messen. Gleichzeitig erhalten die Teilnehmenden am Experiment vollständige und wahrheitsgemäße Informationen über die Funktionsweise des Algorithmus.

In unserer ersten Serie von Experimenten wollten wir herausfinden, ob Bewerber solche Algorithmen eher akzeptieren, die Informationen zum Geschlecht nicht verwenden dürfen und daher nicht in der Lage sind, aufgrund des Geschlechts zu diskriminieren. (Das schließt natürlich nicht aus, dass es indirekte Diskriminierung gibt – etwa, weil verwendete Eigenschaften mit dem Geschlecht korrelieren.) Eine zweite Versuchsserie sollte uns helfen, besser zu verstehen, warum einige Unternehmen bei der Verwendung von Einstellungsalgorithmen zögerlicher sind als andere. Eine mögliche Erklärung könnte die Selbstüberschätzung des Managements sein, was seine eigenen Fähigkeiten bei der Auswahl der Bewerber angeht. Eine dritte Experimentanordnung schließlich untersucht, ob die Transparenz des Algorithmus für Bewerber und Managerinnen von Bedeutung ist. Die Teilnehmenden erfahren in dieser Version des Experiments, wie der Algorithmus die vorhandenen Informationen über die Bewerber gewichtet, um eine Prognose zu machen.

In allen Experimenten nehmen die Teilnehmenden entweder die Rolle von Bewerbern



Dorothea Kübler ist Direktorin der Abteilung Verhalten auf Märkten am WZB und Professorin für Volkswirtschaftslehre an der Technischen Universität Berlin. dorothea.kuebler@wzb.eu

Foto: © WZB/David Ausserhofer, alle Rechte vorbehalten.

oder Managerinnen ein. Die Bewerber führen drei Aufgaben aus: Aufgabe 1, Aufgabe 2 und die Arbeitsaufgabe, die eine Kombination aus Aufgabe 1 und 2 darstellt. Für alle drei Aufgaben erhalten die Teilnehmenden in der Rolle der Bewerber Geld, aber für die Bezahlung des Managements ist nur die Arbeitsaufgabe relevant. Die Bewerber wählen zudem, ob eine Teilnehmerin in der Rolle der Managerin oder ein Algorithmus die Entscheidung zwischen ihm und einem anderen Bewerber trifft. In einer weiteren Versuchsanordnung müssen die Managerinnen entscheiden, ob sie die Entscheidungen an den Algorithmus delegieren oder selbst treffen wollen. In allen Experimentanordnungen gilt: Sowohl die Managerin als auch der Algorithmus kennen von den Bewerbern nur ihre Leistungen aus Aufgabe 1 und 2 sowie das Geschlecht. Sie kennen dagegen nicht die Leistung der Bewerber aus der Arbeitsaufgabe, auf die es für die Auszahlung der Geldprämie für die Teilnehmenden ankommt.

Bevor die Einstellungsentscheidungen getroffen werden, erhalten die Managerinnen und der Algorithmus Informationen über eine Auswahl von Bewerbern, inklusive ihrer Leistung aus der Arbeitsaufgabe. Diese Informationen reichen aus, um zu erkennen, dass weibliche Arbeitnehmer bei gleichen Leistungen in Aufgabe 1 und 2 männliche Arbeitnehmer bei der Arbeitsaufgabe übertreffen. Das könnte beispielsweise darauf zurückzuführen sein, dass Frauen mehr aus den beiden ersten Aufgaben lernen als Männer oder dass sie in unseren Experimenten schlichtweg ausdauernder sind. Infolgedessen bevorzugt der Algorithmus weiblich-

che gegenüber männlichen Bewerbern, wenn sie gleiche Leistungen bei Aufgabe 1 und 2 erzielt haben. Und wir beobachten, dass auch die Managerinnen in solchen Fällen häufig weibliche Arbeitskräfte bevorzugen. Allerdings ist der Anteil der richtigen Einstellungsentscheidungen bei Algorithmen deutlich höher als bei Managerinnen. Wir haben damit im Labor eine Situation geschaffen, in der die Managerinnen schlechtere Entscheidungen treffen als der Algorithmus, weil sie die vorhandenen Informationen weniger gut verarbeiten. Außerdem erkennen die Managerinnen nicht, dass der Algorithmus ihnen überlegen ist. Wir bezeichnen deswegen die Neigung der Managerin, die Einstellungsentscheidung selbst zu treffen, als „Aversion“ gegen Algorithmen.

