



Arbeitspapier

Diskriminierende KI? Risiken algorithmischer Entscheidungen in der Personalauswahl

November 2021

Arbeitspapier

Diskriminierende KI? Risiken algorithmischer Entscheidungen in der Personalauswahl

November 2021

INHALTSVERZEICHNIS

Editorial	3
Digitalisierung geschlechtergerecht gestalten	4
Wenn die Maschine zum Bewerbungsgespräch bittet	7
Vermeidung von Diskriminierung durch KI – Rechtliche Ankerpunkte und Ausblick auf die KI-Regulierung der EU	10
Impressum	15

Editorial

Auch in Deutschland kommen zunehmend KI-basierte Systeme für die Personalauswahl auf den Markt. Sie sollen eine möglichst effiziente, sichere und vorurteilsfreie Auswahl geeigneter Bewerber*innen ermöglichen. Die Praxis zeigt jedoch, dass entsprechende Systeme erhebliche Diskriminierungsrisiken bergen. [1] Diese Risiken werden von den einschlägigen Gesetzen, allem voran der Datenschutz-Grundverordnung (DSGVO) und dem Allgemeinen Gleichbehandlungsgesetz (AGG), nur unzureichend adressiert. Zu diesem Ergebnis kommt nicht nur die rechtswissenschaftliche Analyse des Forschungsprojekts „ExamAI – KI Testing & Auditing“ [2], sondern auch das Gutachten zum Dritten Gleichstellungsbericht der Bundesregierung [3].

Vor diesem Hintergrund stand die Frage nach den Diskriminierungsrisiken, die mit dem Einsatz von KI im Arbeitsbereich einhergehen können, im Zentrum unseres WebTalks „Diskriminierende KI? Risiken algorithmischer Entscheidungen in der Personalauswahl“ am 25. Mai 2021. Gemeinsam mit vier Expert*innen aus Informatik, Recht und Zivilgesellschaft diskutierten wir über die rechtlichen Leerstellen beim Diskriminierungsschutz, über KI als Werbeversprechen und über mögliche Ansätze für die Regulierung algorithmischer Systeme. Die Diskussionsrunde wurde aufgezeichnet und steht allen Interessierten zur Verfügung. [4]

Dieses Arbeitspapier stellt eine Zusammenfassung der vielfältigen Perspektiven unserer vier Diskutant*innen dar. Dr. Andrea Knaut (Geschäftsstelle für den Dritten Gleichstellungsbericht) plädiert in ihrem Beitrag dafür, dass Diskriminierungsfreiheit ein obligatorisches Ziel guter Systemgestaltung für KI sein sollte. Elisa Lindinger (Superr Lab) spricht sich dafür aus, die Entwicklung und den Einsatz von KI-Systemen interdisziplinär zu gestalten, um die gesellschaftlichen Folgen dieser Systeme erfassen zu können. In ihrem gemeinsamen Beitrag schlagen Dr. Andreas Sesing (Universität des Saarlandes) und Angela Tschech (MLU Halle-Wittenberg) eine Testpflicht für KI-Systeme vor, um Diskriminierung vorzubeugen. Außerdem untersuchen sie, inwiefern der Schutz vor Diskriminierung im Vorschlag der EU-Kommission zur Regulierung von KI-Systemen berücksichtigt wird.

Das Team des Projekts ExamAI wünscht Ihnen viel Freude und spannende Einsichten beim Lesen.

Pauline Junginger & Nikolas Becker

[1]

S. hierzu: Dastin, Jeffrey (11.10.2018): [Amazon scraps secret AI recruiting tool that showed bias against women.](#) (Abruf: 30.09.2021).

[2]

Borges, Georg; Hoffmann, Robert; Sesing, Andreas (2021): KI-Systeme im Personal- und Talentmanagement. Rechtsfragen im Überblick. Publikation befindet sich im Rahmen des Projekts ExamAI in Erscheinung.

[3]

Bundesregierung (2021): [Dritter Gleichstellungsbericht der Bundesregierung.](#) BT-Drucksache 19/30750, Berlin.

[4]

<https://testing-ai.gi.de/meldung/diskriminiert-durch-ki-gi-web-talk-diskutiert-rechtliche-aspekte>

Digitalisierung geschlechtergerecht gestalten

Von Dr. Andrea Knaut, Geschäftsstelle für den Dritten Gleichstellungsbericht der Bundesregierung

Die Schaffung diskriminierungsfreier digitaler Systeme ist eine Gestaltungsaufgabe für Staat, Wirtschaft und Wissenschaft.

