GENDER OPEN REPOSITORIUM

AI Powered Recruiting? Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst : Expertise für den Dritten Gleichstellungsbericht der Bundesregierung

Jaume-Palasi, Lorena; Lindinger, Elisa; Kloiber, Julia

2020

https://doi.org/10.25595/2404

Veröffentlichungsversion / published version
Working Paper

Empfohlene Zitierung / Suggested Citation:

Jaume-Palasi, Lorena; Lindinger, Elisa; Kloiber, Julia: *AI Powered Recruiting? Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst : Expertise für den Dritten Gleichstellungsbericht der Bundesregierung*. Berlin: Geschäftsstelle Dritter Gleichstellungsbericht der Bundesregierung; Institut für Sozialarbeit und Sozialpädagogik e.V., 2020. DOI: https://doi.org/10.25595/2404.

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY 4.0 Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:

This document is made available under a CC BY 4.0 License (Attribution). For more information see:

https://creativecommons.org/licenses/by/4.0/deed.en

GenderOpen – Repositorium für die Geschlechterforschung: www.genderopen.de

FREIE UNIVERSITÄT BERLIN

HUMBOLDT-UNIVERSITÄT ZU BERLIN

DRITTER GLEICHSTELLUNGSBERICHT

DRITTER
GLEICHSTELLUNGSBERICHT

Lorena Jaume-Palasi, Elisa Lindinger, Julia Kloiber

AI Powered Recruiting?

Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst

Expertise für den Dritten
Gleichstellungsbericht der Bundesregierung

AI Powered Recruiting?

Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst.

Expertise für den Dritten Gleichstellungsbericht der Bundesregierung

Lorena Jaume-Palasi, Elisa Lindinger, Julia Kloiber

31.10.2020

Inhalt

1. Fragestellung und Forschungsstand zum Einsatz algorithmischer Assistenzsysteme im Recruiting-Bereich in Deutschland 5

1.1. Fragestellung 5

1.2. Definition und Terminologie 6

1.3. Disziplinäre Verortung der Forschungsfrage 6

1.4. Grundlegende Fragen der einzelnen Disziplinen 8

2. Recruiting im Kontext 10

2.1. Was ist Recruiting? 10

2.2. Die Automatisierung des Recruitingbereichs 11

2.3. Gängige Produktarten von Recruitingsoftware 11

2.4. Der Einsatz von algorithmischen Assistenzsystemen im Recruitingbereich: Marketingversprechen versus Realität 13

2.4.1. Sourcing 14

2.4.2. Screening 16

2.4.3. Auswahlverfahren 17

2.5. Diskriminierungspotenziale im Bewerbungsprozess 18

2.5.1. Sourcing 18

Wer sieht die Stellenanzeige? 18

Wer fühlt sich angesprochen? 19

2.5.2. Screening 20

2.5.3. Auswahlverfahren 20

Kandidat*innenmanagement 20

Bewerbungsgespräche 21

2.6. Geschäftsmodelle 22

2.6.1. Plattformisierung 22

2.6.2 Weitere Beobachtungen 24

3. Methodologische Vorüberlegungen und Ansätze für Taxonomien 25

3.1. Verortung der Ursachen für Diskriminierungen 25

3.1.1. Optimierungsziel 26

3.1.2. Datengrundlagen 26

3.1.3 Format 27

3.1.4 Hardware und Sensorik 27

3.1.5 Sichtbarkeit 28

3.1.6 Kriterien-Angabe 28

3.1.7 Konzeptualisierung des Systems 29

3.1.8 Abbildung des Verfahrens 30

3.1.9 Einbettung der Entscheidung 31

3.2. Das Ökosystem 31

3.2.1. Systeme 32

3.2.2. Metriken 32

3.2.3. Wer ist in das Verfahren involviert? 33

3.2.4. Der Prozess 33

3.2.5. Die Regeln 34

Soziale Regeln 34

Gesetzliche Regeln 35

4. Algorithmen- und Datenmärkte 35

4.1. Algorithmenmärkte 36

Erwartungen an digitale Lösungen unterscheiden sich. 37

Häufigste Nennungen 37

Das Undokumentierte 37

4.2. Datenmärkte: Über die algorithmische Materialisierung und Ökonomisierung von Affekten und Emotionen 41

Die Ökonomisierung von Emotionen 42

Verdatung der Emotionen 44

Was kann verdatet werden? 44

Biometrie und Verhaltensanalyse 45

5. Empfehlungen **50**

5.1. Empfehlungen zur Governance von ASR-Systemen 50

5.2. Empfehlung zu zusätzlicher Forschung 50

5.2.1. Studien zur Entwicklung von interdisziplinären Methodologien und wissenschaftlicher Sprache 50

5.2.2. Studien zum Nutzungsverhalten 50

5.2.3. Studie zum Einsatz von algorithmischen Assistenzsystemen 51

5.2.4. Methodologische Einordnung der eingesetzten Verfahren 51

5.3.5. Sprachwissenschaftliche Erforschung 52

6. Ausblick und Fazit **53**

6.1. Fehlende Reflexion und Berücksichtigung der bestehenden Strukturen und Asymmetrien im Arbeitsmarkt 54

6.2. Die Ökonomisierung von Emotionen und Charakter 54

Executive Summary **56**

Literaturverzeichnis **57**

Über die Autorinnen **67**

Anhang **69**

1. Fragestellung und Forschungsstand zum Einsatz algorithmischer Assistenzsysteme im Recruiting-Bereich in Deutschland

1.1. Fragestellung

Zunehmend werden in Deutschland neue Softwareprodukte für Prozesse im Personalmanagement angeboten, die sowohl direkt als auch indirekt Recruiting-Funktionen besitzen. Die Evaluation dieser Produkte in Hinblick auf geschlechtsdiskriminierende Faktoren und Prozesse beschränkt sich derzeit weitgehend auf die Qualität der Datengrundlage einerseits und der Funktionsweise der Algorithmen andererseits. In unserer Expertise erweitern wir die Betrachtung von Recruiting-Softwareprodukten um eine intersektionale Perspektive (im Sinne von WALGENBACH 2012) auf Geschlechterdiskriminierung. Die vorliegende Studie hat ihren Ausgangspunkt in einer Recherche der derzeit angebotenen Software-Produkte im Bereich Personalrecruiting. Das Ergebnis dieser Recherche ist einer Produktliste im Anhang des Gutachtens zu entnehmen. Diese Produktliste wurde nach drei Grundbetrachtungen und Rastern klassifiziert:

Raster 1: Recruiting unterteilt sich in verschiedene Tätigkeiten und Arbeitsschritte (siehe Tabelle Anhang 1: Recruitingtätigkeiten). Um potenzielle Auswirkungen auf Gleichstellungsfragen punktuell bestimmen zu können, differenzieren wir nach diesen Tätigkeiten und Arbeitsschritten.

Raster 2: Welche Art von Kund*innen nutzen diese Technologien? Dabei unterscheiden wir auf einer Skala zwischen High-Volume, High-Frequency-Recruitingbedarf und Low-Volume-, Low-Frequency-Recruitingbedarf.

Raster 3: Neben Dienstleistungen oder Plattformen mit einem expliziten Recruitingcharakter gibt es andere Plattformen (SAP, Oracle, Slack) und Dienstleistungen, die faktisch als Recruitingplattform fungieren und auf denen algorithmische Verfahren zum Einsatz kommen.
Die Liste ist exemplarisch und dokumentiert Beispiele mit verschiedenen Produktlösungen und Annahmen (von Anwerbe-Erfolgs-Analysen über Culture-Fit-Analytics, Referenzchecks, Chatbots etc.). Viele der Anbieter sind US-Unternehmen und haben

europäische Unternehmen als Kunden. Darüber hinaus bieten die meisten Anbieter Integrierbarkeit mit Architekturlösungen wie SAP, Oracle oder IBM an.
Wir betrachten außerdem den weiteren technischen und marktlichen Kontext, in dem Recruiting stattfindet, und betrachten auch Softwareprodukte, die nicht primär für den Recruitingbereich entwickelt wurden, aber dort zum Einsatz kommen.
Ziel der Expertise ist es, die Diskriminierungspotenziale beim Einsatz von Technologien für Personaleinstellungsprozesse zu beschreiben.
Um sich einer systematischen Darstellung der sozialen und marktlichen Einbettung digitaler Recruiting-Werkzeuge anzunähern, entwerfen wir methodologische Vorüberlegungen und Ansätze für Taxonomien, mit denen das Untersuchungsfeld weitergehend analysiert werden kann.

1.2. Definition und Terminologie

In der vorliegenden Expertise verwenden wir den Begriff ***a****lgorithmische Assistenz****s****yste-me im* ***R****ecruitingbereich* (ASR) für den Einsatz von Softwarelösungen im Bereich der Personalgewinnung (*Human Resources*, HR). Damit umfasst der Begriff ein breites Feld von digitalen Werkzeugen, die Daten anhand von Modellen analysieren und auswerten. Ähnlich wie beim Begriff der algorithmischen Entscheidungssysteme (ADM) (KRUEGER & LISCHKA 2018, 13) verstehen wir algorithmische Systeme nicht als eine Form von automatisierter Entscheidung, sondern als technische Ausführung von vorab getroffenen menschlichen Entscheidungen. Die in der Recruiting-Branche für diese Produkte übliche Bezeichnung *Künstliche Intelligenz* ist in ihrer Verwendung sehr unscharf. In der vorliegenden Expertise verwenden wir den Begriff nur, wenn er auch von den besprochenen Unternehmen verwendet wird.

1.3. Disziplinäre Verortung der Forschungsfrage

In der Frage, wie sich ASR auf die Gleichstellung am Arbeitsmarkt in Deutschland auswirken, treffen mehrere unterschiedliche akademische Fächer aufeinander:

- die **Informatik**, um die algorithmischen Grundlagen der ASR zu analysieren;
- die **Rechtswissenschaften**, um juristische Fragen der Gleichstellung im Kontext zu erörtern;
- die **Techniksoziologie**, um die gesellschaftliche Einbettung zu untersuchen, in der Softwareprodukte entwickelt und eingesetzt werden;

- die **Genderstudien**, um die Konstruktion der Kategorie „Geschlecht“ im Recruiting-, im Rechts- und im KI-Kontext zu beleuchten.

Zusätzlich zur Verschneidung dieser Fachbereiche stellt der Bezugsrahmen Deutschland mit seiner speziellen, stark mittelständisch geprägten Unternehmenslandschaft eine weitere Besonderheit dar: 95 % der Unternehmen in Deutschland sind kleine oder mittlere Unternehmen (KMU). Auch aus Sicht der Arbeitnehmer*innen sind KMU relevant: 58 % der sozialversicherungspflichtig Beschäftigten arbeiten in einem KMU (BVMW o. J.).

Die bisherige Forschung, die in den Themenbereich dieser Expertise fällt, ist vielfältig, deckt jedoch nicht alle genannten Dimensionen gleichzeitig ab. Außerdem befasst sie sich überwiegend mit Fragestellungen, die sich auf die rechtlichen und sozio-ökonomischen Gegebenheiten in den USA beziehen. Die erste tiefergehende Studie für den europäischen Raum bieten SÁNCHEZ-MONEDERO ET AL. (2020) mit ihrer Untersuchung von Bias-Vermeidung bei einer Auswahl von *automated hiring systems,* einer Untergruppe der in dieser Expertise betrachteten ASR*,* die nachweislich in Großbritannien eingesetzt werden. Hier wird die mangelnde Übertragbarkeit von Studien mit US-amerikanischem Bezug bestätigt, insbesondere unter rechtlichen Gesichtspunkten. Die Studie befasst sich jedoch explizit nicht mit der sozio-ökonomischen Einbettung der Softwareprodukte, sondern konzentriert sich auf die Analyse, mit welchen Designkonzepten sie „Bias“ und Diskriminierung im Recruiting-Prozess entgegenwirken wollen. Eine Erfolgsanalyse dieser Designkonzepte bleibt ebenfalls aus.

Studien mit einem Fokus auf den deutschen oder deutschsprachigen Raum beschränken sich meist auf quantitative Untersuchungen zu Akzeptanz und Einsatz von ASR in HR-Abteilungen. Grundlegende rechtliche, ökonomische und techniksoziologische Fragen werfen KNOBLOCH & HUSTEDT (2019) in ihrem Papier auf, betrachten aber weder die psychologischen Modelle, die einigen ASR zugrunde liegen (siehe Abschnitt 3.1.3.), noch den techniksoziologischen Kontext der Softwareprodukte. Die Studie „Algorithmen und Gesellschaft“ (SCHÜNEMANN & LEBERT 2019, 24–30) bietet eine gute Zusammenfassung des Forschungsstandes in Deutschland und den USA und erweitert den Blick auf ASR um eine Perspektive aus der Zivilgesellschaft, beschränkt sich in ihrer Bewertung jedoch weitgehend auf die Qualität der Daten und Ergebnisse und hinterfragt weder die Zweckmäßigkeit noch die Rechtmäßigkeit von ASR.

Einen allgemeinen Überblick über die Entwicklung von ASR und ihre Diskriminierungspotenziale bietet die Studie „Help Wanted“ (BOGEN & RIEKE 2018), die sich jedoch auf den US-amerikanischen Raum bezieht. Diese und andere Bias-Analysen (z. B.

MCCURDY & SERBETÇI 2020) fokussieren besonders auf zwei Bereiche: Diskriminierung nach Geschlecht einerseits und Diskriminierung aufgrund von Herkunft oder Hautfarbe andererseits. Dabei gerät außer Acht, dass Diskriminierung in der Regel intersektional wirkt: Von einem algorithmischen System, das beispielsweise gleichzeitig nach Geschlecht und Hautfarbe diskriminiert, sind Women of Color deutlich schwerer betroffen als weiße Frauen. Während eine intersektionale Betrachtung von Diskriminierungsmustern (allerdings allgemein, nicht in Bezug auf algorithmische Diskriminierung, siehe Abschnitt 4.) im Bereich der Genderstudien und der Rechtswissenschaften bereits durchgeführt wird (zusammengefasst bei BAER ET AL. 2010, 10–28), ist sie im soziotechnischen Kontext noch die Ausnahme. Wir sind der Überzeugung, dass gerade an der Schnittstelle von Gesellschaft und Technologie eine intersektionale Betrachtungsweise essentiell ist.
Dafür ist es unabdingbar, das soziotechnische System (vgl. ROPOHL 2009, 135), das den Einsatz einer Software erst bedingt und ihre Ergebnisse in realweltliche Entscheidungen übersetzt, in den Vordergrund zu rücken. Damit werden erstmals auch die Menschen, die im Recruitingbereich tätig sind, in solche Analysen von KI-Systemen mit einbezogen und ihr Wissen, ihre Handlungsspielräume und ihre Motivationen berücksichtigt (siehe hierzu Abschnitt 4.2.3.).