In der Ausgangskonstellation des Experiments beobachten wir, dass etwas weniger als die Hälfte der Bewerber den Algorithmus für die Entscheidung über den Erfolg der Bewerbung bevorzugen. Je mehr Selbstvertrauen die Bewerber in ihre eigene Leistung haben, desto eher entscheiden sie sich für den Algorithmus. Sie vermuten also offenbar, dass der Algorithmus meritokratischer vorgeht als das Management. Wir stellen außerdem fest, dass die Entscheidungen zumindest in gewissem Maße von Eigeninteressen geleitet sind: Männliche Bewerber, die glauben, dass Managerinnen im Vergleich zum Algorithmus Bewerbungen von Männern bevorzugen, wählen eher die Managerin als Entscheidungsinstanz. Dasselbe trifft auf Bewerberinnen zu, die glauben, dass Managerinnen Bewerbungen von Frauen bevorzugen.

Für männliche und weibliche Bewerber gilt dagegen, dass sie eher den Algorithmus wählen, wenn sie wissen, dass dieser die Zugehörigkeit zu einem Geschlecht nicht zur Vorhersage der Arbeitsleistung heranzieht. Das bedeutet, dass das Diskriminierungsverbot in den existierenden gesetzlichen Regelungen die Akzeptanz von Algorithmen insgesamt erhöhen sollte. Von

den Teilnehmenden in der Rolle der Managerin delegieren nur etwa ein Drittel die Einstellungsentscheidungen an den Algorithmus – eine Bestätigung der Aversion gegen Algorithmen, denn der Algorithmus trifft im Durchschnitt die besseren Einstellungsentscheidungen als die Managerinnen. Wie erwartet, tendieren diejenigen Managerinnen dazu, die Entscheidung nicht an den Algorithmus zu delegieren, die glauben, selber viele richtige Einstellungsentscheidungen zu treffen. Sobald jedoch die Managerinnen über ihre wahren Fähigkeiten bei der Auswahl der besten Bewerber informiert werden, erhöht sich ihre Bereitschaft erheblich, die Aufgabe an den Algorithmus abzugeben. Dies gilt insbesondere für Managerinnen, die die Rückmeldung erhalten, dass sie deutlich weniger richtige Einstellungsentscheidungen getroffen haben, als sie dachten. Übermäßiges Selbstvertrauen im Management eines Unternehmens kann demnach ein wichtiges Hindernis für den Einsatz von Algorithmen sein.