„Digitalisierung geschlechtergerecht gestalten“ ist der programmatische Titel des aktuellen Gutachtens für den Dritten Gleichstellungsbericht der Bundesregierung. Gleichstellungsberichte stellen fest, wie es um die im Grundgesetz festgeschriebene Gleichstellung von Männern und Frauen in Deutschland bestellt ist. Die 2018 einberufene elfköpfige interdisziplinäre Sachverständigenkommission [5] beschäftigte sich für den aktuellen Bericht schwerpunktmäßig damit, wie auch in der Digitalisierung gleiche Verwirklichungschancen für alle Menschen unabhängig vom Geschlecht sichergestellt werden können. [6]

Im Gutachten wird die Digitalisierung aus einer soziotechnischen Perspektive betrachtet: Die Gestaltung eines algorithmischen Systems ist immer an einen sozialen Kontext gebunden. Dieser beeinflusst die Auswahl der zu verarbeitenden Daten, die Entscheidung für eine bestimmte Art ihrer mathematischen Interpretation sowie die Deutung ausgegebener neuer Daten. Dieser Kontext ist geprägt von den gesellschaftlichen Verhältnissen, in denen diejenigen leben, die Software- oder Hardwaresysteme beauftragen, entwerfen oder nutzen. Weltanschauungen, Menschenbilder und Stereotype fließen bewusst oder unbewusst in digitale Technik ein.

Im Anwendungsbereich algorithmischer Personalauswahlssysteme wird dies besonders deutlich. Ein häufig genanntes Beispiel ist der Fall des internen Amazon-Recruiting-Systems, das Männer aufgrund bestimmter Sprachmuster im Lebenslauf bevorzugte. [7] Sprachmuster und Geschlecht hängen stark zusammen: Männer bewerben z. B. alle möglichen Fähigkeiten häufiger wortreich, während Frauen oft nur Qualifikationen angeben, die sie auf hohem Niveau haben. Die Geschlechterdiskriminierung steckt schon in den Daten und verhindert eine korrekte Bewertung der Fähigkeiten. [8] Verwenden Entwickler*innen eines Personalauswahlsystems aus Lebensläufen extrahierte Da-

[5]

<https://www.dritter-gleichstellungsbericht.de/de/to-pic/44.sachverst%C3%A4ndigenkommission3.html>
(Abruf: 21.07.2021).

[6]

Bundesregierung (2021):
Dritter Gleichstellungsbericht der Bundesregierung.
BT-Drucksache 19/30750,
Berlin. (Abruf: 21.07.2021).

[7]

Dastin, Jeffrey (2018): [Amazon scraps secret AI recruiting tool that showed bias against women](#). In: Reuters, 11.10.2018. (Abruf: 21.07.2021).

[8]

Jersin zit. n. Wall, Sheridan/ Schellmann, Hilke (2021): [AI was biased. The company's solution? More AI](#). In: MIT Technology Review, 23.06.2021. (Abruf: 21.07.2021).

ten zum Training von Machine-Learning-Modellen – zumal in einer männerdominierten Branche –, wird dieses System Geschlechterstereotype verstärken. Fatourou et al. erläutern systematisch, welche Formen von Stereotypen sich auf diese und ähnliche Weise in algorithmische Systeme einschleichen. [9]

Vorurteilsfreie algorithmische Systeme für die Personalauswahl zu schaffen, stellt Entwicklungsteams vor hohe Anforderungen. Gender- und Diversity-Kompetenz ist gefragt, um das Problem zu verstehen oder gar formalisieren zu können. Bisher sind die Anreize und Verpflichtungen gering, Diskriminierungsfreiheit sicherzustellen. Bspw. ist seit mehr als zwei Jahren belegt, dass Facebook Stellenanzeigen entlang von Geschlechterstereotypen verteilt (Gender Targeting), verändert hat sich nichts. [10]

Geschlechtergerechte Gestaltung digitaler Systeme bedeutet, politische und in Industrienormen kodifizierte Rahmenbedingungen zu setzen, die Diskriminierungsfreiheit algorithmischer Systeme zu einem obligatorischen Ziel guter Systemgestaltung machen.

Im Gutachten für den Dritten Gleichstellungsbericht empfiehlt die Sachverständigenkommission, bei algorithmischen Personalauswahlssystemen zwingend Datenschutzfolgeabschätzungen vorzunehmen und hierbei Diskriminierungsrisiken zu berücksichtigen. Außerdem soll das Allgemeine Gleichbehandlungsgesetz so erweitert werden, dass Arbeitgebende verpflichtet sind offenzulegen, wenn ein algorithmisches System im Rahmen personeller Maßnahmen zum Einsatz kommt. Es werden zudem u. a. Präzisierungen der Datenschutzgrundverordnung (DSGVO) bezüglich der Verarbeitung geschlechtsbezogener Daten, ein DSGVO-konformes Arbeitsrecht, Verbandsklagerecht bspw. für Antidiskriminierungsstellen und -verbände sowie Schulungen von Betriebs- und Personalräten gefordert. [11]