1.4. Grundlegende Fragen der einzelnen Disziplinen

Nicht nur durch die multidisziplinäre Betrachtung von ASR ergeben sich neue Forschungsfragen. Auch innerhalb der einzelnen Disziplinen gibt es noch Forschungslücken, die dazu führen, dass die Problematik von ASR in Bezug auf Gleichstellungsfragen am Arbeitsmarkt nur in Grundzügen aufgezeichnet, aber nicht abschließend bewertet werden kann. Beispiele für derartige offene Fragen sind folgende:

Informatik: Ist automatisiertes De-Biasing von Algorithmen möglich, und wenn, welche Design-Ansätze sind denkbar (z. B. SANCHEZ-MONEDERO ET AL. 2020)? Welche grundlegenden ethischen Aspekte müssen bei der Entwicklung und beim Einsatz von Methoden aus dem Bereich Algorithmic Decision-Making (ADM) beachtet werden?

Datenwissenschaften: Ist algorithmisches/automatisiertes De-Biasing von Datenbeständen möglich, und wenn, welche Design-Ansätze sind denkbar?

Genderforschung: Welche Auswirkungen haben die Konzepte „genderneutral“ und „genderinklusiv“ auf sprachlicher Ebene? Inwieweit können Befunde und Best Practices aus dem Englischen auf den deutschen Sprachraum übertragen werden, wenn dort das grammatikalische Geschlecht viel tiefer in der Sprache verankert ist?

Rechtswissenschaften: Wie beeinflussen neue Berufsfelder wie *clickwork* und *crowdwork* Arbeitnehmer*innenrechte?

2. Recruiting im Kontext

2.1. Was ist Recruiting?

Recruiting (zu deutsch auch Personalbeschaffung) beschreibt den Prozess der Suche und Auswahl von Kandidat*innen für Jobpositionen in Unternehmen oder Organisationen. Das Ziel von Recruiting ist einerseits die Gewinnung, andererseits die Auslese von Kandidat*innen durch den *„Hiring Funnel“ (BOGEN & RIEKE 2018)* oder Einstellungstrichter. Innerhalb dieses bildlich gedachten Trichters kommen in den verschiedenen Stufen unterschiedliche Arten von algorithmischen Verfahren zum Einsatz (siehe 2.3). Die Personalauswahl ist ein mehrstufiger Prozess, im Laufe dessen eine Vielzahl von Entscheidungen getroffen werden (BOWER ET AL. 2017). Die Entscheidungen umfassen unter anderem die Auswahl der Plattform(en), auf der die Ausschreibung platziert wird, welche Informationen die Ausschreibung enthält, wer in den Auswahlprozess involviert ist und welche Softwarelösung verwendet wird. Lange bevor die ersten Bewerbungen eingehen, wird somit bereits eine Reihe an Entscheidungen getroffen, die Einfluss auf den Prozess und die Auswahl haben.
Der Personalauswahlprozess lässt sich in mehrere Stufen und Verfahren gliedern:

- **Sourcing:** In dieser Phase der Personalbeschaffung suchen Arbeitgeber*innen nach Kandidat*innen, die sich auf ihre Stellenangebote bewerben. Dafür wird ein Anforderungsprofil für die offene Position erstellt, das die Bewerber*innen erfüllen müssen, und eine inhaltlich fundierte Stellenausschreibung verfasst. Diese wird veröffentlicht, beworben und in vielversprechende Kanäle ausgespielt. Alternativ suchen Headhunter*innen aktiv nach passenden Kandidat*innen und machen sie auf die Ausschreibung aufmerksam.
- **Sichten und Prüfen von Unterlagen (Screening):** Eingereichte Unterlagen wie Lebensläufe, Zeugnisse, Arbeitsproben und Motivationsschreiben werden auf die Eignung der Kandidat*innen hin gesichtet. Eine erste Vorauswahl wird getroffen.
- **Auswahlverfahren:** Es gibt eine Vielzahl an unterschiedlichen Verfahren: Interviews, Eignungstests (z. B. über Assessment Center), Fragebögen etc. Die Verfahren zielen darauf ab, zusätzliche aussagekräftige Informationen über die Kandidat*innen und deren Eignung zu sammeln, um eine gute Auswahl treffen zu können. Die Verfahren können mehrstufig sein.
- **Finale Auswahl:** Auf Basis der Informationen aus den Unterlagen und den Auswahlverfahren wird eine finale Entscheidung getroffen.

Je nach Unternehmen und Branche weicht das Vorgehen beim Auswahlprozess ab. Es gibt aber Prozessschritte, wie das Sichten von Unterlagen und das Durchführen von Interviews, die als Basis für die Auswahl dienen und nahezu von allen Unternehmen genutzt werden. In allen Schritten des Recruiting-Prozesses kann es zur Ungleichbehandlung oder Diskriminierung von Kandidat*innen kommen (BOGEN & RIEKE 2018, 13, vgl. auch Abschnitt 2.5.).

2.2. Die Automatisierung des Recruitingbereichs

Recruiting ist traditionell ein zeitaufwändiger Prozess, der viele Ressourcen bindet und dessen Erfolg für Unternehmen direkte wirtschaftliche Auswirkungen haben kann. Mit der Verbreitung von Computern entstand die Möglichkeit, den Recruitingprozess zu digitalisieren und beispielsweise Tests am Computer durchzuführen und von ihm auswerten zu lassen. Mit dem Aufkommen des Internets wurde Online-Recruiting populär. Unternehmen begannen damit, Jobbeschreibungen online zu posten, über E-Mail mit den Bewerber*innen zu kommunizieren oder Bewerbungsgespräche per Videotelefonat durchzuführen. Online-Recruiting ist für die Unternehmen eine Kosten- und Zeitersparnis durch die deutlich erhöhte Effizienz (OKOLIE & IRABOR 2017, 118). Der Einsatz von ASR setzt diesen Trend weiter fort. Vereinfacht gesagt können sie drei wichtige Geschäftsanforderungen unterstützen: die Automatisierung von Geschäftsprozessen (*Process Automation*), die Gewinnung von Erkenntnissen durch Datenanalyse (*Cognitive Insights*) und die Interaktion mit Kund*innen und Mitarbeiter*innen (*Cognitive Engagement*) (nach DAVENPORT & RONANKI 2018).

2.3. Gängige Produktarten von Recruitingsoftware

Im Rahmen der Expertise haben wir eine nicht abschließende Liste von Softwareprodukten für den Recruitingbereich erstellt, die mit dem Einsatz von KI werben oder nachweislich KI nutzen (s. Anhang). Das Ziel der Liste ist es, die Bandbreite der technischen Verfahren, der Einsatzgebiete und der Marketingversprechen darzustellen. Unsere Recherche zeigt, dass insbesondere Standalone-Lösungen, also selbständige Softwareprodukte, die nicht in Plattformen oder Produkte eingebettet sind, häufig nur kurz auf dem Markt bestehen und selbst Produkte von 2018 heute oft nicht mehr erhältlich sind. Aufgrund der hohen Beweglichkeit dieses emergenten Marktes und der Entstehung neuer Beschäftigungsformen und -felder (z. B. *clickwork*) ist zu erwarten,

dass sich dieser Trend weiter fortsetzt. Eine kontinuierliche Beobachtung neu angebotener Produkte und neu entstehender Produktkategorien ist deshalb notwendig. In der Tabelle 1 haben wir eine Übersicht gängiger ASR-Produktkategorien, die international eingesetzt werden, erstellt und geben an, ob es sich dabei um ein speziell für den Recruiting-Bereich entwickeltes Produkt handelt (S) und in welchem Recruiting-Schritt es zum Einsatz kommt (1: Sourcing, 2: Screening, 3: Interview, 4: Auswahl). Ob die ASR tatsächlich auch in Deutschland *eingesetzt* werden, ist auf Basis der Datenlage nur in Ausnahmefällen nachvollziehbar. Details zu den einzelnen Produktbeispielen sind in der Tabelle Anhang 1 zu finden.

Produkt- kategorie	Beschreibung	Beispiele	S	1	2	3	4
Recruitment Marketing	Unterstützt bei der Formulierung von Stellenanzeigen und gibt Feedback oder Verbesserungsvorschläge für Formulierungen, die nachweislich Personengruppen abschrecken.	Textio, Unbias	teils	x	-	-	-
Plattformen und Plugins	Teils gezielt für Recruiting entwickelt, teils für allgemeinere Zwecke. Suchfilter und Recommendation-Algorithmen haben enormen Einfluss, sind aber intransparent; zusätzlich gibt es Meta-Werkzeuge, die auf den Plattform-Datenbanken aufbauen.	Facebook Jobs, Facebook Marketplace, LinkedIn, Xing, Semantic Search, Carerix Search!	teils	x	-	-	-
Datenan-reicherung	Gleichen Bewerber*innendaten ab und greifen dafür auf externe Datenbestände zurück (z. B. Korrektheit von Zertifikaten, Abschlüssen, Vorstrafenregister, Gesundheitsdaten und Kreditwürdigkeit) oder generieren eigene Datenpunkte (Verhaltens-, Kommunikationsanalyse).	Checkr, Skillsurvey, Ideal Screening Software, Prescreen, Veremark, Orange Tree	ja	-	x	-	-
Matching-Werkzeug	Analysieren und kategorisieren die Bewerber*innen nach unterschiedlichen psychologischen Modellen und empfehlen Bewerber*innen, die zum Unternehmen passen sollen.	MoBerries, Eightfold, Prescreen	ja	x	x	-	-
Chatbot	Überwiegend administrative Unterstützung bei repetitiven Tätigkeiten von Recruiter*innen (Terminfindung etc.), seltener Datenbeschaffung für psychometrische Analysen.	XOR, HR Chatbots, Olivia Capture, Mya, Ideal Recruiter Chatbots	teils	x	x	-	-
Management-Tool	Dient der Bearbeitung der Kandidat*innen-Akten und der Nachverfolgung des Bewerbungsprozesses, sammelt Daten zur Qualität der Betreuung und Kommunikation. Die Management-Tools werden im Recruiting-Kontext	Greenhouse Recruiting, Prescreen, Metaview, Bullhorn, Cornerstone, Oracle, SmartRecruiters	ja	-	x	x	-

	als ATS- (Applicant Tracking System) oder HCM-Systeme (Human Capital Management) bezeichnet. Sie ähneln CRM-Systemen (Customer Relationship Management).						
Assessment-Tool	Dient als Plattform für Assessments oder Bewerbungsgespräche per Video, wertet gleichzeitig die Bewerber*innen nach unterschiedlichen Verfahren im Hinblick auf Effizienz, sog. „weiche" Fähigkeiten oder Persönlichkeitsmerkmale aus.	Viasto Interview Suite, Metaview, Cammio, Precire, Seedlink, HireVue Video interviewing software, Humantic	ja	-	-	x	x

Tabelle 1: ASR-Produktkategorien

2.4. Der Einsatz von algorithmischen Assistenzsystemen im Recruitingbereich: Marketingversprechen versus Realität

Der Begriff „Künstliche Intelligenz" (KI) wird in der Beschreibung von und Werbung für Softwarelösungen im Recruitingbereich ubiquitär verwendet. Kaum ein Werkzeug kommt heute in der Außendarstellung ohne den Begriff aus. In welchem Maße beispielsweise selbstlernende KI-Systeme oder eher einfachere ASR tatsächlich zum Einsatz kommen, lässt sich von Anwender*innenseite häufig schwer einschätzen. Genaue Erläuterungen dazu, welche konkreten Verfahren angewendet werden, findet man in den Marketing-texten kaum. Keines der von uns betrachteten ASR ist Open-Source, so dass es nicht möglich ist, ihren Quelltext zu untersuchen und sich so einen Eindruck des Funktions-umfangs zu verschaffen.

Die Liste der Versprechen, was die Softwarelösungen leisten können, ist lang. Sie erstreckt sich von Effizienzsteigerung im Bewerbungsprozess über De-Biasing der Ausschreibung und eine bessere Passgenauigkeit der eingehenden Bewerbungen bis hin zur Entdeckung geeigneter Kandidat*innen anhand von psychometrischen Echtzeit-Tests während des Bewerbungsgesprächs.

In digitalen Werkzeugen und Plattformen für Recruiting werden Maschinelle Lernver-fahren auf unterschiedliche Weise eingesetzt, z. B. in Form von *Natural Language Processing* (NLP), bei dem Computer „lernen", Text oder Sprache zu verstehen und/ oder zu manipulieren (grundlegend bei CHOWDHURY 2005). Ein Beispiel dafür sind Chatbots oder Werkzeuge, die Bias aus Daten filtern oder in Texten erkenntlich machen sollen.

ASR werden auch zur Optimierung von Prozessen eingesetzt, zum Beispiel um Bewerber*innen zu kategorisieren und zu matchen. Die Anbieter versprechen mit dem Einsatz von ASR effizientere, kostengünstigere und fairere Recruitingprozesse.[1] Auf den Webseiten finden sich Werbeslogans wie „Source faster than a human“[2], „Recruit while you sleep & wake up to interested candidates“[3], „Technology built to avoid unconscious bias“[4].

BOWER ET AL. (2017) zeigen in ihrem Papier mit dem Titel „Fair Pipelines“, dass die Anwendungen nicht immer ihren Versprechen gerecht werden. Ein Grund dafür sei, dass die Probleme manchmal deutlich komplexer sind als die vorgeschlagenen Lösungen. Häufig seien implizite Annahmen, dass automatisierte Entscheidungsfindung Zeit spare und Bias eliminiere, wenig überprüft.

Sie beschreiben zudem, wie sich wissenschaftliche Arbeiten zur algorithmischen Fairness auf die Fairness einzelner Entscheidung konzentrieren, nicht aber auf Prozesse, in denen viele Entscheidungen aufeinander aufbauen. Da Recruiting eine Abfolge von Entscheidungen darstellt, ist diese Art der Betrachtung aber essentiell, wenn man auf Gleichbehandlung und Fairness abzielt. (BOWER ET AL. 2017, 3)

Wie genau Unternehmen algorithmische Assistenzsysteme in den angebotenen Werkzeugen einsetzen und welche Daten sie für das Training ihrer Modelle heranziehen, oder ob sie extern erstellte Modelle einkaufen, erschließt sich aus den Produktvorstellungen der für diese Studie untersuchten Anwendungen nur im Ausnahmefall. Die Texte auf den Webseiten der Unternehmen sind Marketingtexte, die kaum Informationen über die genaue Technologie hinter verwendeten Buzzwords wie KI enthalten. Das ist kaum verwunderlich, da es sich bei den Technologien ggf. um Geschäftsgeheimnisse handelt, die gegenüber potenziellen Konkurrenten geschützt werden müssen.

Im Folgenden wird erläutert, für welche Schritte des Recruitingprozesses ASR zu welchem Zweck angeboten werden.

2.4.1. Sourcing

Für den Bereich Sourcing gibt es eine Vielzahl an ASR-Produkten. Sie helfen bei der Platzierung und Optimierung von Stellenanzeigen, benachrichtigen Arbeitssuchende

1 https://www.moberries.com/de/

2 https://goarya.com/

3 https://www.fetcher.ai/

4 https://hiretual.com/solutions/by-role/in-house-recruiter/

über potenziell attraktive Positionen und identifizieren geeignete Kandidat*innen. Mithilfe diverser Werkzeuge wird ein Kandidat*innenpool aufgebaut, der zu den Anforderungen des jeweiligen Unternehmens passt.