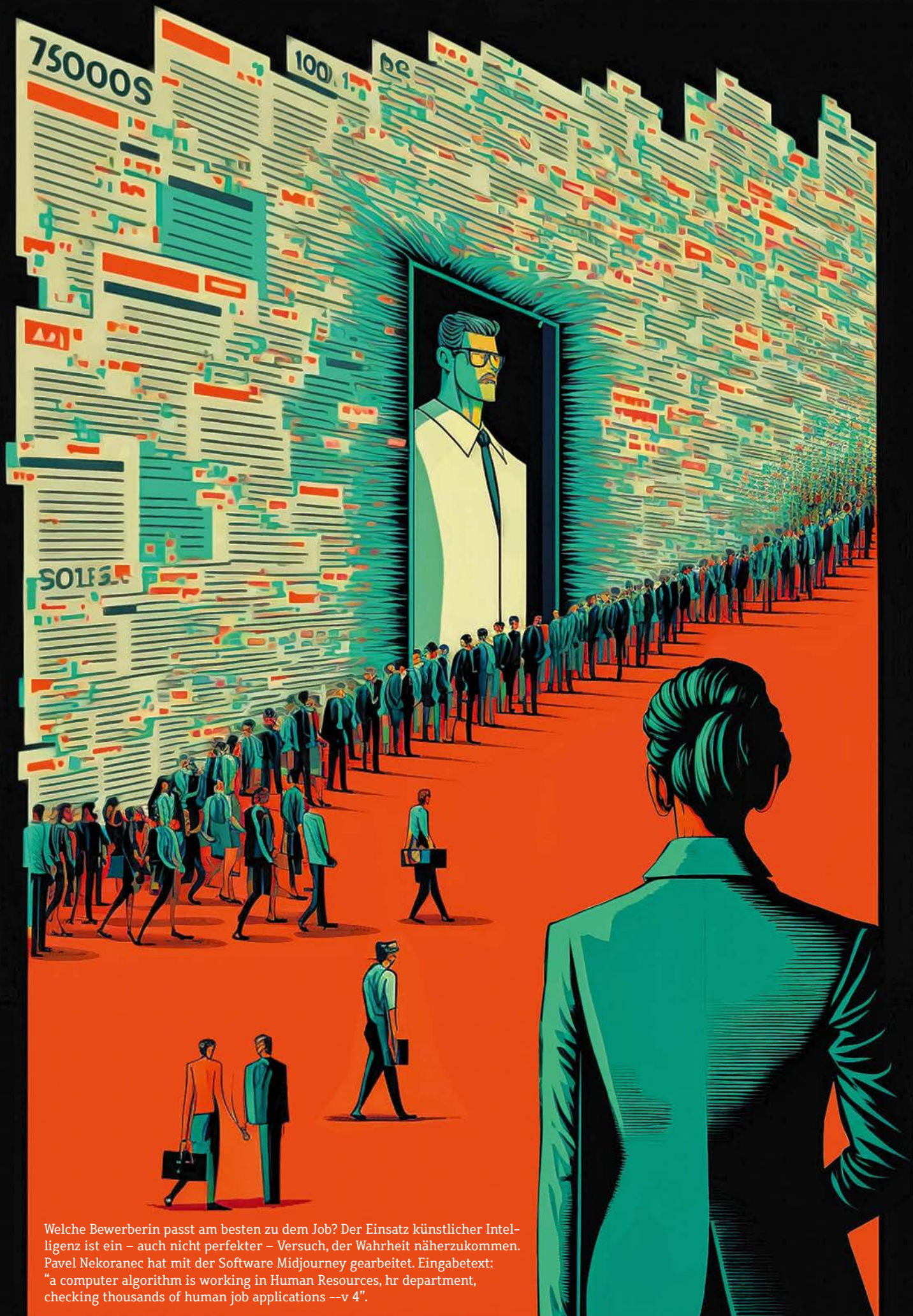
Transparenz, also die Bereitstellung von Informationen über die technische Funktionsweise des Algorithmus, hat dagegen keine Bedeutung für die Akzeptanz, weder aufseiten der Bewerber noch der Managerinnen. Transparenz ist jedoch aus anderen Gründen geboten – zum Beispiel um etwaige Verstöße gegen das Diskriminierungsverbot oder andere gesetzliche Regeln sichtbar zu machen.

Es gibt viele offene Fragen, zum Beispiel: Was wäre, wenn Managerinnen den Algorithmus zu Rate ziehen können, aber ihm nicht folgen müssen? Studien zeigen, dass Algorithmen, deren Empfehlungen von Menschen korrigiert werden können, attraktiver sind. In Zukunft wird es auf eine vernünftige Balance zwischen menschlichen Entscheidungen und algorithmischer Beratung ankommen – und darauf, die eigenen Entscheidungsfehler möglichst gut zu kennen.

Literatur

Dargnies, Marie-Pierre/Hakimov, Rustamdjan/Kübler, Dorothea: Aversion to Hiring Algorithms: Transparency, Gender Profiling, and Self-confidence. WZB Discussion Paper SP II 2022-202. Berlin: WZB 2022.

Metz, Cade/Satariano, Adam: „An Algorithm that Grants Freedom or Takes it Away“. In: The New York Times, 6.2.2020.



Welche Bewerberin passt am besten zu dem Job? Der Einsatz künstlicher Intelligenz ist ein – auch nicht perfekter – Versuch, der Wahrheit näherzukommen. Pavel Nekoranec hat mit der Software Midjourney gearbeitet. Eingabetext: "a computer algorithm is working in Human Resources, hr department, checking thousands of human job applications --v 4".