Mit Verweis auf den risiko-adaptierten Ansatz der Datenethikkommission (2019) wird im Gutachten zudem vorgeschlagen, Softwaresysteme einer unabhängigen Risikoprüfung hinsichtlich Diskriminierung zu unterziehen. [12] Werden Personalauswahl-Technologien bspw. als Hochrisikotechnologien eingestuft, wie im KI-Regulierungsentwurf der EU-Kommission vorgesehen [13], würde dies neben Risikoprüfungen zahlreiche Maßnahmen zur Schadensminimierung nach sich ziehen: u. a. gesteigerte Transparenz- und Auskunftspflichten für Betroffene, Veröffentlichung der in die Berechnung einfließenden Faktoren, ihrer Gewichtung, der Datengrundlage und des Entscheidungsmodells in nachvollziehbarer Form. Die Sachverständigenkommission für den Gleichstellungsbericht hat für algorithmische Systeme im Bereich der Personalauswahl eine Einstufung in die höchste Kritikalitätsstufe und damit auch deren Verbot nicht ausgeschlossen.

[9]

Fatourou, Panagiota/Han-kin, Chris/Knowles, Brian: Gender Bias in Automated Decision Making Systems. Policy Paper, [ACM Europe Technology Policy Committee](#), (Abruf: 21.07.2021).

[10]

Imana, Basileal/Korolova, Aleksandra/Heideman, John (2021): [Auditing for Discrimination in Algorithms Delivering Job Ads](#). In: Proceedings of the Web Conference 2021 (WWW'21), S. 3767–3778. (Abruf: 21.07.2021).

[11]

Bundesregierung (2021): [Dritter Gleichstellungsbericht der Bundesregierung](#). BT-Drucksache 19/30750, Berlin. (Abruf: 21.07.2021).

[12]

[Gutachten der Datenethikkommission der Bundesregierung](#), Berlin. (Abruf: 21.07.2021)

[13]

Europäische Kommission: [Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz](#) (Gesetz über Künstliche Intelligenz) und zur Änderung bestimmter Rechtsakte der Union, COM (2021) 206 final. (Abruf: 21.07.2021).

Risikoprüfungen oder Auditierungen sind sehr voraussetzungsvoll und prüfen konkrete Anwendungen in konkreten Szenarien. Sie verlangen den Entwicklungsteams ab, Diskriminierungsfreiheit als prüfbares Kriterium in ihrer Software zu konzeptualisieren. [14] Doch wie sich oben bereits angedeutet hat, lässt sich Geschlecht nicht einfach heraus- oder hineinrechnen.

[14]

Wilson, Christo/Ghosh, Avijit/Jiang, Shan/Mislove, Alan/Baker, Lewis/Szary, Janelle/Trindel, Kelly/Polli, Frida (2021): Building and Auditing Fair Algorithms: A Case Study in Candidate Screening. In: FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 666–667. (Abruf: 21.07.2021).



Dr. Andrea Knaut ist Informatikerin und arbeitet derzeit als wissenschaftliche Mitarbeiterin in der Geschäftsstelle für den Dritten Gleichstellungsbericht der Bundesregierung. Außerdem leitet sie die Fachgruppe „Internet und Gesellschaft“ in der Gesellschaft für Informatik (GI) und hat in den Wissenschaftsjahren 2018 und 2019 das mobile Bildungsprojekt Turing-Bus für die GI geleitet.

und 2019 das mobile Bildungsprojekt Turing-Bus für die GI geleitet.

Wenn die Maschine zum Bewerbungsgespräch bittet

Von Elisa Lindinger, Superrrr Lab [15]

[15]

<https://superrrr.net/>

<https://elisalindinger.de/>

<https://twitter.com/elizabOt>

Immer mehr Softwareprodukte im Personal- und Recruitingbereich werben mit dem Einsatz von „Künstlicher Intelligenz“. Doch welche Auswirkungen sie auf Chancengerechtigkeit am Arbeitsmarkt haben, ist zu wenig untersucht.

Im Trend: Künstliche Intelligenz im Recruiting-Bereich

Verfahren der Künstlichen Intelligenz (KI) sind im Recruitingbereich zwar noch nicht weit verbreitet; bisher kommen sie vor allem bei großen Unternehmen zum Einsatz. Doch zu Recht werden sie schon heute diskutiert und analysiert: Ihr möglicher Impact auf eine gerechte Arbeitswelt ist enorm, und ihr Ruf, gesellschaftliche Ungerechtigkeiten zu reproduzieren, eilt ihnen voraus. [16]

[16]

Dieser Beitrag basiert auf der Expertise „AI Powered Recruiting? Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst“ von Lorena Jaume-Palasi, Elisa Lindinger, Julia Kloiber

Schon heute sind zahlreiche Produkte auf dem Markt, die mit dem Schlagwort „KI“ werben. Ihre Spannweite ist enorm: Gemeint sind damit Plugins für die Plattform LinkedIn, die Suchergebnisse für Recruiter*innen optimieren sollen, oder Chatbots, die z.B. auf Firmenwebseiten eingebettet werden und potenzielle Bewerber*innen ansprechen sollen, oder die sie durch den Bewerbungsprozess begleiten und z.B. bei der Terminvereinbarung assistieren.