Stellenausschreibungen: Fast jede Kandidat*innensuche beginnt mit einem Ausschreibungstext. Dieser hat maßgeblichen Einfluss darauf, wer sich bewirbt. Je nachdem, wie die Ausschreibung formuliert ist, spricht sie bestimmte Personen an oder schreckt andere davon ab, sich zu bewerben. Studien zeigen, dass sich weniger Frauen auf eine Position bewerben, wenn in der Ausschreibung Wörter verwendet werden, die stereotyp maskulin sind (vgl. COLLIER & ZHANG 2016, 3).
ASR werden deshalb beim Verfassen von Stellenbeschreibungen eingesetzt, um ihre Sprache auf bestimmte Zielgruppen anzupassen und so induzierten Bias zu reduzieren. Ein Beispiel dafür ist die US-amerikanische Firma *Textio*. Sie bietet Instrumente an, um die Sprache in Stellenausschreibungen auf Diversität und Reichweite hin zu optimieren (HARRIS 2016). Das System lernt aus Millionen von Jobbeschreibungen und deren Erfolgsstatistiken und trifft damit Vorhersagen über die zu erwartende Größe und Demografie des Bewerber*innenpools. *Textio* hat ein „Gender-Barometer“, anhand dessen es Feedback zur Gender-Balance in der Sprache der Ausschreibung gibt (ANDREWS 2017). *Textio* ist aktuell nur für Texte in englischer Sprache erhältlich.
Die Software-Werkzeuge für gendergerechte Sprache sind nicht immer explizit für den Recruitingbereich designt. Für die deutsche Sprache bietet *Fairlanguage.com* eine App und eine Browsererweiterung, die alle Arten von Texten auf Gender-Bias hin untersucht und Änderungsvorschläge macht.[5] Neben Grammatik und Rechtschreibkorrektur bietet die Office-Webanwendung Microsoft 365 seit März 2020 auch Inklusivitäts-Einstellungen wie z. B. „Cultural Bias, Disability Bias, Gender Bias, Gender-Specific Language“ an. Das Werkzeug erkennt ungenaue Adjektive, Klischees und weitere Probleme mit der Wortwahl und bietet Alternativen an.[6]

Bewerbung der Ausschreibung: In einer Zeit vor der gezielten Werbung war die englische Phrase *post and pray* gebräuchlich. Sie bezeichnet das Teilen von Stellenausschreibungen auf Plattformen, bei denen wenig Einfluss darauf bestand, welche Menschen sie erreichen. Pray steht dabei für das „Hoffen“ darauf, dass die Ausschreibung die richtige Zielgruppe erreicht (JUNG 2017). In Zeiten von Datenanalyse,

5 https://fairlanguage.com/autokorrektur-tool-fuer-gendergerechte-sprache/

6 https://www.microsoft.com/de-de/microsoft-365/microsoft-editor?market=de

Maschinellen Lernverfahren und Microtargeting können Stellenausschreibungen viel gezielter in Netzwerke und an Kandidat*innen gespielt werden, die dem Profil mit größerer Wahrscheinlichkeit entsprechen und/oder für Stellenausschreibungen gerade empfänglich sind. Plattformen wie *LinkedIn* und *Facebook* bieten die Möglichkeit, die Zielgruppe exakt zu definieren und Personen und Gruppen, die nicht in das gewünschte Profil passen, gezielt auszusparen.

Matching: Um noch gezielter an passende Kandidat*innen heranzutreten, bieten Firmen wie *MoBerries* Dienstleistungen an, in denen sie gezielt Kandidat*innen für Stellenausschreibungen suchen (*matching*). Mit Hilfe von datengetriebenen Methoden werden laut Anbieter vorgefilterte Listen von Bewerber*innen erstellt. *MoBerries* verwendet die von Bewerber*innen angegebenen Informationen über Fähigkeiten und Präferenzen und erstellt damit Profile, die mit der Ausschreibung abgeglichen werden. Erst wenn die sogenannte „Matching-KI“ eine hohe Übereinstimmung feststellt, können sich Bewerber*innen die offene Position ansehen.[7]

2.4.2. Screening

In der Screening-Phase beginnen die Arbeitgeber*innen formell mit der Prüfung der Bewerbungen, wobei sie unqualifizierte oder relativ schwache Bewerber*innen ablehnen und die verbleibenden Bewerber*innen für eine nähere Prüfung priorisieren.
Bei der Profilanreicherung werden Daten aus Bewerbungen mit verschiedenen digitalen Identitäten (z. B. Profilen auf Recruiting-Plattformen) verglichen und die Echtheit von Zertifikaten und Abschlüssen verifiziert (VESTERLING 2018). Im internationalen Kontext leisten Anbieter wie *Veremark*[8] sogar eine Überprüfung der Vorstrafen, der Kreditwürdigkeit und Rechtsverfahren, an denen Kandidat*innen beteiligt sind oder waren. Für die Integration von Daten und Inhalten externer Plattformen sind Programmierschnittstellen essentiell. Über sie können verschiedene Anwendungen automatisiert Daten austauschen, die für die ASR-Werkzeuge benötigt werden. *Orange Tree,* ein in den Vereinigten Staaten ansässiges Unternehmen, das aber für den Privacy Shield zertifiziert ist und

7 https://www.moberries.com/de/

8 https://www.veremark.com/

unter anderem Integrationen für *SAP Success Factors* anbietet, führt Hintergrundüberprüfungen der Bewerber*innen u. a. zu Vorstrafen, Gesundheit oder Fahreignung sowie die Gesundheits- und Drogenchecks durch.[9]
Screening-Werkzeuge bewerten und ordnen außerdem die Bewerber*innen nach ihren Qualifikationen, ihren Soft-Skills und anderen Fähigkeiten ein (*matching* oder *culture fit)*. Diese Einordnung unterstützt Einstellungsmanager*innen bei der Entscheidung, wer in die nächste Phase des Prozesses kommen soll. Mit Hilfe von Screening-Werkzeugen können Arbeitgeber*innen den Pool an Bewerber*innen schnell verkleinern, so dass sie mehr Zeit für die Prüfung der als am besten geeigneten Bewerber*innen aufwenden können.
Als Screening-Werkzeuge dienen z. B. Chatbots, wie sie von *AllyO*[10] oder *Mya*[11] angeboten werden. Diese Bots verwenden NLP, um mit Jobsuchenden zu interagieren und erste Fragen per Chat, Textnachricht oder E-Mail zu beantworten (LI 2016).

2.4.3. Auswahlverfahren

Die Phase der Vorstellungsgespräche ist die Phase, in der Bewerber*innen und potenzielle Arbeitgeber*innen direkt miteinander interagieren. Neben klassischen Interviewsettings, in denen die Gespräche bei einem persönlichen Treffen stattfinden, gibt es auch die Möglichkeit, Interviews über Videotelefonie durchzuführen. Anbieter wie *Metaview*[12] helfen Interviewer*innen dabei, ihre Interviewfragen zu verbessern, indem sie Gespräche aufzeichnen, automatisiert transkribieren und analysieren. Sie weisen Interviewer*innen im Nachgang auf besonders gute Fragen und Interaktionen hin. Plattformen wie *HireVue*[13] gehen einen Schritt weiter und geben an, die Leistung von Bewerber*innen anhand der Antworten, ihres Tonfalls und ihrer Gesichtsausdrücke zu analysieren. Die Versprechen dieser Werkzeuge an Arbeitgeber*innen sind vor allem Zeitersparnis und eine Standardisierung von Prozessen.

9 https://www.orangetreescreening.com/

10 https://www.allyo.com/recruiting-ai-and-automation-platform-allyo-recruit/

11 https://www.mya.com/

12 https://www.metaview.ai/

13 https://www.hirevue.com/

2.5. Diskriminierungspotenziale im Bewerbungsprozess

Um im Detail darzustellen, an welchen Schritten des Recruitings es zu Diskriminierung kommen kann, lohnt es sich, die Perspektive zu wechseln und aus Sicht der Bewerber*innen die Berührungspunkte mit ASR während des Bewerbungsverfahrens herauszuarbeiten.
Das Diskriminierungspotenzial von ASR in den verschiedenen Stufen des *Hiring Funnels* ist unterschiedlich. Ihnen ist jedoch gemein, dass sie in der Regel Ablehnungen automatisieren, nicht Zusagen (BOGEN & RIEKE 2018).
Durch diese Negativauslese erhalten algorithmische Verfahren eine breite Wirkung, bei denen Einzelfälle nicht betrachtet werden, da sie früh im Bewerbungsprozess erfolgt und die Bewerber*innen noch nicht im direkten Austausch mit den zukünftigen Arbeitgeber*innen stehen. Bewerber*innen haben deshalb kaum die Möglichkeit, strukturelle Diskriminierung zu erkennen. Gleichzeitig ist der Großteil der Bewerber*innen davon betroffen, so dass bei einer Fehlfunktion die Auswirkungen enorm sind.
Im Folgenden schlüsseln wir für die einzelnen Schritte des Einstellungsprozesses auf, welche Faktoren dazu führen können, dass ASR selektiv und potenziell diskriminierend wirken.

2.5.1. Sourcing

Wer sieht die Stellenanzeige?

Aus Sicht der Bewerber*innen ist der entscheidende Punkt, ob eine Bewerbung überhaupt möglich ist, ob sie eine Stellenanzeige auf einer Plattform sehen oder nicht. Zwei Faktoren kommen hier vor allem zum Tragen: Einerseits der Zugang zu Plattformen, der durch technische Fähigkeiten (*digital literacy*), zur Verfügung stehende Geräte und digitales Nutzungsverhalten limitiert wird, und andererseits die selektive Art, im digitalen Raum zu werben. Vor allem der regelmäßige Zugang zu digitalen Plattformen ist in Deutschland immer noch abhängig vom ökonomischen Status: Während nur 80 % der Haushalte mit einem Nettoeinkommen unter 1.500 € über einen Internetanschluss verfügen, sind es bei den Haushalten mit über 3.600 € Nettoeinkommen 99 % (DESTATIS 2020, 10). Außerdem können Menschen auf Arbeitssuche mit einem kostenpflichtigen Premium-Account einsehen, wer ihre Profile über welche Schlagwörter besucht hat, und ihr Profil entsprechend anpassen, um die eigenen Chancen zu erhöhen. Auch hier wirkt der ökonomische Status unter Umständen begünstigend.

Die Werbeplatzierung wird häufig durch einfache Algorithmen vorgenommen, die auf teils detaillierten Datenmodellen beruhen (COARD 2020). Wie das Beispiel *Targeting* auf *Facebook* zeigt, kommen dort bei der Auswahl, wer welche Anzeigen zu sehen bekommt, nicht nur gängige Faktoren wie Alter und berufliche Qualifikationen zum Tragen, sondern hochgradig individuelle Eigenschaften: Ob eine Person Haustiere hat, Yoga macht, Alkohol trinkt oder kurz vor der Hochzeit steht – eine umfangreiche Interessenliste kann hier als Filter herangezogen werden. In ihrer Gesamtheit stellen diese Interessen ein komplexes Gesamtbild dar, das Unterschiede zwischen den Geschlechtern, aber auch Altersunterschiede und soziale Herkünfte detailliert abbildet, ohne dass nach dem eigentlichen Datenpunkt (z. B. Wahrscheinlichkeit für baldige Schwangerschaft, Migrationshintergrund, Suchtgefahr) gefragt werden muss. Ansätze für Diskriminierung sind tief im Datenmodell verborgen und als Proxy-Diskriminierung im Einzelnen kaum nachzuweisen (ORWAT 2019, 87; PRINCE & SCHWARCZ 2020, 1273). Auch wenn *Facebook* nicht primär eine Recruiting-Plattform ist, ist es bei Recruiter*innen deutlich beliebter als andere soziale Netzwerke: 76,7 % der befragten Recruiter*innen gaben 2015 an, *Facebook* zu verwenden, dagegen nutzten nur 54,6 % das eigens dafür geschaffene Berufsnetzwerk *LinkedIn* (ONLINE-RECRUITING.NET 2015).

Wer fühlt sich angesprochen?

Eine inklusiv formulierte Stellenanzeige, die unter Umständen mit algorithmischen Werkzeugen der Textanalyse erstellt wurde, um eine inklusive Sprache zu verwenden, ist also durch die Platzierung auf unterschiedlichen Plattformen oder die Ausspielung über verschiedene Datenmodelle nicht für alle sichtbar, sondern nur für ein Subset an Personen – ob diese Auswahl intendiert ist oder nicht. Anders verhält es sich bei aktiver Suche nach Bewerber*innen, die heute ebenfalls überwiegend auf digitalen Plattformen stattfindet. Hier fehlen jedoch umfassende Daten dazu, welche Personengruppen auf welchen Plattformen vertreten sind und welche Daten sie dort über sich preisgeben, insbesondere, wenn die Plattformen nicht direkt für die Arbeitssuche ausgelegt ist (z. B. *Facebook*, aber auch spezialisierte *Slack*-Instanzen, z. B. für Recruiting im IT-Bereich). Geben Personengruppen weniger Daten über sich preis, ist es möglich, dass sie von den Suchalgorithmen und Filtern der Recruiter*innen weniger oft gefunden werden. Auch fehlen Daten darüber, welches Spektrum an Stellen – Niedriglohntätigkeiten, prekäre oder reguläre Anstellungen, Leitungspositionen – über welche Plattformen angeboten wird.

2.5.2. Screening

In der zweiten Phase, dem Screening, wiederholen sich die genannten Filter im Schritt der Profilanreicherung. Personen, deren Daten zur Verfügung stehen, sind tendenziell im Vorteil. Deshalb ist es auch hier möglich, dass Personengruppen, die im Netz weniger Daten von sich preisgeben oder vom Standard abweichende Nutzungsverhalten zeigen, schlechter erfassbar oder gar nicht verifizierbar sind.
Ein weiterer Schritt ist das sogenannte *Skill Matching*[14], in dessen Zuge Bewerber*innen auf Basis ihrer Bewerbungsdaten darauf überprüft werden, wie gut ihre Fähigkeiten und Persönlichkeit zum Unternehmen passen. In den von uns betrachteten Tools (z. B. *MoBerries* oder *Eightfold*) werden weder das Qualitätsmaß definiert noch offengelegt, welche technischen Verfahren dabei zum Einsatz kommen – vage Beschreibungen wie „Algorithmische Verfahren", „mit Software" oder „mit Robot Recruiting" sind die Regel. Inwieweit für diese Produkte tatsächlich Persönlichkeitsmerkmale wie Kommunikations- und Teamfähigkeiten analysiert werden, ist unklar. Sollte sich dieses häufige Verkaufsversprechen jedoch bewahrheiten, liegt hier ein enormes Diskriminierungspotenzial. Psychologische Modellierungsverfahren für Wesenszüge sind, ebenso wie früher Lügendetektortests, im Personalbereich weit verbreitet, aber seit Jahrzehnten methodisch (z. B. der im Personalbereich besonders im US-amerikanischen Raum weit verbreitete Myers-Briggs-Test, vgl. GRANT 2013) und rechtlich umstritten (vgl. dazu HERMANN 1971). Wenn sie digital implementiert werden, ist noch schwerer durchschaubar, welche Eigenschaften zu welcher Klassifizierung führen. Ähnlich verhält es sich mit dem sogenannten *Culture Fit*, also der Überprüfung, ob sich ein*e Bewerber*in auch soziokulturell gut in die Belegschaft einfügt.

2.5.3. Auswahlverfahren

Kandidat*innenmanagement

Nach dem Screening erfolgt ein ausgedehnter Kommunikationsprozess zwischen Recruiter*in/HR-Manager*in und den verschiedenen Bewerber*innen, um Bewerbungsgespräche und Assessments zu planen. Auch in diesem Teil des *Hiring Funnels* versprechen digitale Werkzeuge, die überwiegend Recruiter*innen oder Personalmanager*innen adressieren, effizienteres Arbeiten durch den Einsatz von ASR: Zum

14 https://www.careerplus.ch/blog/jobmatching-und-skillmatching

Beispiel in Form von Chatbots und Sprachassistenzsystemen, die Kandidat*innen administrative Fragen beantworten oder Termine für Bewerbungsgespräche verhandeln. Kaum ein Werkzeug, das mit dem Versprechen „KI“ wirbt, verfügt über eine hohe algorithmische Komplexität; wir vermuten, dass die meisten dieser Angebote zwar möglicherweise Modelle nutzen, die mit KI-Algorithmen entwickelt wurden, aber keine Feedbackschleifen besitzen und sich deshalb auch nicht an das jeweilige Einsatzgebiet anpassen. Die Werkzeuge sind außerdem durchweg administrativer Natur und führen nicht zu einem weiteren Aussieben des Bewerber*innenpools. Mögliche Diskriminierung kann aber z. B. dann entstehen, wenn die Werkzeuge nicht barrierefrei nutzbar sind.

Bewerbungsgespräche

Im letzten Schritt des Bewerbungsprozesses finden die Bewerbungsgespräche statt. Hier ist eine breite Palette an Produkten entstanden, die diesen Schritt unterstützen sollen. Ähnlich wie schon im Screening konzentrieren sie sich vor allem auf *Skill Matching* und *Culture Fitting*, analysieren also Daten der Bewerber*innen nach einem definierten Modell daraufhin, ob die Person in ihren Fähigkeiten das Team gut ergänzt, und prognostizieren, ob sie sich sozial in das Team einfügen wird. Es gibt aber eine Reihe von Unterschieden: Anders als beim Screening ist die Datengrundlage vielfältiger, da nun Audio- und Videostreams oder -mitschnitte von Bewerbungsgesprächen analysiert werden oder sogar zugeschnittene Assessmentaufgaben entwickelt und nach behavioristischen Modellen auf Effizienz, Fähigkeiten und Persönlichkeitszüge ausgewertet werden: die Lautsprache gibt Aufschluss über Bildung, Selbstbewusstsein und Kommunikationsqualitäten; die Körpersprache soll Unsicherheiten oder Unwahrheiten im Gespräch erkennen und den Interviewer*innen entsprechendes Feedback in Echtzeit geben. Die Ergebnisse von Assessmentaufgaben werden nicht nur fachlich analysiert, sondern auch „Persönlichkeitszügen“ nach unterschiedlichen psychologischen Modellen zugewiesen.

Wie schon bei den ASR-Produkten, die das Screening auf ähnliche Weise unterstützen, sind das Qualitätsmaß, die konkreten technischen Verfahren und die zugrundeliegenden psychologischen Persönlichkeitsmodelle nach außen nicht zu erkennen. Analysen aus dem englischsprachigen Raum zeigen, dass beispielsweise Verfahren aus dem Bereich des NLP dazu tendieren, afroamerikanisch klingende Namen weniger positiv zu konnotieren als europäisch-amerikanische Namen, und männliche Pronomen mit „maths and sciences“ verbinden, weibliche Pronomen stärker mit dem Begriff „arts“ (SUTTON ET

AL. 2018, 4). Wenn NLP-basierte ASR Interviewsituationen analysieren und Anhaltspunkte für die Eignung von Kandidat*innen für eine Stelle liefern, fallen solche Biases enorm ins Gewicht.
Auch hier zeigen sich die möglichen Effekte komplexer Diskriminierungsmuster. Für Sprachen mit grammatischem Geschlecht, wie das Deutsche, ergeben sich weitere Interferenzen zwischen grammatikalischem Bias und Sachbias (*topical bias*), die bei der Bewertung einzelner NLP-Modelle untersucht werden müssen (MCCURDY & SERBETÇI 2020, 2–3). Studien, die sich konkret auf den deutschsprachigen Raum und den Kontext von Recruiting beziehen, fehlen bisher. Deshalb konstatieren wir für diese Produktkategorien ein besonders hohes Diskriminierungspotenzial, das durch die vermeintliche Wissenschaftlichkeit der propagierten Verfahren verschleiert wird (vgl. Abschnitt 3.).

2.6. Geschäftsmodelle

In der Recherche für diese Expertise konnten wir mehrere Trends und Kategorien von Geschäftsmodellen und Märkten beobachten. Einer der dominantesten Trends ist die Plattformisierung und die Tatsache, dass wenige große Plattformen den Recruitingmarkt dominieren und viele Dienstleister auf ihre Daten und Services aufbauen. Das bedeutet, dass viele Plattformen miteinander verbunden sind oder Daten austauschen.

- Social-Media-Plattformen wie *Facebook*, *LinkedIn* oder *Xing* sind Dreh- und Angelpunkt von Sourcing und Matching im Recruitingbereich.
- Viele *Applicant Tracking Systems* (ATS) und Plattformen für *Human Capital Management* (HCM) integrieren Social-Media-Plattformen und deren Daten.
- Es gibt eine Reihe von Anbietern, deren Plattformen sowohl Integrationen zu ATS als auch zu sozialen Medien und Recruitingnetzwerken haben.

2.6.1. Plattformisierung

Vor allem im Sourcing führt fast kein Weg an den großen Social-Media-Plattformen vorbei. Plattformen wie *Facebook*, *LinkedIn* und *Xing* haben Millionen von Nutzer*innen und sind mit ihrer Reichweite und Größe prädestiniert für Sourcing von Kandidat*innen. Das Kerngeschäft der Plattformen sind Suchalgorithmen, die mit maschinellen Lernverfahren kontinuierlich angepasst werden. Die Plattformen haben nicht nur eine große Anzahl an Nutzer*innen, sondern vor allem deren Daten, die notwendig für die Anwendung von ASR sind. Auch über das Sourcing hinaus bieten einige von ihnen digitale

Werkzeuge für das Recruiting an. *Facebook Jobs* bietet Unternehmen neben dem Posten von Stellenanzeigen auch Werkzeuge an, über die sie mit Bewerber*innen kommunizieren und Bewerbungen managen können.[15] *LinkedIn Recruiter* wirbt damit, dass man in den über 690 Millionen *LinkedIn*-Nutzer*innen schnell die richtigen Kandidat*innen für eine Ausschreibung findet. Das Geschäftsmodell der Plattformen bezieht sich dabei nicht ausschließlich auf Recruiter*innen und Unternehmen, auch für Jobsuchende gibt es kostenpflichtige Dienstleistungen.[16] Recruiter*innen und Jobsuchende können Verträge für unterschiedliche Premium-Accounts abschließen, die ihnen Zugriff auf weitere Werkzeuge geben. Für Recruiter*innen gibt es eine Reihe an Filteroptionen, darunter Algorithmen, die beim Matching helfen, und Werkzeuge für das Talent-Pipeline-Management. Jobsuchende bekommen mit dem Premium-Account Zugriff auf zusätzliche Profilfunktionen und können besser nach passenden Unternehmen suchen.[17]
Die Social-Media-Plattformen sind so mächtig, dass fast alle Anbieter von digitalen Recruiting-Lösungen sie integrieren und nutzen.
Eine weitere Plattformkategorie bilden sogenannte Applicant-Tracking-Systeme (ATS) und Human-Capital-Management-Plattformen (HCM). Unternehmen wie *Oracle*, *Cornerstone*, *Bullhorn* und *Greenhouse* bieten Softwarelösungen für ATS und HCM an. Diese Systeme sind auf die Recruiting- und HR-Abläufe der jeweiligen Kunden abgestimmt. Auch sie bieten Integrationen zu Plattformen wie z. B. *LinkedIn*. Mehrere der oben genannten Unternehmen sind auf der *LinkedIn*-Webseite als Partner aufgeführt.[18]
Ein dritter Typ von Anbietern wie *XOR* [19] und *MoBerries*[20] integriert sowohl Social-Media-Plattformen sowie ATS und HCM-Systeme. *MoBerries* bietet beispielsweise ein „Talentnetzwerk“, in dem es Jobportale, Soziale Netzwerke, Recruitingagenturen und externe Talentnetzwerke zusammenführt. Kund*innen können damit an einer zentralen Stelle ihre Ausschreibungen über viele Plattformen auf einmal ausspielen. *MoBerries* kann auch in bestehende Bewerbungsmanagementsysteme integriert werden. Die Dienstleistungen, die es ermöglichen, auf mehreren Plattformen von einer Stelle aus zu suchen und zu inserieren, werden Unternehmen und Jobsuchenden zugleich angeboten.

15 https://www.facebook.com/business/news/take-the-work-out-of-hiring

16 https://business.linkedin.com/talent-solutions/recruiter

17 https://premium.linkedin.com/content/premium/global/en_us/index/jobsearch

18 https://business.linkedin.com/talent-solutions/ats-partners#all

19 https://www.xor.ai/

20 https://www.moberries.com/de/

2.6.2 Weitere Beobachtungen

Neue Werkzeuge und Dienstleistungen werden von Unternehmen häufig in großen Märkten wie den USA oder für den angelsächsischen Raum erprobt, bevor sie auf andere Märkte und damit auch in andere Sprachen übertragen werden. Ein Beispiel dafür ist das Inclusiveness-Feature von *Microsoft 365*, das zuerst auf Englisch angeboten und 2020 auf weitere Sprachen ausgeweitet wurde.[21]
Startups fokussieren sich auf Nischen und experimentieren mit neuen Ansätzen und Technologien. Das Startup *Scoutible*[22] verwendet immersive Computerspiele, um geeignete Kandidat*innen für Stellen zu finden. Dabei versuchen sie, kognitive Persönlichkeitsmerkmale der Spieler*innen durch deren Spielverhalten zu identifizieren und gleichen diese mit Eigenschaften von Top-Performer*innen des jeweiligen Unternehmens ab.[23] Das 2016 gegründete Unternehmen *Headstart*[24] hat sich auf Diversität und auf Rekrutierungssoftware für Hochschulabsolvent*innen und Berufseinsteiger*innen spezialisiert. Das Startup *Vetty*[25] setzt laut eigener Beschreibung Deep-Learning-Bots und Blockchain-Technologie ein, um Identitäts- und Hintergrundchecks von Kandidat*innen durchzuführen. Das in Berlin ansässige Startup *Bunch*[26] erstellt auf Basis von kurzen Online-Fragebögen ein Profil über die Verhaltenstendenzen von Kandidat*innen und ermittelt damit, wie gut sie kulturell zu einem Unternehmen passen.

21 https://www.microsoft.com/de-de/microsoft-365/microsoft-editor?market=de

22 https://www.scoutible.com/

23 https://www.scoutible.com/about-us

24 https://www.headstart.io/

25 https://vetty.co/

26 https://bunch.ai/

3. Methodologische Vorüberlegungen und Ansätze für Taxonomien

Wie oben aufgeführt verdient das Feld von ASR-Recruiting und Gleichstellung eine differenzierte Betrachtung, in die eine Reihe unterschiedlicher wissenschaftlicher Disziplinen und Ansätze einfließen müssen. Um die Bandbreite der möglichen Perspektiven darzustellen, entwerfen wir im folgenden verschiedene Ansätze für Taxonomien.
Intersektionale Diskriminierung ist ein fließender Prozess, der Gesellschaften und die darin sich befindenden Kollektive strukturiert. Diese Strukturierung ist das Resultat sozialer, ökonomischer, historischer und politischer Macht-Asymmetrien. Durch diese Strukturierung werden die Formen des Zugangs und der Verteilung von Macht, Teilhabe, Gemeinwohl und Handlungsmöglichkeiten festgelegt. Gesellschaftliche Asymmetrien sind nicht determiniert, sondern kontingent (vgl. BINDER & HESS 2011). Die Position einer christlichen Frau sieht je nach Land, Kontext und historischem Zeitpunkt anders aus als die Position einer Muslima. Die Positionen sind relational. Veränderungen einer Position in der Struktur verändern infolgedessen auch die Positionen der anderen in der Struktur. Ansätze, um Diskriminierung in Kontexten zu evaluieren, in denen algorithmische Systeme angewendet werden, müssen diesen prozessualen Charakter von Diskriminierung berücksichtigen und einen strukturbezogenen Blick auf den Gesamtprozess werfen. Intersektionale Genderpositionen sind im algorithmischen Kontext an weitere Attribute korreliert, so genannte relationale Daten. Diese können bekannt sein, wie etwa Körperform, Kleidung, familiäre Status, Haare, oder auch weniger bekannt. Der prozessuale Blick hilft zu verstehen, welche intersektionale Genderprofile in einem algorithmischen System diskriminiert werden, welche (unerwartete) relationale Daten als Indikator einer Genderasymmetrie gelten können, und welche Programmier-Aspekte in ihrer Kombination zur Diskriminierung führen.
Im folgenden Kapitel werden Einstiegspunkte für die Einführung von Verzerrungen und Asymmetrien in sozio-technischen Systemen aufgelistet. Darüber hinaus werden Kategorien identifiziert, mit denen Technologien im Recruiting als hybride menschlich-maschinelle Prozesse und Ökosysteme in Ansätzen systematisiert werden könnten.

3.1. Verortung der Ursachen für Diskriminierungen

Im Folgenden versuchen wir eine Aufschlüsselung darüber, worin das Diskriminierungspotenzial im Einzelnen begründet ist.

3.1.1. Optimierungsziel

Algorithmische Systeme werden mit einer spezifischen Zielrichtung gebaut. Sie sind die Antwort auf die Frage: Was soll im Prozess optimiert werden? Bereits bei der Identifizierung des Sachverhaltes, der optimiert werden soll, oder bei der Übersetzung dieses Sachverhalts in einer algorithmischen Äquivalenz kann Bias entstehen. So macht es einen Unterschied, ob Systeme eingesetzt werden, um die Anzahl von weiblichen Bewerbungen, oder eher die Anzahl der weiblichen Mitarbeiterinnen, oder der weiblichen Führung zu erhöhen. Eine Erhöhung der Anzahl von weiblichen Bewerberinnen durch algorithmische Systeme führt nicht unmittelbar zu mehr weiblichen Angestellten. Eine Erhöhung der Anstellungszahlen von weiblichen Mitarbeiterinnen bedeutet nicht unmittelbar langfristig eine Erhöhung von weiblichen Angestellten, wenn nicht entsprechende weitere Maßnahmen ergriffen werden (bspw. Mutterschutz, Förderung etc.). Punktuell an einem Schritt des Anstellungsprozesses zu optimieren, ersetzt nicht weitere Maßnahmen an allen anderen Stellen.