Daneben gibt es auch Produkte, die Stellenausschreibungen analysieren und alternative, genderinklusive Formulierungen vorschlagen. Solche Werkzeuge helfen also sogar, mögliche sprachliche Diskriminierung zu erkennen und zu vermeiden. Andere Werkzeuge aggregieren Daten über Bewerber*innen von Social Media, Datenbanken oder Nachrichtenwebseiten. In Ländern, in denen das erlaubt ist, greifen sie sogar auf Vorstrafenregister oder Gesundheitsdaten von Versicherungen zurück.

Komplexe Produktsuiten dienen dazu, Bewerbungen oder sogar videobasierte Bewerbungsgespräche daraufhin zu analysieren, ob die Person geeignet ist – und ob sie zum Arbeitgeber passt. Dabei werden physische Aktionen und Reaktionen von Bewerber*innen klassifiziert und Eigenschaften wie Selbstbewusstsein, Teamfähigkeit, Durchsetzungsfähigkeit oder Aufrichtigkeit zugeordnet. Diese Verfahren entbehren wissenschaftlichen Grundlagen und bergen vor allem in interkulturellen Kontexten, die die physischen Reaktionen beeinflussen können, enormes Potenzial für Diskriminierung.

Wo „KI“ draufsteht, ist nicht immer Maschinelles Lernen drin

Doch welche Algorithmen und Modelle bei den Softwareprodukten eingesetzt werden, ist so gut wie nicht nachvollziehbar – die angebotenen Produkte sind nicht Open Source, welche Verfahren des Maschinellen Lernens sie anwenden, bleibt unklar. So bleibt der Verdacht, dass einige Produkten sich vor allem aus Marketinggründen mit dem Label „KI“ schmücken. Eine Analyse einzelner Produkte wäre zwar prinzipiell möglich, jedoch ist der Markt derzeit so schnelllebig, dass ein Werkzeug wahrscheinlich nicht mehr angeboten würde, wenn es ausgewertet ist. Die Vielfalt der Anwendungen und die übertriebenen Versprechen machen es daher schwierig, pauschale Aussagen darüber zu treffen, inwieweit KI heute bereits im Recruiting eingesetzt wird – und ob sie diskriminierende Strukturen in der Arbeitswelt verstärkt oder nicht.

Der Bias ist nicht nur ein Problem der Daten

Betrachtet man Diskriminierung bei Einstellungsverfahren, wird es schnell komplex. Diskriminierung erfolgt nicht basierend entweder auf Gender oder auf Alter, entweder auf Klasse oder auf Herkunft. All diese Faktoren bilden Interferenzmuster, die sich je nach Situation ganz unterschiedlich auswirken können. Solche Interferenzen in Diskriminierungsmustern – oder intersektionale Diskriminierung – sind statistisch nicht einfach abzubilden, was eine Analyse von KI-Verfahren im Hinblick auf ihr Diskriminierungspotenzial erschwert. Umso wichtiger ist es, KI-Softwareprodukte im Recruitingbereich mit hohen Transparenz- und Sorgfaltspflichten zu belegen, wie in der KI-Verordnung der EU vorgesehen ist. Diese Sorgfaltspflichten dürfen sich nicht auf die technische Entwicklung der Softwareprodukte oder die zugrundeliegenden Datensätze beschränken. Denn als soziotechnische Systeme unterliegen diese Produkte vielen Faktoren, die ihre Wirksamkeit beeinflussen, zwischen denen jedoch kein Wissenstransfer besteht: Datenerhebung, Modellierung, Implementierung, Akquise von Softwareprodukten und ihr tatsächlicher Einsatz erfolgen in ganz unterschiedlichen Kontexten, und durch Menschen, die zwar über die jeweilige Expertise verfügen mögen, aber kaum die technologischen, soziologischen und rechtlichen Folgen allesamt überblicken können.

Interdisziplinarität als notwendige Voraussetzung

Diese Komplexität macht es notwendig, dass der Einsatz von KI-Softwareprodukten im Recruitingbereich interdisziplinär diskutiert, hinterfragt und – sollte ein Einsatz notwendig und sinnvoll sein – kontinuierlich begleitet wird. Wissenstransfer und disziplinübergreifende Zusammenarbeit ist die Grundvoraussetzung dafür, dass nicht schon bald Technologien im großen Stil eingesetzt werden, die die Chancengerechtigkeit in der Arbeitswelt massiv negativ beeinflussen. Die vielbeschworene digital literacy – also ein Wissenstransfer aus technischen Disziplinen in die Breite – allein ist dafür nicht die Lösung; auch soziologisches und rechtliches Wissen muss schon in den Schaffungsprozess von Software mit einfließen, z.B. in Form von Technik- und Datenfolgenabschätzungen.



Elisa Lindinger ist an der Schnittstelle von Technologie, Kunst und Geisteswissenschaften aktiv, sowohl innerhalb als auch außerhalb der Wissenschaft. Sie ist Mitbegründerin von Superrr Lab, einem in Berlin ansässigen femi-

nistischen Labor, das sich dem Entwurf diverser und gerechter Zukünfte widmet – im Tech-Bereich und darüber hinaus. Ihre Forschung konzentriert sich auf offene digitale Infrastrukturcommunities, die digitale Zivilgesellschaft und die sozialen Auswirkungen neuer Technologien.

Vermeidung von Diskriminierung durch KI – Rechtliche Ankerpunkte und Ausblick auf die KI-Regulierung der EU*

Von Dr. Andreas Sasing [17], Universität des Saarlandes und Angela Tschech [18], Martin-Luther-Universität Halle-Wittenberg

Mit KI sind Hoffnungen ebenso verbunden wie die Angst vor „Diskriminierung“. Dabei ist die Einhegung sozial unerwünschter Risiken eine gesamtgesellschaftliche Aufgabe – und das Testen von KI ein vielversprechendes Instrument hierzu, was voraussichtlich künftig auch rechtlich festgeschrieben wird.

Status quo: Drohende Diskriminierung durch KI im Beschäftigungskontext

Der Einsatz von KI und intelligenten Systemen im Beschäftigungskontext nimmt stetig zu. Der unreflektierte Einsatz solcher Systeme kann jedoch zu erheblichen Risiken für die Verwirklichung gleichberechtigter Teilhabe am Erwerbsleben führen. [19] Diese Risiken für Freiheits- und Gleichheitsrechte haben die bisherige rechtspolitische Diskussion über die Regulierung von intelligenten Systemen auf europäischer Ebene entscheidend geprägt. [20] Der nachfolgende Überblick zeigt mögliche rechtliche Anknüpfungspunkte auf und gibt einen knappen Ausblick auf die bevorstehende KI-Regulierung auf der Ebene der Europäischen Union.

Rechtliche Ankerpunkte: AGG und Transparenzpflichten

Rechtlichen Schutz vor Diskriminierung durch KI bietet in Deutschland im Wesentlichen das Allgemeine Gleichbehandlungsgesetz (AGG). Das Gesetz soll die ungerechtfertigte Ungleichbehandlung (vulgo: „Diskriminierung“) wegen sog. „verpöner Merkmale“ verhindern. Zu den (zugeschriebenen) Merkmalen, wegen derer Bewerber*innen für ein Beschäftigungsverhältnis nicht benachteiligt werden dürfen, zählen die „Rasse“ [21]

*

Alle digital verfügbaren Quellen wurden zuletzt am 2.7.2021 abgerufen.

[17]

E-Mail: andreas.sasing@uni-saarland.de

[18]

E-Mail: angela.tschech@jura.uni-halle.de

[19]

Fröhlich/Spiecker gen. Döhm, Können Algorithmen diskriminieren?, <https://verfassungsblog.de/koennen-algorithmen-diskriminieren> Orwat, Diskriminierungsrisiken durch Verwendung von Algorithmen, Hg.: Antidiskriminierungsstelle des Bundes, Berlin 2019; Bundesregierung (2021): Dritter Gleichstellungsbericht der Bundesregierung, BT-Drucksache 19/30750, Berlin.

[20]

Siehe nur Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, <https://rm.coe.int/09000016809e1154>.

[21]

Der Gesetzgeber geht in § 1 AGG nicht von einer tatsächlichen Existenz menschlicher „Rassen“ aus, vielmehr ist damit die Diskriminierung wegen einer rassistischen Zuschreibung gemeint. Dazu ausführlicher: BUG e.V.

und die ethnische Herkunft, das Geschlecht, Religion und Weltanschauung, etwaige Behinderungen, das Alter sowie die sexuelle Identität (vgl. § 1 AGG).

(Rechts-)Folge einer unzulässigen Ungleichbehandlung ist ein Anspruch auf Schadensersatz bzw. Entschädigung der Betroffenen (§ 15 AGG). Dieser setzt dreierlei voraus: Eine Ungleichbehandlung, die Kausalität eines der „verpönten“ Merkmale sowie das Fehlen einer Rechtfertigung der Ungleichbehandlung. Dabei tragen im Grundsatz die Benachteiligten die Beweislast für das Vorliegen dieser Voraussetzungen. Dies ist bereits jenseits des Einsatzes von KI außerordentlich schwierig, da keine generelle Pflicht zur Begründung ablehnender Entscheidungen existiert. [22] Betroffenen kommt das Gesetz daher ein Stück weit entgegen: Können Betroffene ausreichende Indizien für eine verbotene Benachteiligung nachweisen, wird die Beweislast umgekehrt (§ 22 AGG) – der Arbeitgeber muss dann beweisen, dass eine Ungleichbehandlung überhaupt nicht vorlag, nicht wegen eines verpönten Merkmals erfolgte oder dass die Ungleichbehandlung durch einen sachlichen Grund gerechtfertigt war.