3.1.2. Datengrundlagen

Die zunehmende Bedeutung von Plattformen im Recruiting (vgl. Abschnitt 2.6.1.) hat zur Folge, dass die An- oder Abwesenheit einzelner Personengruppen auf den jeweiligen Plattformen bereits ein Filter darstellt. Auch wenn die meisten Plattformen kostenfrei zumindest mit Basisfunktionalitäten nutzbar sind, sprechen sie unterschiedliche Nutzer*innengruppen an. Allein die digitale Natur der Plattformen kann dazu führen, dass wenig technikaffine Personengruppen (beeinflusst z. B. von Alter, sozialem Hintergrund oder Bildungsstand, aber auch Gender) eine Stellenanzeige deutlich seltener zu Gesicht bekommen. Eine detaillierte, öffentlich zugängliche Demographie der Nutzer*innen der einzelnen Plattformen wäre für eine weitere Betrachtung wünschenswert.
Dabei ist zu beachten, dass nicht alle Plattformen ausschließlich dem Recruiting dienen. Soziale Netzwerke wie *Facebook* werden oft weniger zur beruflichen als zur privaten Kommunikation verwendet. Die detaillierten Filtermöglichkeiten bei *Facebook* können dahingehend wirken, dass Nutzungs- und Kommunikationsverhalten der Mehrheitsgesellschaft gegenüber dem migrantischer Communities (z. B. Mehrsprachigkeit) bevorzugt werden. Hierfür fehlt jedoch der Zugang zur Datengrundlage der *Facebook*-Interessensfilter einerseits, und andererseits, wie diese von Unternehmen bei Stelleninseraten verwendet werden.

3.1.3 Format

Nicht alle Bewerber*innen haben dieselben Präferenzen bei der Nutzung von digitalen Formaten. Laut einer Studie der Manpower Gruppe 2017 (MANPOWER 2017, 9) fühlen sich fast zwei Drittel der Frauen weltweit (58 %) unwohl mit Video-Interviewformaten (z. B. *Skype*) oder der Bereitstellung einer Video-Einführung über ihre Berufserfahrung und Karriere. Im Gegensatz dazu, fühlt sich mehr als die Hälfte der Männer mit dieser Technologie sehr wohl. 56 % der Männer gab an, Interview über Videokonferenztechnologien zu bevorzugen. Zudem sind nicht alle Formate gleich zugänglich für alle Bewerber*innen (z. B. aufgrund von Behinderung, Alter, notwendigen Betriebssystemen, Hardwarenanforderungen und Bandbreite).

3.1.4 Hardware und Sensorik

Auch Hardware kann keineswegs als neutrales Artefakt betrachtet werden. Ein kleines, aber bekannteres Beispiel für Diskriminierungseffekte in diesem Bereich ist ungeeignete oder schlecht getestete Sensorik in Hardware-Technologien, die etwa dunklere Haut schlechter reflektiert oder schlecht dafür kalibriert ist. Die Kalibrierung erfolgte traditionell durch sogenannte Shirley Cards. Wie Jan Kasoff, ein ehemaliger Kameramann des Fernsehsenders NBC, berichtet, erfolgt der Farbabgleich einer Fernsehkamera in mehreren Stufen. Zwei Balkenkarten – eine in Schwarzweiß, die die Graustufen zeigt, die andere in Farbe – werden nacheinander vor die Kameras gelegt, um die Genauigkeit und Sättigung der Luminanz- und Farbdarstellung zu messen. Die resultierenden Videosignale werden auf einem Wellenform-Monitor und einem Vektorskop eingestellt, um Verzerrungen zu eliminieren. Anschließend werden die Studiokameras auf die Gleichmäßigkeit der Farbdarstellung abgestimmt. Die Techniker verwenden das menschliche Auge für einen subjektiven Farbtest, basierend auf der Annahme, dass das menschliche Auge Farben anders wahrnimmt als ein Vektorskop. Laut Jan Kasoff, ist es so, dass "a good VCR person will have a colour girl stand in front of the camera and stay there while the technicians focus on her flesh tones to do their fine adjustments to balance the cameras. This colour girl is always white." (ROTH 2009, S. 12).
So setzt sich ROTH (2009) historisch mit der Entwicklung von Belichtungstechniken, Sexismus, Rassismus, deren Intersektionalität und kognitiver Ungerechtigkeit auseinander, die nicht nur die Entwicklung von nicht-repräsentativen Datenbanken, sondern auch die Entwicklung von Hardware beeinflussen. Auch PLENKE (2015) zeigt, warum bestimmte Infrarot-Hardware nicht in der Lage ist, dunklere Hauttypen zu erkennen.

Auch die Entwicklung von Robotern und deren Einsatz in Recruiting Situationen erfordert eine Überprüfung in Bezug auf Gendergerechtigkeit: Sind die Roboter gegendert? Haben sie eine menschenähnliche Körperform mit gegenderten Attributen?
Auch in Falle von Kiosks oder Terminals können Unterschiede im Zugang entstehen, die, zusammen mit anderen Faktoren, mit Gender korrelieren. So ist bspw. die Aufnahme eines Video-Bewerbungsgesprächs in einem Kiosk für Frauen im Rollstuhl grundsätzlich schwierig, nicht nur aufgrund des Formates (siehe Abschnitt 3.1.3), sondern auch aufgrund der Hardware, wenn die Kamera des Kiosks in einer Durchschnittshöhe fixiert ist, die Menschen in Rollstühlen außer Acht lässt. In diesem Beispiel sehen wir, dass die Kombination eines von vorherein diskriminierenden Formats und der zusätzlichen Nutzung einer bestimmten Hardware auf Frauen in Rollstühlen noch komplexer diskriminierend wirkt.

3.1.5 Sichtbarkeit

Menschen mit unterschiedlichen Demographien nutzen digitale Werkzeuge und insbesondere öffentliche Soziale Netzwerke unterschiedlich. Für von Stalking betroffenen Personen (von denen 88 % Frauen sind, BUNDESKRIMINALAMT 2019, S. 6) kann das bedeuten, dass sie ihre Arbeitsbiografie nicht online mitteilen und generell weniger Daten über sich preisgeben oder ihr Profil nicht öffentlich machen. Im ersten Fall kann das bedeuten, dass ihren Profilen nur Anzeigen mit einem vagen Filter angezeigt werden, während Personen mit detailliertem Profil und intensivem Nutzungsverhalten auch die stärker eingeschränkten Anzeigen zu sehen bekommen.

3.1.6 Kriterien-Angabe

Viele ASR, sowohl im Screening als auch bei der Analyse der Interviews, versprechen, ein *Matching* durchzuführen, also die Kandidat*innen daraufhin zu überprüfen, wie gut sie zum Unternehmen und zur ausgeschriebenen Stelle passen (vgl. Abschnitte 2.4.2 und 2.5.2.). Dies geschieht entweder über eine Analyse der Fähigkeiten (*Skill Matching*) oder eine Untersuchung der kulturellen Ähnlichkeit (*Culture Fit*) zur Belegschaft des Unternehmens. Beide Ansätze sind aus Gleichstellungsperspektive problematisch.
Beim *Skill Matching* werden auch abstrakte Fähigkeiten untersucht, beispielsweise Führungsqualitäten. Diese für einen analytischen Abgleich zu modellieren ist schwierig; die für die vorliegende Expertise untersuchten Produkte verwenden bspw. das OCEAN Modell um anhand von Selbst-Einschätzungen festzustellen, welche Eigenschaften gute

Führung ausmachen. Das OCEAN Modell arbeitet mit stark gegenderten Emotionen und Charaktereigenschaften. Sie dienen der Beurteilung nicht fachlicher Fähigkeiten, sondern der emotionalen Verfasstheit (vgl. Abschnitt 4.2.).
Noch offensichtlicher problembehaftet ist der *Culture Fit.* Dabei wird abgeglichen, inwieweit sich Verhalten, Ziele, Motivationen und Hintergründe von Bewerber*innen mit denen der Belegschaft einer Firma decken. Diese Herangehensweise führt mit hoher Wahrscheinlichkeit zu mehr Homogenität und verhindert aktiv Diversität. Durch die Szientifizierung und Digitalisierung der Verfahren wird es jedoch kaum möglich, eventuelle Diskriminierungen zu erkennen oder gar zu vermeiden.
Auch Genderneutralität fällt unter die Kriterien, die überprüft werden müssen. Verfahren, die auf die Verschleierung des Geschlechts und anderer Merkmale setzen, normalisieren die Marginalisierung von bestimmten Merkmalen, anstelle diese durch Anerkennung zu inkludieren.
Genderneutralität und genderneutrale Sprache sind sowohl in der Wirklichkeit als auch technisch schwer implementierbar. Gender und andere Merkmale können durch andere relationale Informationen inferiert werden (vgl. Abschnitt 3.2.1.). Auch ist hier zu hinterfragen, wer sich von einer genderneutralen – anstelle einer genderinklusiven bzw. gendergerechten – Sprache angesprochen fühlt (HEISE 2000).

3.1.7 Konzeptualisierung des Systems

Die Prämissen und Theorien, auf deren Grundlage automatisierte Recruiting-Systeme gebaut werden, können ebenfalls Bias im Verfahren integrieren. Dies geschieht beispielsweise durch die Mathematisierung scheinbar wissenschaftlicher Annahmen.
Viel beworben, aber höchst umstritten ist der Einsatz von psychologischen Modellierungsverfahren oder *Psychometric Tests* im Bewerbungsprozess. Diese kamen zunächst vor allem in den USA verstärkt zum Einsatz, nachdem der *Employee Polygraph Protection Act* 1988 den Einsatz von Lügendetektoren in Bewerbungsgesprächen weitgehend untersagte. Die psychometrischen Tests setzen diese Idee jedoch fort: Bewerber*innen werden vermeintlich durchleuchtet und ihre Neigungen, Interessen, Stärken und Schwächen offengelegt.
Bei den zugrundeliegenden Tests handelt es sich zum einen um Intelligenztests: Diese werden dafür kritisiert, Intelligenz als solche anhand von Faktoren normativ festzuschreiben, die einer bestimmten Denk- und Bildungstradition entsprechen und insofern nur in einem konkreten kulturellen Kontext Bestand haben. Kulturelle Unterschiede werden so

unter Umständen als Unterschiede in der kognitiven Leistungsfähigkeit interpretiert (umfänglich zusammengefasst bei FORD 2004).
Zum anderen sind psychometrische Persönlichkeitstests weit verbreitet. Der genannte *Employee Polygraph Protection Act* führte dazu, dass sie statt der Lügendetektortests zunehmend eingesetzt wurden.
Zu den beliebtesten psychometrischen Tests im Personalmanagementbereich gehören der *Myers Briggs Type Indicator* (benannt nach seinen Erfinderinnen Katharine Cook Briggs und Isabel Briggs Myers), und die darauf basierenden *Big Five* und *DISC Behavior Inventory Tests*, die Wesenszüge wie Dominanz, Offenheit und Urteilskraft abbilden sollen, gemeinhin aber eher als pseudowissenschaftlich gelten (vgl. GRANT 2013, SPILLANE 2012). Auch hier gilt die Kritik, dass diese Persönlichkeitsmodelle in einem weiß und angelsächsisch geprägten Umfeld entstanden sind und bestenfalls dessen soziale Ordnung widerspiegeln, aber keine Allgemeingültigkeit besitzen. Durch ihren Einsatz wird also kulturelle Diversität ebenfalls benachteiligt.
Die Digitalisierung dieser Verfahren verleiht ihnen einerseits eine höhere Glaubwürdigkeit und verschleiert andererseits die eingesetzten Verfahren.
In Deutschland gehören Culture-Fit-Tests bei 15 % der Unternehmen zum Auswahlprozess. Dazu kommen bei ca. 10 % der Unternehmen noch Gesichts- und Sprachanalyse-Werkzeuge (STEPSTONE 2020, ausführlich dazu Kapitel 3.2.2.)
Konzepte, die dazu dienen, ein gesamtes Verfahren zu erfassen, können reduktionistisch wirken und durch den Fokus auf bestimmte Aspekte von anderen Faktoren ablenken. So sind Softwareprodukte, die wie Sprachkorrektoren eine De-Biasing-Funktion besitzen, kein Ersatz für eine Überprüfung aller anderen Aspekte, die zu Genderdiskriminierung führen.
Ebenfalls problematisch sind Konzepte, die Homophilie, also „das Prinzip, wonach Ähnlichkeit Verbindungen erzeugt“ (CHUN 2018), oder den Status Quo als optimale Ausgangsprämisse haben. Culture-Fit-Tests oder ASR-Produkte, die Bewerber*innen auf ihre Ähnlichkeit mit Top-Angestellten eines Unternehmens überprüfen, setzen nicht auf eine Veränderung der Belegschaft und der Arbeitskultur, sondern perpetuieren bereits bestehende strukturellen Asymmetrien.

3.1.8 Abbildung des Verfahrens

Wie in den Abschnitten 2.4. und 3. beschrieben, ergibt sich durch die Verknüpfung verschiedener Schritte zu einem Verfahren eine Komplexität von Wirkungsweisen, die ganzheitlich betrachtet werden muss, um die Ursachen möglicher Biases aufzudecken

und das Verfahren als solches bewerten zu können. Einstellungsverfahren bestehen aus verschiedenen Handlungsschritten und Tätigkeiten. Das gesamte Verfahren von der Stellenanzeige bis zur Vertragsunterzeichnung kann nicht vollständig automatisiert werden. Einzelne Schritte im Verfahren werden als digitale Dienstleistung angeboten, von Assistenz beim Schreiben von Anzeigen über das Screening bis hin zum Anonymisieren von Lebensläufen.
Ist das automatisierte Verfahren eine Abbildung des tatsächlichen Verfahrens? Oder optimiert das System basierend auf dem Output das Input, ohne das tatsächliche Verfahren zu berücksichtigen? Lassen sie sich im Gesamtverfahren einbetten? Oder kompartimentalisieren sie eher das Verfahren in einzelne Software-Module?
Werden bspw. Lebensläufe anonymisiert, ohne zu überprüfen, ob die vorausgehende Jobanzeige gendergerecht formuliert wurde und die Screening-Kriterien gendergerecht sind?