Kernvoraussetzung für einen wirksamen Schutz (auch) gegen Diskriminierung durch KI ist damit, dass Betroffene Kenntnis darüber erlangen, dass KI über sie (mit)entscheidet. Im Fall der vollautomatisierten Entscheidungsfindung verpflichtet unstreitig bereits das Datenschutzrecht zu einer entsprechenden Information. [23] Eine vergleichbare Informationspflicht für den Fall der bloßen Entscheidungsassistenz ist im Datenschutzrecht jedoch allenfalls entgegen der herrschenden Lesart begründbar, [24] im Vertragsrecht könnte die Annahme einer entsprechenden (vor-)vertraglichen Auskunftspflicht Abhilfe schaffen. Im Sachverständigengutachten für den Dritten Gleichstellungsbericht der Bundesregierung wird empfohlen, Transparenzpflichten explizit im AGG zu verankern. Danach soll u. a. offengelegt werden, wenn Arbeitgebende im Zusammenhang mit personellen Maßnahmen algorithmische Systeme einsetzen. [25] Verstoßen sie gegen diese Pflicht, soll das hiernach – vergleichbar mit einer gegen das Benachteiligungsverbot verstoßenden Stellenausschreibung [26] – als widerlegbares Indiz für eine unzulässige Ungleichbehandlung zu werten sein. [27]

Eine weitere Voraussetzung für individuellen Rechtsschutz ist das Bestehen von Anhaltspunkten dafür, dass ein System Entscheidungen in diskriminierender Weise vorbereitet oder selbst trifft. Dies wiederum setzt zuverlässige Informationen über ein System, dessen Funktionsweise und dessen Diskriminierungsgeneigtheit voraus – eine Information, über die Betroffene im Regelfall nicht verfügen. Im Kontext von KI ist darüber hinaus fraglich, ob Anwender*innen und Betroffene diese Informationen durch Einblick in das System überhaupt erhalten können.

[22]

Vgl. in diesem Sinne EuGH, NJW 2012, 2497 ff. – Meister/Speech Design Carrier Systems GmbH.

[23]

Grundlage hierfür ist Art. 22 Abs. 1 DSGVO i.V.m. Art. 13 Abs. 2 lit. f) DSGVO.

[24]

Sesing, Multimedia und Recht (MMR) 2021, S. 288 ff.; in diese Richtung zum Auskunftsrecht nach Art. 15 DSGVO auch Hacker, NJW 2020, 2142, 2144.

[25]

Bundesregierung (2021): Dritter Gleichstellungsbericht der Bundesregierung. BT-Drucksache 19/30750, Berlin.

[26]

Vgl. § 11 AGG.

[27]

Bundesregierung (2021): Dritter Gleichstellungsbericht der Bundesregierung. BT-Drucksache 19/30750, Berlin: S. 169.

Zur Verbesserung können insoweit das Testen von KI-Systemen vor deren Einsatz und die kontinuierliche Überprüfung im Einsatz beitragen. Naheliegender erschien es aus der Perspektive des deutschen Rechts, die Pflichten zur Diskriminierungsprävention in § 12 AGG um eine Testpflicht für KI-Systeme zu erweitern und entsprechend zu dokumentieren. Bei Nichtvorliegen einer entsprechenden Testdokumentation wäre es sachgerecht, diesen Umstand als ausreichendes Indiz anzusehen, um die Beweiserleichterung des § 22 AGG zur Anwendung zu bringen. Die Sachverständigenkommission für den Dritten Gleichstellungsbericht verfolgt einen ähnlichen Ansatz, wenn vorgeschlagen wird, die präventiven Organisationspflichten des § 11 AGG um eine konkrete Regelung für den Einsatz von algorithmischen Systemen zu ergänzen. Neben der angesprochenen Offenlegungspflicht soll eine diskriminierungsspezifische Prüfpflicht normiert werden, für deren Einhaltung Arbeitgebende beweisbelastet sein sollen. [28]

Ausblick: Der Vorschlag für eine EU-Verordnung über Künstliche Intelligenz

Im April 2021 hat die Europäische Kommission einen Vorschlag für eine Verordnung zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (KI-VO) veröffentlicht. [29] Die vorgeschlagene Regelung unterscheidet zwischen unzulässigen KI-Systemen, Hochrisiko-KI-Systemen und sonstigen KI-Systemen. Von Bedeutung für den Schutz vor ungerechtfertigter Ungleichbehandlung ist, dass der Einsatz von KI im hochregulierten Bereich der Anbahnung und Durchführung von Arbeitsverhältnissen insgesamt als Hochrisiko-Szenario eingestuft wird. Dies folgt nicht unmittelbar aus Art. 6 Abs. 1 der KI-VO, sondern aus dem von Art. 6 Abs. 2 KI-VO in Bezug genommenen Anhang III des Vorschlags. Hiernach sind einerseits KI-Systeme als Hochrisiko-Systeme einzustufen, die bestimmungsgemäß für die Einstellung oder Auswahl natürlicher Personen verwendet werden sollen, insbesondere für die Bekanntmachung freier Stellen, das Sichten oder Filtern von Bewerbungen und das Bewerten von Bewerber*innen in Vorstellungsgesprächen oder Tests, [30] sowie ferner solche KI-Systeme, die bestimmungsgemäß für Entscheidungen über Beförderungen und über Kündigungen von Arbeitsvertragsverhältnissen, für die Aufgabenzuweisung sowie für die Überwachung und Bewertung der Leistung und des Verhaltens von Personen in solchen Beschäftigungsverhältnissen verwendet werden sollen. [31]