3.1.9 Einbettung der Entscheidung

Doch die Evaluation eines ASR muss über die reine Bewertung der Daten, die in die Software einfließen, der Ergebnisse und der Technologien hinausgehen. Bias kann entstehen, wenn die Ergebnisse der ASR (z. B. Korrelationen, Scores etc.) in die organisationseigene Policy für das Personalmanagement eingebettet werden. Wenn sie als normativer Maßstab genommen werden, können sie zur mathematischen Rechtfertigung und Verschleierung struktureller Diskriminierung werden. Wenn sie dagegen als Indikator für die Notwendigkeit von Inklusionsmaßnahmen benutzt werden, werden diese Systeme zur Assistenz, um über die eigenen strukturellen Genderasymmetrien zu reflektieren.
Wie werden die Ergebnisse eines ASR eingesetzt? Sollen sie mehr Diversity erzielen oder eine ökonomische und produktionstechnische Optimierung erreichen? Werden bspw. Informationen über den Wohnort von Bewerber*innen im ASR mitbewertet? Wenn ja, wie wird dieser Faktor bei Bewerber*innen bewertet, die in armen Vierteln oder ländlichen Regionen mit schlechtem öffentlichen Nahverkehr und ohne private Fortbewegungsmittel leben? Fließen die gewonnenen Erkenntnisse als negativer Faktor in die Bewertung ein, oder um Bewerber*innen einen Firmenwagen anzubieten?

3.2. Das Ökosystem

Die künstliche und isolierte Betrachtung von ASR-Produkten führt dazu, dass Probleme, die sich aus der Interaktion mit anderen Systemen sowie mit den involvierten Menschen

ergeben, nicht abgebildet werden. Nachfolgend werden die Ansatzpunkte aufgelistet, entlang derer ASR mit einem systemischen Blick hinsichtlich intersektionaler Gendergerechtigkeit betrachtet werden können.

3.2.1. Systeme

ASR werden in bereits bestehende, komplexe rechnergestützte Architekturen integriert: Sie werden in Hardware eingeführt, die bereits mit bestimmten Betriebssystemen arbeiten (wie etwa Microsoft, die über eigene Inklusivitäts-Sprachkorrekturen und Suchlogiken verfügen), in Personalmanagement-Software (mit eigenen Cluster-Kategorien und algorithmischen Logiken) und ggf. über Schnittstellen mit anderen Softwareprodukten verknüpft (wie etwa die Software für die Finanz- oder Rechtsabteilung). Diese Systeme verarbeiten und kategorisieren nach jeweils eigenen Logiken und Regeln. Aus diesem Zusammenwirken ergeben sich eine Reihe von Fragen, die für die Bewertung des Ökosystems relevant sind:

- Welche Logiken arbeiten gender-neutral, welche gender-inklusiv? Sind diese Logiken kompatibel? Wie werden Informationen von einem in das andere System übertragen? Sind dabei alle Informationen übertragbar?
- Welche anderen algorithmischen Systeme stehen in der Umgebung des ASR?
- In welchem System ist das ASR integriert (Bewerbermanagement-, Personalmanagementsystem)?
- Welche Schnittstellen gibt es?
- In welchen anderen Systemen werden die Outputs anderweitig verarbeitet? (Beispielsweise Finanzabteilung, Rechtsabteilung)

3.2.2. Metriken

Metriken sind eine Mathematisierung von Vorgaben, Regeln und Grenzdefinitionen.

- Welche Metriken sind Teil des Recruiting Systems?
- Welche Metriken sind Teil des Recruiting-Ökosystems?
- Welche Metriken regeln das Recruiting-Verfahren sowohl analog (z. B. psychologische, nicht digitalisierte Tests) als auch maschinell?
- Welche Metriken haben die Schnittstellen?
- Auf welchen Annahmen basieren diese Metriken?
- Inwiefern berücksichtigen diese Metriken intersektionale Gender-Asymmetrien und Unterschiede?

- Wie werden diese berücksichtigt?
- Welche Aspekte/Sachverhalte werden in Metriken übersetzt? Welche nicht?

3.2.3. Wer ist in das Verfahren involviert?

Wer kann alles ein System beeinflussen? Neben Software-Entwickler*innen, Algorithmenentwickler*innen und Datenwissenschaftler*innen gibt es eine Reihe von Akteur*innen, die mitentscheiden, welche Technologien zu welchen Zwecken und auf welche Art eingesetzt werden. Das fängt an bei den Akteur*innen, die für oder gegen den Kauf einer bestimmten Software oder einer Dienstleistung entscheiden, geht über die Mitarbeiter*innen, die die Produkte nutzen und sie ggf. einer anderen Funktionalität zuordnen, und endet bei den Kandidat*innen, die mit der Software interagieren und versuchen, ebenfalls in ihrem Sinne mit der Software umzugehen. So versuchten Bewerber*innen bei Lebenslauf-Screening-Software ein höheres Ranking zu erzielen, indem sie im Lebenslauf in weißer und damit für das Auge nicht erkennbarer Schrift bestimmte Schlüsselbegriffe einfügten. Die Methode funktioniert zwar nicht, kann aber weitere negative Konsequenzen haben (FORBES 2016). Aus diesem Grund ist es hinsichtlich intersektionaler Gendergerechtigkeit wichtig, nicht nur die Metriken und Technologien zu verstehen, sondern auch die Gruppe der Akteur*innen, die diese Technologien aus verschiedenen Positionen mitgestalten. Dabei ist es relevant zu betrachten, welche menschlichen Handlungen von welcher Akteursgruppe vom System erfasst werden und welche nicht.

- Welche Daten und welche Handlungen werden erfasst?
- Erfasst das System auch die situativen Machtasymmetrien (*studying up* vs *studying down*, BARABAS ET AL. 2020), d. h. werden neben der Bewerber*innen auch die Recruiter*innen, das Personalmanagement, ggf. alle anderen interagierenden Abteilungsmitarbeiter*innen (bspw. Finanz- oder Rechtsabteilung) erfasst?
- Gibt es Akteur*innen, die die Produkte mitgestalten und nicht von System erfasst werden?

3.2.4. Der Prozess

Üblicherweise wird unter dem Verfahren eines ASR verstanden, wie der Algorithmus oder die Algorithmen im System operieren, und ggf. die Art und Weise, wie sie trainiert wurden. Diese Sicht ist künstlich und reduzierend. Sie berücksichtigt nicht die weiteren

(manuellen und maschinellen) Verfahren bei der konkreten Implementierung. Ein Lebenslauf-Anonymisierungstool setzt beispielsweise eine Stellenanzeige mit bestimmten Kriterien voraus, die sich idealiter im Lebenslauf wiederfinden müssen. Dem Anonymisierungsprozess folgt zudem ein manuelles oder digitales Screening, das die Kriterien der Anzeige aufnehmen und als Filter für die manuelle oder digitale Prüfung der Bewerbungen dienen soll. Alle diese drei Sachverhalte sind rein inhaltlich voneinander abhängig und sollten nicht ausschließlich getrennt entworfen und implementiert werden. Technisch gesehen benötigen alle drei Sachverhalte unterschiedliche Techniken, wenn sie digital durchgeführt werden. Eine Analyse sollte sowohl die inhaltliche als auch technische Ebene durchdringen; sowohl im Einzelnen als auch architektonisch. Es ist sinnvoll, jede Technik, aber auch jeden Verfahrensschritt (ob digital oder analog) individuell und ganzheitlich zu betrachten.

- Wie sieht der gesamte Lebenszyklus einer Arbeitskraft von Recruiting bis zum Ausscheiden aus dem Unternehmen aus?
- Welche Prozesse gibt es im Einzelnen? Welche davon sind digital, welche manuell?
- Wie werden diese vermengt?
- Wie werden (maschinelle) Ergebnisse kontextualisiert und interpretiert?
- Welche Feedback- und Aufsichtsverfahren existieren?
- In welchen Schritten des gesamten Personalmanagement-Zyklus gibt es Controlling-Mechanismen? In welchen nicht?

3.2.5. Die Regeln

Neben den Metriken geben soziale und gesetzliche Regeln den Rahmen vor, in dem Technologie eingesetzt wird (vgl. Abschnitt 3).

Soziale Regeln

Die tatsächlichen menschlichen Gewohnheiten, Erwartungshaltungen und Einstellungen zur Technologie sind Teil der sozio-technischen Dimension von algorithmischen Systemen. Diese zu identifizieren und zu verstehen, ist essentiell, um Effekte und Verzerrungen bei der Implementierung und Nutzung einer Software vollständig evaluieren zu können. Softwareprodukte werden nicht unmittelbar so benutzt, wie sie von den Entwickler*innen konzipiert wurden. Manche Funktionen sind zu kompliziert und werden mit einer anderen Software oder einem anderen Produkt erledigt. Manche Ergebnisse des Systems werden antizipiert, so dass durch Modifikation des Verhaltens (z. B. durch Nut-

zung von bestimmten Schlüsselwörtern) eine Veränderung der Berechnungen angestrebt oder erzielt wird. Dies kann wiederum neue Arten von Verzerrungen im System bringen.
Auch die Funktion des Softwareproduktes ist in dieser Hinsicht zu betrachten: soll das Produkt einem bereits bestehenden Prozess im Personalmanagement assistieren, oder geht es darum, mit dem Produkt eine bestimmte menschliche Tätigkeit im Prozess maschinell zu ersetzen? Werden die bestehenden Verfahren berücksichtigt oder diese komplett umgestellt?

Gesetzliche Regeln

Auch die gesetzlichen Regeln beeinflussen die Architektur von algorithmischen Systemen. Neben arbeitsrechtlichen Vorgaben kommen andere sektorale Gesetze je nach Kontext in Betracht (z. B. Medizin, Agrarrecht), sowie datenschutzrechtliche oder allgemein gleichstellungsrechtliche Vorgaben. Diese verschiedenen Gesetze können kollidieren bzw. in der kontextuellen Vermengung Inkompatibilitäten aufweisen. Sie können durchaus technische Spezifikationen über die Datenverarbeitung oder mathematische Vorgaben enthalten, die die Governance und Architektur der Systeme beeinflussen. Auch hier ist die kontextuelle Betrachtung aus allen rechtlichen Disziplinen in Verflechtung mit der einzusetzenden Technologie notwendig, da diese Gesetze eine normative Gestaltungskraft über die konkrete Technologieumsetzung besitzen.
Wie den verschiedenen angerissenen Gesichtspunkten zu entnehmen ist, bedarf eine Auswertung von ASR eines systemischen, interdisziplinären Blicks und zeigt die Komplexität und Eingriffstiefe einer solchen Technologie in die Prozesse einer Organisation.

4. Algorithmen- und Datenmärkte

Um Genderdiskriminierung in algorithmischen Systemen im Allgemeinen und ASR im Besonderen zu verstehen, ist nicht nur eine Analyse der Softwareprodukte notwendig, sondern auch der normativen Vorgaben einer Gesellschaft, in der solche Systeme implementiert werden. Wissenschaftliche Theorien, Ethik und Recht bilden die Leitplanken beim Einsatz von Technologien, sie kreieren den Rahmen, die Anreize und die Regeln, in denen die Technologien in Märkten angeboten und entwickelt werden. All diese Aspekte sollten bei der Untersuchung von Algorithmen- und datenbasierten Geschäftsmodellen und Märkten berücksichtigt werden.

Die folgende Analyse widmet sich ausgewählten Aspekten der Algorithmen- und datenbasierten Entwicklung von ASR. Sie hat keinen Anspruch auf Vollständigkeit. Die Studie ist vielmehr als Plädoyer für eine Vertiefung der Forschung in den unten skizzierten Aspekten und Kontexten zu verstehen.

4.1. Algorithmenmärkte

Eine Befragung von 400 Führungskräften in Deutschland, die von der *Initiative Chefsache* im Jahr 2018 durchgeführt wurde, ergab:

> „Nur 29 % der Befragten berichten von standardisierten Kriterien zur Identifikation von Talenten. Auch digital unterstütztes Feedback, z. B. Feedback-Apps, nutzen nur 15 % der Befragten." (CHEFSACHE REPORT 2018, 11)

Diese Zahlen galten für digitale Recruitingprodukte im Allgemeinen bis vor der Corona-Pandemie. Dies belegen zwei verschiedene Studien des *Bundesverbands der Personalmanager* und *Stepstone*, die in zwei Teilen jeweils in Januar und Mai 2020 geführt wurde, sowie eine Studie von *Stellenanzeigen.de* in Kooperation mit der Unternehmensberatung *Meta HR. Ca*., wonach ein Drittel der Unternehmen – darunter auch kleine und kleinere mittelständische Unternehmen – digitale Recruiting-Instrumente verwenden. Management-Werkzeuge werden in der Studie nicht als digitale Instrumente gezählt. Sie sind allerdings bei der Hälfte der Befragten in Benutzung (KOSCHICK 2020). Bewerbermanagement-Tools können cloudbasierte, digitale Werkzeuge sein, die zum Teil mit komplexen algorithmischen Systemen operieren. Darunter befinden sich Produkte wie *SAP Success Factors*, das für Recruiting-Tätigkeiten auf Verfahren aus dem Maschinellen Lernen und dem NLP sowie dem *Natural Language Understanding* (NLU) rekurriert. Diese Produkte bieten Leistungen wie

- Sprachscans zur Vermeidung von Vorurteilen in Stellenausschreibungen,
- Einsatz von Maschinellem Lernen für Gehaltsempfehlungen,
- Aufstellungen von Angebotsregeln für eine faire und gerechte Bezahlung unabhängig von Geschlecht, Rasse, sexueller Orientierung und anderen Aspekten.[27]

Insgesamt sehen Personalmanager*innen laut unterschiedlichen Studien seit 2018 ihr Arbeitsfeld regelmäßig als digitalisierbar an und rechnen mittelfristig mit einem Einzug digitaler Produkte und einer Automatisierung von Prozessen (CHEFSACHE REPORT

27 https://www.sap.com/germany/products/recruiting-software.html

2018, JÄGER & MEURER 2018, GADE ET AL. 2018, RECRUITING 2019/2020, STEPSTONE 2020, ATHANAS & BARISLOVITS 2020). Doch die bereits vorhandene Verfügbarkeit von digitalen Instrumenten lässt auf eine viel breitere Digitalisierung im Bereich Recruiting schließen (PESCH 2018, 25-27). Mögliche Erklärungen für diesen Befund versucht die Studie von GADE ET AL. (2018) zu liefern. Großunternehmen verwenden ihr zufolge mehr digitale Werkzeuge als KMU, da sie offenbar über eine hinreichende Menge an Daten und Profilen, insbesondere auf Fachkräfte-Ebene, verfügt. Primär werden digitale Werkzeuge dabei im Sourcing benutzt. Mittleres Management und vor allem Top-Management-Positionen sind ebenfalls weniger durch digitale Recruiting-Strategien betroffen. Der Studie nach entsprechen die derzeitigen Werkzeuge womöglich nicht den Erwartungen der Unternehmen.

Erwartungen an digitale Lösungen unterscheiden sich.

Häufigste Nennungen

Unternehmen*	Personalberater
Datensicherheit 88 %	Beschleunigung der Prozesse 48 %
Schnittstelle zu eigenen Systemen 75 %	Vereinfachung der Abläufe 29 %
Validität 64 %	Entlastung 27 %
Mobile Verfügbarkeit 59 %	Beschleunigung 19 %

*Kein Unterschied zwischen großen und KMUs

Abb. 1: Erwartungen an digitale Lösungen (GADE ET AL. 2018)

Das Undokumentierte

Die in der vorliegenden Expertise untersuchten Produkte dokumentieren nicht die Anzahl der abgelehnten Bewerber*innen und deren Geschlecht oder weitere Merkmale. Darüber hinaus sind in Deutschland laut der Studie von JÄGER & MEURER (2018) Controllingverfahren im Bereich Recruiting noch nicht Standard: Über 43 % der Befragten gaben an, kein Recruiting-Controlling zu betreiben.
Die Verwendung einiger der aktuell verfügbaren ASR können durchaus als eine konsistente Form der Dokumentierung von Diskriminierungsmustern im Verlauf eines Recruiting-Prozesses gesehen werden. Die Werkzeuge zeigen dabei eine gewisse Ambivalenz.

In Deutschland werden seit Ende 2019 vermehrt von großen Software-Anbietern wie SAP, IBM und Microsoft digitale Produkte angeboten, die als Assistenz gegen Klischees, Vorurteile und Diskriminierung helfen sollen. Doch zugleich nimmt nicht nur die Debatte der impliziten Diskriminierung in algorithmischen Systemen an Gewicht zu (vgl. ORWAT 2019). Es entstehen zusätzliche (auch rechtliche) Risiken hinsichtlich mittelbarer Diskriminierung, die sich in den auf dem Markt angebotenen Produkte widerspiegeln. Durch die Zusammenführung verschiedener Softwareprodukte und somit verschiedene algorithmische Diversity-Ansätze auf Plattformen (z. B. SAP HANA) können unvorhergesehene Formen von algorithmischen Kollisionen, Inkonsistenzen und in der Folge zusätzlichen Diskriminierungen entstehen. So gibt es beispielsweise zwischen der Bewerbermanagement-Software von SAP, *Success Factors*, welche Sprachscans auf der Suche nach Vorurteilen im Recruiting ausführt, und *Cammio*, einer Software, die automatisierte Video-Interviews ermöglicht, eine Schnittstelle, deren Funktionsweise in Bezug auf Diskriminierungen unklar ist. *Cammio* bietet in diesem Zusammenhang *Xpress Analytics* an, ein Feature, das unter Nutzung von *IBM Watson* den gesprochenen Text der Kandidat*innen analysiert und basierend auf dem OCEAN-Modell die Persönlichkeitsmerkmale in ein Persönlichkeitsprofil einordnet. Wir haben dementsprechend in diesem Beispiel mindestens: 1) die Logik der Sprachscans von SAP, 2) die Logik von *Cammio* und 3) die Logik von *IBM Watson* nebst 4) der Logik hinter des psychologischen OCEAN-Modells sowie 5) die Logik der Plattform, auf der alle Systeme betrieben werden, gegebenenfalls – je nach Architekturmodell – sogar unter 6) Hinzuziehung einer analytischen Zwischenschicht. Damit erhöht sich das Risiko einer Kollision der Logiken: So basiert die Video-Analyse von *Cammio* etwa auf einer Kategorisierung, die Verhaltensweisen eng mit Gender assoziiert, während die genderspezifische Sprachlogik von SAP nach Möglichkeit eine genderneutrale Sprache empfiehlt (vgl. SAP 2019). Dies erschwert nochmals die Evaluation und Zuordnung von Verantwortung und Haftung, die schon bei einer individuellen Betrachtung der jeweiligen Software schwierig ist. Dies gilt nun umso mehr, als wir es mit verschiedenen Anbietern zu tun haben, deren Produkte zusammengeführt werden, ohne dass die jeweiligen Wechselwirkungen und Effekte der Zusammenführung bekannt wären. Wer ist demnach für eine potentielle intersektionale Diskriminierung hierzu haftbar: das Unternehmen, dass eine solche Konstellation nutzt, SAP, *Cammio* oder IBM? Auch der Nachweis einer mittelbaren Diskriminierung wird bei

der Kombination von drei proprietären Software-Codes erheblich erschwert. Das wiederum senkt das Risiko einer Entschädigung, da Vorsatz oder Verschulden nachgewiesen werden muss.[28]

Neben dem zusätzlichen Risiko der Sicherheit und Integrität der genutzten sowie der generierten Daten bestehen auch datenschutzbedingte Probleme sowie Notwendigkeiten von Anpassungen. Damit gehen Kosten bei der Einführung einer neuen technischen Infrastruktur einher: anwaltliche Kosten, Personalkosten, technische Vorkehrungen etc. Vor dem Hintergrund der datenschutzrechtlichen Herausforderungen besteht ein Anreiz, den Anwendungsbereich des Datenschutzrechts zu verlassen und mit anonymisierten Techniken im Verfahren des Personalmanagements zu arbeiten. So startete die Diskriminierungsstelle des Bundes 2010 ein Pilotprojekt zur Anonymisierung von Bewerbungsverfahren mit der Entwicklung eines standardisierten Verfahrens mitsamt online-Version (vgl. BÖSCHEN ET AL. 2012). Eine Verarbeitung von anonymisierten, d. h. nicht personenbezogenen oder personenbeziehbaren Daten fällt nicht in den Rahmen des Datenschutzrechts (siehe Erwägungsgrund 26, letzter Satz, zu Art. 4 Nr. 5 DSGVO). Dies vermeidet rechtliche Risiken, schafft aber neue Probleme. Zwar werden Anonymisierungsverfahren auch als eine Maßnahme gegen Diskriminierung betrachtet. Doch Anonymisierung ist ein dehnbarer Begriff. Aus einem Schreiben können z. B. Alter, soziale Herkunft und Gender etc. durch den Gebrauch bestimmter Redewendungen, Wörter und Ausdrücke, die von bestimmten Generationen, sozialen Gruppen usw. verwendet werden, abgeleitet werden. Anhand eines Lebenslaufs können ebenfalls bestimmte Annahmen über die Herkunft und Profil eine*r Bewerber*in vorgenommen werden, wie etwa ob die Person in einer religiösen Einrichtung ausgebildet wurde. Durch Angabe von Hobbys und ehrenamtlichen Tätigkeiten können der Bewerbung auch soziale Herkunft sowie Sozial- und Bildungskapital entnommen werden. So finden wir auf dem deutschen Markt verschiedene sektorale anonyme Recruiting-Apps wie *truffls*[29] für Berufseinsteiger*innen, *omnium*[30] für Fachkräfte, *legalhead*[31] für Jurist*innen oder *placing you*[32] für Hörakustiker*innen und Augenoptiker*innen. Diese Apps sind primär für

28 Siehe §15 und §21 AGG.

29 https://truffls.de/de/

30 https://www.omnium.de/

31 https://legalhead.de/

32 https://placing-you.de/

mobile Bewerbungen gedacht. Einige dieser Apps sehen vom Upload eines konventionellen Lebenslaufs ab. Die Bewerber*innen bleiben anonym, bis sie dem Match mit einem bestimmten Unternehmen proaktiv zustimmen. Doch diese Anonymität ist keineswegs vollständig und betrifft gerade nicht die gruppenbezogenen Diskriminierungsmerkmale. Bei der Angabe des Bildungsabschlusses bzw. des Lebenslaufes können mehrere Rückschlüsse über den Hintergrund der Person gezogen werden, insbesondere wenn der Ausbildungs- oder Studienort angegeben wird. Es können z. B. das Geschlecht oder politische und religiöse Positionierungen abgeleitet werden, wenn es sich um einen Abschluss einer Mädchen- oder Jungenschule, einer religiösen Universität oder der Bundeswehr handelt. Eine vollständige Anonymisierung, die keine diskriminierenden Rückschlüsse zulässt, wird von keinem Anonymisierungswerkzeug erreicht.
Dem wirken auch Mechanismen einer künstlichen Verschleierung nur bedingt entgegen. Zum einen können künstliche Verschleierungen nicht während des gesamten Kommunikationszyklus aufrechterhalten werden. Teilweise ist es sogar das Ziel der Personalmanager*innen oder der Unternehmensleitung, Verschleierungen am Ende entgegenzuwirken, etwa, wenn es ihnen darum geht, Personen mit ähnlichen Profilen und Eigenschaften, die den ihrigen ähneln, anzustellen. So hebt der CHEFSACHE REPORT (2018, 9) hervor:

> „Von den Vorteilen von Vielfalt ist nur die Minderheit der Befragten überzeugt: 67 % gaben an, es sei ihnen bei Beförderungen oder Teamzusammenstellungen wichtig, dass die Personen ähnlich arbeiten und denken wie sie selbst."

Darüber hinaus helfen anonymisierte Lebensläufe bei der ersten Sichtungshürde nicht, wenn die Anstellungsmaßstäbe auf Kriterien setzen, die in der Bevölkerung nicht homogen verteilt sind. Wenn beispielsweise Führungserfahrung als einziges Kriterium für die Feststellung von Führungskompetenz abverlangt wird, werden damit nur die Profile bevorzugt, die historisch die Führungsetagen geprägt haben. Das wird mitunter Frauen, insbesondere Frauen mit weiteren potenziellen Diskriminierungsmerkmalen, ausschließen. Allein diese erste Einstiegshürde bei der Bewerber*innenauswahl sollte wesentlich prozeduraler verstanden werden. Sie umfasst mehr als die Sichtung von Bewerbungen und fängt bereits bei der Erstellung einer Anzeige an. Wenn die Sprache der Anzeige, einen männlich konnotierten Wortschatz verwendet, werden sich weniger Frauen angesprochen fühlen. So zeigte jüngst eine Studie, dass

> „strukturelle Ungleichheiten in der Sprache von Stellenanzeigen bestehen, die vorhandene Geschlechterunterschiede aufrechterhalten und verfestigen. Vor dem Hintergrund des oben beschriebenen Mechanismus der Zugehörigkeit, wonach sich Männer und Frauen vorrangig für Stellenausschreibungen interessieren, die

> in ihrer Geschlechtersprache formuliert sind (weil sie eben dadurch Passung und Zugehörigkeit erleben) sind gerade Befunde, die wir sie im Rahmen der Untersuchung von Berufen mit unterschiedlicher Bezahlung machten, von Brisanz.“ (BUREL ET AL. 2019, S. 8).

Weiterhin sind die Ausschlusskriterien für die nächste Runde zu beachten: wenn als Qualifikation für ein Gesprächsinterview nur Kriterien gelten, die wenig auf Potential, sondern auf Performance setzen, werden strukturell benachteiligte Gruppen weiterhin Schwierigkeiten haben, ihr Potential zu entfalten. So stellte eine Studie von MCKINSEY (2011, 6) fest, dass Männer aufgrund ihres Potenzials angestellt und gefördert wurden, Frauen dagegen basierend auf Performance und Erfolgsbilanz. Eine Anonymisierung der Bewerbungen als Maßnahme, um unbewussten Bias beim Sourcing abzumildern, verkennt die Komplexität und Verfahrenshaftigkeit dieses Arbeitsschrittes. Sie kann vielmehr zu nicht intendierten, gegenteiligen Ergebnissen führen (BEHAGHEL ET AL. 2015).

Aus ethischer Sicht kann das Verschleiern der Merkmale daher durchaus problematisiert werden. Die Verschleierung setzt auf Teilhabe durch Verstecken von Persönlichkeitsmerkmalen. Damit werden diese Persönlichkeitsmerkmale nicht normalisiert, sondern fortwährend problematisiert. Doch die Teilhabe einer schwangeren Person am Arbeitsleben wird nicht dadurch befähigt, dass ihre Schwangerschaft versteckt wird, sondern dadurch, dass ihr entsprechend gleiche Maßstäbe und gesonderte Rechte (z. B. Kündigungsschutz) gewährt werden, damit sie ihre Schwangerschaft zeigen kann, ohne Angst vor Diskriminierung zu haben.

4.2. Datenmärkte: Über die algorithmische Materialisierung und Ökonomisierung von Affekten und Emotionen

In der oben erwähnten Studie (STEPSTONE 2020) gehören Cultural-Fit-Tests bei 15 % der Unternehmen zum Auswahlprozess. Dazu kommen bei ca. 10 % der Unternehmen noch Gesichts- und Sprachanalyse-Tools. In einigen weiteren der angegebenen Werkzeuge (wie etwa *predictive analytics*) sind möglicherweise auch Komponenten enthalten, die eine Analyse der Fähigkeiten oder der emotionalen Intelligenz sowie der sozialen Eigenschaften der Kandidat*innen durchführen. Damit wird die Passform der Kandidat*innen zur Unternehmenskultur hinsichtlich ihrer Wertvorstellungen, ihrer Motivation und ihres Engagements geprüft. Die jüngsten Softwareprodukte im Bereich Recruiting tendieren dabei stark in Richtung der Entwicklung von Metriken und setzen psychologische Methoden ein, um menschliche Eigenschaften jenseits der inhaltlichen

Fachkompetenz zu analysieren. Die potentiellen intersektionalen Diskriminierungsrisiken bei der Automatisierung psychologischer Testmethoden liegen bereits in der Konzeptualisierung der psychologischen Tests. Diese werden dann bei der spezifischen Implementierung in Software-Lösungen verstärkt. Eine Analyse der inhärenten Gender Bias psychologischer Modelle ist vonnöten, um die technikspezifischen Diskriminierungsrisiken bei der Umsetzung zu verstehen.

Die Ökonomisierung von Emotionen

Die Arbeitswelt war bis zu den 1960er Jahren durch eine strikte Trennung zwischen den wirtschaftlichen rationalen und menschlichen Leidenschaften charakterisiert. Gefühle waren dem privaten Bereich zugeordnet und das rationale Handeln dem Beruflichen. Wie die Soziologin Eva ILLOUZ (2009a) anmerkt:

> „Gefühle und Leidenschaften, Lebensgefühl und Lebensweise waren bestimmt für den privaten Lebensbereich und wurden dort zur emotionalen Basis für eine neue Kultur und Lebensweise des Konsums."

Die Definition und Unterscheidung zwischen Gefühlen, Emotionen und Affekten ist in der Wissenschaft umstritten. In diesem Kontext wird aus funktionalen Gründen die Unterscheidung von SHOUSE (2005, 1) verwendet:

> „Feelings are personal and biographical, emotions are social, and affects are prepersonal."