Die Einordnung als Hochrisiko-KI-System führt zur Geltung der Anforderungen der Artt. 8 ff. KI-VO. Hiernach ist insbesondere eine Risikobewertung durchzuführen und laufend zu aktualisieren, [32] zudem sind derartige Systeme vor dem Inverkehrbringen zu

[28]

Bundesregierung (2021):
Dritter Gleichstellungsbericht der Bundesregierung.
BT-Drucksache 19/30750,
Berlin: S. 168f.

[29]

COM(2021) 206 final, <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52021PC0206>.

[30]

Vgl. Ziffer 4 Buchstabe a)
Anhang III zur KI-VO.

[31]

Vgl. Ziffer 4 Buchstabe b)
Anhang III zur KI-VO.

[32]

Die Risikobewertung ist gemäß Art. 9 Abs. 2 KI-VO als „kontinuierlicher iterativer Prozess“ konzipiert.

testen. [33] Wenngleich als Testziel die Erfüllung der Anforderungen der Artt. 8 ff. KI-VO genannt wird (Art. 9 Abs. 5 Satz 2 KI-VO), bleibt unklar, welche Anforderungen nach welchen Kriterien zu testen sind. Für datengetriebene Systeme schreibt Art. 10 KI-VO zudem gewisse Qualitätsanforderungen für Trainings-, Validierungs- und Testdatensätze vor. Die Anforderung, mögliche Verzerrungen ex ante zu vermeiden (vgl. Art. 10 Abs. 2 Buchstabe f KI-VO), ist in der Sache unterstützenswert – die Umsetzung hingegen ist technisch alles andere als trivial und derzeit Gegenstand intensiver Forschung.

[33]

Vgl. Art. 9 Abs. 7 KI-VO.

Im Hinblick auf die bestehenden Befürchtungen einer gesellschaftlich unerwünschten und/oder rechtlich unzulässigen Diskriminierung durch KI ist der Vorschlag der KI-VO jedoch sehr zurückhaltend: Die Schutzziele der Verordnung werden zwar in der Begründung und den Erwägungsgründen [34] angesprochen, aber in ihrem bindenden verfügenden Teil nicht explizit definiert. Auch das Problem fehlender Transparenz wird nicht umfassend adressiert, da die maßgebliche Transparenzvorschrift in Art. 13 KI-VO lediglich auf die Nutzer*innen von KI abzielt. Die Konstellation, dass Nutzer*innen KI-Systeme zur Entscheidung über Dritte einsetzen, wird nicht angesprochen. Im Kontext des Personalmanagements wäre eine Adressierung eines derartigen Mehrpersonenverhältnisses jedoch von entscheidender Bedeutung, zumal die Regelung des Art. 27 KI-VO – in dem Pflichten der Nutzer*innen von KI-Systemen geregelt sind – keine Pflichten gegenüber „Betroffenen“ enthalten.

[34]

Vgl. nur Erwägungsgründe 5, 28, 36 der KI-VO.

Unklar ist zudem, wie sich das Regelungsinstrument der KI-VO – die als „horizontale“ Querschnittsregelung mit Ausnahme weniger Sektoren alle Wirtschafts- und Lebensbereiche erfassen soll – zu anderen Rechtsakten innerhalb der EU verhält. Weder werden die Anliegen der Antidiskriminierungs-Richtlinien der EU im Entwurf konkret benannt, noch wird das Verhältnis zu diesen Richtlinien klargestellt. Aufgeworfen ist darüber hinaus die Frage, ob für die hier oder im Gutachten zum Dritten Gleichstellungsbericht erwogenen Ergänzungen des AGG mit Blick auf KI-spezifische Risiken weiterhin Raum bleibt. Ausweislich des Erwägungsgrundes 2 der KI-VO soll eine Fragmentierung des Binnenmarkts durch unterschiedliche nationale Rechtsakte unterbunden werden; zugleich enthält sich der Vorschlag jedoch einer klaren Positionierung zum angestrebten Harmonisierungsgrad.

Insoweit wäre durchaus zu befürworten, dass nationale sektorspezifische Regelungen weiterhin möglich bleiben, wenn sie – in Übereinstimmung mit der Zielsetzung der KI-VO [35] – einerseits dazu bestimmt sind, durch die EU-Grundrechtecharta geschützte Rechte wie die Würde des Menschen (Art. 1), die Nichtdiskriminierung (Art. 21) und die Gleichheit von Frauen und Männern (Art. 23) zu gewährleisten, und sich andererseits

[35]

Siehe dazu die Begründung zur KI-VO, S. 12.

nicht in Widerspruch zu den Grundsätzen der KI-VO setzen. Eine Blaupause für ein derartiges Ineinandergreifen von unionaler Harmonisierung und nationaler Konkretisierung könnte das in der DS-GVO angelegte Konzept von Öffnungsklauseln sein – der Übergang zu einer „KI-Grundverordnung“ erscheint zumindest diskutabel.



Dr. Andreas Sasing ist Geschäftsführer des Instituts für Rechtsinformatik und wissenschaftlicher Mitarbeiter an der Universität des Saarlandes, zugleich Lehrbeauftragter an der Fachhochschule Südwestfalen. Seit Abschluss des

Studiums der Rechtswissenschaften ist er – unterbrochen durch den juristischen Vorbereitungsdienst und zweites Staatsexamen – als wissenschaftlicher Mitarbeiter tätig, zunächst an der Ruhr-Universität Bochum, seit 2015 an der Universität des Saarlandes. Im Jahr 2021 erfolgte die Promotion zu einer urheberrechtlichen Fragestellung. Dr. Andreas Sasing ist Autor zahlreicher Veröffentlichungen auf dem Gebiet des Urheber-, IT- und Datenschutzrechts.



Angela Tschech ist wissenschaftliche Mitarbeiterin und Doktorandin am Lehrstuhl für Bürgerliches Recht, Arbeitsrecht und Recht der Sozialen Sicherheit an der MLU Halle (Saale). Sie forscht zu arbeits- und antidiskriminierungs-

rechtlichen Anforderungen an den Einsatz von algorithmischen Systemen im Personalmanagement. Für das Gutachten zum Dritten Gleichstellungsbericht „Digitalisierung geschlechtergerecht gestalten“ hat sie die Lehrstuhlinhaberin und Sachverständige Prof'in Dr'in Katja Nebe inhaltlich insb. bei den arbeitsrechtlichen Einschätzungen zu Plattformarbeit und softwaregestützter Personalauswahl unterstützt. Sie engagiert sich im Forum Legal Gender Studies und im Netzwerk IT & Recht.

Impressum

Eine Veröffentlichung aus dem Projekt „ExamAI – KI Testing & Auditing“ <https://testing-ai.gi.de>

November 2021

Herausgeberin

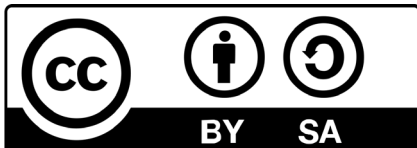
Gesellschaft für Informatik e.V. (GI)
Spreepalais am Dom
Anna-Louisa-Karsch-Straße 2
10178 Berlin

Projektleitung

Nikolas Becker
nikolas.becker@gi.de

Gestaltung

Gabriela Kapfer
<http://smileinitial.plus>



Dieser Beitrag unterliegt einer Creative-Commons-Lizenz (CC BY-SA). Die Vervielfältigung, Verbreitung und Veröffentlichung, Veränderung oder Übersetzung von Inhalten der Gesellschaft für Informatik e.V., die mit der Lizenz „CC BY-SA“ gekennzeichnet sind, sowie die Erstellung daraus abgeleiteter Produkte sind unter den Bedingungen „Namensnennung“ und „Weiterverwendung unter gleicher Lizenz“ gestattet. Ausführliche Informationen zu den Lizenzbedingungen finden Sie hier: <http://creativecommons.org/licenses/by-sa/4.0/>

ExamAI – KI Testing & Auditing

Dieses Arbeitspapier erscheint als Teil des Forschungsprojekts „ExamAI – KI Testing und Auditing“, das sich der Erforschung geeigneter Test- und Auditierungsverfahren für KI-Anwendungen widmet. Es steht unter der Leitung der Gesellschaft für Informatik e. V. und wird von einem interdisziplinären Team bestehend aus Mitgliedern der TU Kaiserslautern, der Universität des Saarlandes, des Fraunhofer-Instituts für Experimentelles Software Engineering IESE und der Stiftung Neue Verantwortung getragen und im Rahmen des Observatoriums Künstliche Intelligenz in Arbeit und Gesellschaft (KIO) der Denkfabrik Digitale Arbeitsgesellschaft des Bundesministeriums für Arbeit und Soziales (BMAS) gefördert.

Informationen zum Projekt und weitere Veröffentlichungen finden Sie unter: <https://testing-ai.gi.de/>

Projektpartner*innen:

GESELLSCHAFT
FÜR INFORMATIK



INSTITUT FÜR
RECHTSINFORMATIK
UNIVERSITÄT DES SAARLANDES



Gefördert durch:



Im Rahmen des:



Observatorium Künstliche Intelligenz
in Arbeit und Gesellschaft