Die Kommodifizierung von Emotionen und Affekten fand Ende der 1960er und während der 1970er Jahre statt. Sie wurden als neue „Gebrauchs- und Funktionswerte (…) zu Potentialen der Ware Arbeitskraft" (ILLOUZ 2009a) bezeichnet, die kontrolliert und zu produktiven Zwecken eingesetzt werden sollten. Emotionen, Affekte und Rationalität standen in einer produktiven, instrumentellen Wechselwirkung. Durch die Integration der Psychologie in den Arbeitsmarkt (ausgehend von der Markensoziologie), entstehen neue Normen von professionellem Verhalten von Fachkräften und Management. Zwischen 1927-1932 führten die Professoren Mayo und Roethlisberger die Hawthorne Experimente aus. Daraus skizzierte Elton MAYO (1949) den Grundsatz einer Betriebssoziologie, die auf eine starke Fokussierung auf Zugehörigkeitsgefühle und Kooperation der Fachkräfte mit dem Betrieb und auf Streikprävention über Vermeidung von konflikttreibenden Emotionen setzte. Mayo baute seine Verfahren auf die Theorien der Gesprächstherapie Piagets auf. Demnach wurden Arbeitnehmer*innen zur Kooperation als

Zeichen einer angemessenen sozialen Integration verpflichtet. Diejenigen, die als "unkooperativ" eingestuft wurden, konnten von den Studien ausgeschlossen werden und letztlich auch vom Arbeitsplatz selbst – so wie es mit zwei Frauen geschah, die im Relais-Raum der Hawthorne-Anlage arbeiteten. Mayos Theorien erlaubten es Managern, sich in autoritativer Weise unter dem Deckmantel eines paternalistischen Interesses zu verhalten, so Andrzej HUCZYNSKI (1993, 44) in seinem Buch *Management Gurus.* Der Manager als Vaterfigur verkörperte ein fürsorgliches Wohlwollen, das unterwürfige Arbeiter*innen belohnte. Die Vaterfigur hatte anderen Emotionsnormen als ein*e Arbeiter*in zu entsprechen. Als Gesprächstherapeut seinen Arbeitskräften gegenüber, richtete er über deren emotionalen Haushalt und Zugehörigkeitsgefühle.
Kompetenz wurde so durch die bewusste, distanzierte Kontrolle von Emotionen erlangt. Emotionen wurden klassifiziert und als Motiv für wünschenswerte oder negative Verhaltensweisen im Beruf eingeordnet. Die Körpersprache, die Mimik, der sprachliche Ausdruck von Emotionen wurden systematisiert und standardisiert als weiteren Bewertungsmaßstab für Führungsqualität oder Teamfähigkeit gedeutet. Weinen, Gefühlsausbrüche – nicht fachliche Inkompetenz – wurden zu unprofessioneller Haltung deklariert.
Dieser Entwicklung liegt die Annahme zugrunde, dass psychologische Zustände mehr sind als individuelles Empfinden. Sie werden zu sozialen und kulturellen Praktiken in diesem Kontext in der Arbeitswelt. Diese Praktiken und Normierungen der Emotionen sind Gegenstand feministischer Philosophie. Wie Sara AHMED (2004, 3) anmerkt, hat sich die feministische Philosophie seit langem mit der Kontrolle und Subordination der Emotionen auseinandergesetzt:

> „Feminist philosophers have shown us how the subordination of emotions also works to subordinate the feminine and the body."

Emotionen werden im feministischen Kontext von Ahmed als soziale Praktiken verstanden und sind somit von Natur aus relational. Dabei hat dieser relationale Charakter einen emanzipatorischen Grund. Die Relationalität scheint aus dem Bestreben heraus zu wachsen, die Arbeit und Reflexionen um Emotionen von der Aufklärung, und dessen Ideen über die Natur des Individuums, zu distanzieren.
Mit dem Einzug der Psychologie in die Arbeitswelt und die Vermarktung von Emotionen als Arbeitskapital entstanden indessen grundlegende Friktionen: Relationale, statistische Methoden wurden ausgerechnet dafür eingesetzt, universelle Aussagen über die emotionale Natur des Menschen zu treffen und standardisierte berufliche Profile über emotionale Kompetenz zu erstellen. Emotionen wurden durch diesen Passus nicht nur im aufklärerischen Sinne rationalisiert, sondern essentialisiert – aus feministisch-philosophischer Sicht ein der Natur der Emotionen nach widersprüchliches Bestreben.

Verdatung der Emotionen

In der US- und der kontinentaleuropäischen Psychologie suchten verschiedene Psycholog*innen nach universellen, allgemeingültigen Persönlichkeitsaspekten, nach denen Emotionen eingeordnet werden können. Verschiedene Forschergruppen seit Ende der 1930er Jahre und insbesondere während der 1960er und 1980er Jahre (z. B. ALLPORT 1937, NORMAN 1963 und COSTA 1985) präsentierten ähnliche Klassifizierungen, die aus fünf Persönlichkeitsdimensionen bestanden. GOLDBERG (1981) nannte diese Klassifikation die „Big Five“. Auf Englisch sind diese Modelle als *OCEAN Model* bekannt. Der Name ergibt sich aus den Anfangsbuchstaben der fünf Persönlichkeitsdimensionen (Openness, Conscientiousness, Extraversion, Agreeableness und Neuroticism). Bestimmte Verhaltensweisen wurden Emotionen zugeordnet und diese wiederum klassifiziert und mit einem bestimmten Datum versehen. Dieser Prozess lässt sich als Verdatung von Emotionen beschreiben.

In den 2010er Jahren, mit der Fortentwicklung der Hardware, konnten psychologische Modelle, die große Datenmengen benötigten, für den Arbeitssektor als Dienstleistung skaliert und profitabel digitalisiert werden. Mit dem Versprechen, Prozesse im Personalmanagement effizienter zu gestalten und Assistenz anzubieten, insbesondere bei höherer Anstellungsfrequenz, erlangte die Psychometrie – sprich psychologische Bemessungstheorien –, insbesondere in Kombination mit dem „Big five“-Modell, Bekanntheit:

> „In den letzten Jahren haben die Big-Five-Faktoren auch das Interesse vieler Unternehmen geweckt, da führende Wirtschaftspsychologen wiederholt zeigen konnten, dass die Personalauswahl mit einem einfachen, 20-minütigen Big-Five-Persönlichkeitstest in Hinblick auf die prognostische Validität häufig einem Assessment Center ebenbürtig ist.“ (SATOW 2012, 5)

Was kann verdatet werden?

In Deutschland werden unterschiedliche Online-Varianten und Softwarprodukte basierend auf dem OCEAN Modell angeboten. Mal werden sie als Culture-Fit-Tests gekennzeichnet, mal sind sie als Persönlichkeitstests wie etwa *B5T* oder *JobPersonality* zu finden, mal fallen sie unter die Rubrik *social Skills*. Gegebenenfalls können sie in allen Versionen auch mit *behavioral biometrics* angereichert werden – sprich der Assoziation von Emotionen und/oder Persönlichkeitsdimensionen zu bestimmten körperlichen Muster

(Bewegungen, Mimik, Physiognomie)[33]. Ziel der Nutzung solcher Tests ist eine Evaluation der nicht fachlichen Eigenschaften der sich bewerbenden Person. Diese Evaluation zielt auf eine Überprüfung der persönlichen Kompatibilität der bewerbenden Person im Unternehmen.

Biometrie und Verhaltensanalyse

Über die grundsätzliche Gender- und Racial-Bias von biometrischer Software im Allgemeinen findet seit 2017 eine rege Debatte statt. So bewies Joy BUOLAMWINI 2017) in ihrem Paper „*Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers*“, dass biometrische Systeme aufgrund einseitiger Datenbanken und Hardware, die dunkle Hautfarbe schlecht reflektieren, eine hohe Fehlerquote bei weiblichen, nicht als weiß gelesenen Menschen besaß. CRAWFORD & PAGLEN (2019) arbeiteten die Ursprünge der theoretischen Konzeptionen hinter Verhaltensanalysen mit biometrischen Systemen heraus und zeichneten eine direkte Linie zur Phrenologie, Kraniologie und Eugenik. Allein die Klassifikation und Kategorisierung von menschlichen Körpern und die Assoziation von körperlichen Merkmalen oder Bewegungen zu eindeutigen Emotionen oder Charaktereigenschaften konstituiere eine Übertragung szientistischer, nicht wissenschaftlicher, phrenologischer Annahmen. Emotionen sind komplex und benötigen Kontextualisierung. Ohne Kontext kann ein Ausdruck von Freude schwer zu unterscheiden sein von einem Ausdruck von Schmerz.



Abb. 2: Tennisspielerin Osaka gewinnt ein Tennis-Match (nach RENFRO 2019, Kyodo News/Getty Images)



Abb. 3: Tennisspielerin Azarenka verletzt sich bei einem Training (nach ASIF 2015).

33 Siehe beispielsweise Produkte wie *HireVue*, *Cammio*, *humantic* oder *Precire*.

Die professionelle Körpersprache einer als Frau gelesenen Person wird anderen Normen unterzogen als bei einem als männlich gelesenen Menschen. Die Normen der beruflichen Körpersprache sind zudem kulturell unterschiedlich. Die derzeitigen Technologien auf dem Gebieten sind nicht in der Lage diese Ambivalenzen und Kontextabhängigkeiten mathematisch zu erfassen. Ganz im Gegenteil, sie berufen sich auf Theorien, die von der Wissenschaft als Szientismus nicht nur aufgrund wissenschaftlicher Mangelhaftigkeit, sondern auch aus ethischen Gründen abgelehnt werden.
In diesem Kontext stellt sich die Frage, ob Persönlichkeitsdimensionen algorithmisch gemessen werden können. Das OCEAN-Modell ist die wissenschaftliche Methode, die am häufigsten als Basismodell in der untersuchten Recruiting-Software angegeben wird (z. B. bei *Cammio*, *Precire*, *humantic*, oder *ideal*). Das OCEAN-Modell basiert in erster Linie auf Selbsteinschätzung (und hilfsweise auf Fremdeinschätzung). Es ist ein Modell, das grundsätzlich immer eine implizit komparative Frage stellt: „Sehen Sie sich als hilfsbereiter Mensch?“ Dabei stellte Leon Festinger in den 1950er Jahren fest (HOFFMAN ET AL. 1954), dass Menschen dazu neigen, ihre eigene Leistung im Vergleich zu anderen zu bewerten, die sie für vergleichbar halten. BIERNAT & FUEGEN (2001) wiederum wiesen nach, dass Geschlechterstereotypen die Art und Weise prägen, wie Menschen einander beurteilen. Wenn Frauen für ihre Selbstbeurteilung sich mit anderen Menschen vergleichen, die sie für vergleichbar halten (bspw. andere Frauen), wird ihre Selbsteinschätzung anders ausfallen, als wenn sie sich mit Menschen in ihrem allgemeinen Umfeld vergleichen oder mit Menschen im Unternehmen (wenn es um internes Recruiting geht). Diese relationalen Überlegungen spielen nicht nur eine inhaltliche Rolle bei der Bewertung der Frau. Auch bei der Implementierung der Software muss diese relationale Frage genauer betrachtet werden. Werden die Antworten der Frauen mit allen anderen Bewerber*innen verglichen oder nur mit Frauen? Sollte Letzteres der Fall sein, werden die – wohlgemerkt: relationalen – Rankings bei Frauen und Männern sich unterscheiden. Diese Erfahrung machte die Wissenschaftsjournalistin Olivia GOLDHILL (2018) von Online-Magazin *Quartz*, als sie verschiedene Online-Persönlichkeitstests machte und feststellte, dass die Auswertung bei gleicher Beantwortung je nach Geschlechtsangabe unterschiedlich ausfiel.
Aus methodologischer Sicht ist die inhärente konzeptuelle Spannung des OCEAN-Modells im Allgemeinen zu bedenken. Die Methodik ist in der Anwendung relational, doch die verwendeten Kategorien (die Persönlichkeitsdimensionen) wurden bereits als Ergebnis relationaler Untersuchungen entwickelt. Der Anspruch der Methodik und Metrik ist es jedoch in erster Linie über Selbsteinschätzungen allgemeingültige Aussagen über

den Menschen zu treffen. Doch ein Selbst-/Fremdreport kann nicht mit tatsächlichem Verhalten gleichgesetzt werden: Angst zu fühlen ist nicht dasselbe wie sich zu ängstlich zu verhalten.

Das Berichten über sich selbst spiegelt vielmehr die Narrative von sozialen Positionen in einer Gesellschaft. Es zeigt nicht nur eine Selbstwahrnehmung, sondern enthält implizit auch eine gesellschaftliche Wahrnehmung. Diese Wahrnehmung ist keineswegs deskriptiv, sie ist auch normativ und kommuniziert eine Erwartungshaltung.

So überrascht es kaum, dass die verschiedenen Persönlichkeitsdimensionen in vielen Studien mit dem Geschlecht korreliert werden, um Unterschiede in der Bezahlung oder in der Arbeitswelt zwischen Mann und Frau zu psychologisieren. So werden beispielswiese höhere Ausprägungen von Agreeableness (Konfliktscheue) und Neuroticism (Selbstzweifeln) bei Frauen festgestellt – diese Züge stehen für eine höhere Teamfähigkeit oder für ein Interesse an Care-Arbeit. Männer korrelieren wiederum mit höherer Ausprägung an Narzissmus und niedrigerem Neuroticism und Agreeableness. Diese Ausprägungskonstellation am Arbeitsplatz wird als niedrigere Teamfähigkeit und höhere Führungsfähigkeit gedeutet (GIOLLA & KAJONIUS 2019; LIPPA 1995; KARWOWSKI ET AL. 2013).

Aus dieser Vorkategorisierung werden durchaus schwerwiegende Konsequenzen auch auf der abstrakt-generellen Ebene gezogen, der sich das Individuum kaum entziehen kann. So wird der ungleiche Lohn von Frauen und Männern regelmäßig mit der Angst vor Konflikt und einem tendenziell höheren Grad an Neurosis der Frau begründet. Anstatt die Praxis der Gehaltsverhandlungen zu hinterfragen, wird auf eine Verhaltenshomogenisierung gesetzt. Ein höherer Grad an Narzissmus bei Frauen wird als psychologisch vorteilhafte Anpassung gesehen.

“Unfortunately women are not engaging in these behaviours as often as men, but when they do, they appear to benefit in a similar way”, bemerkt Professor Peter O’Connor in einem Interview bei der Vorstellung der Studie über Persönlichkeitsmerkmale, Gender und Ungleicher Lohn (WIDDOWSO 2018).

So werden Unterschiede im Gehalt oder Arbeitsplätze auf die Persönlichkeit zurückgeführt (JONASON ET AL. 2018), die nun strukturell oder gesellschaftlich bestimmte narrative Zuordnung über die Geschlechter vornimmt und in Algorithmen abgebildet und perpetuiert werden können. Die Zuordnungen werden jedoch nicht nur von den Algorithmen, sondern mehr oder weniger auch von den Geschlechtern selbst internalisiert. Ob dies tatsächlich zutrifft, ist umstritten. Die tatsächlichen Korrelationen zwischen den Persönlichkeitsdimensionen und den Geschlechtern haben sich in Studien als statistisch klein bis moderat erwiesen (WEISBERG ET AL. 2011). Entsprechende Folgen der aktuellen

Recruiting-Werkzeuge sind also weniger Grund als vielmehr Ausdruck der Ungleichbehandlung per default. Im Recruiting-Kontext haben die Tests die Funktion zu identifizieren, 1) welche Individuen ein Potential zur Führung haben und wer eher teamfähig ist, 2) welche Bewerber*innen sich dem Betrieb zugehörig fühlen werden, und 3) die Arbeitsbedingungen der Bewerber*in zu verhandeln. Dies erfolgt jedoch auf der Basis einer Methodik, welche die strukturellen Diskriminierungen psychologisiert und nicht nur als deskriptiven, sondern als normativen Ausgangspunkt nimmt, um eine Anstellungsentscheidung zu begründen. Somit wird Diskriminierung nicht nur in Kauf genommen, sondern geradezu gerechtfertigt und unter digitalen algorithmischen Metriken outgesourct und versteckt. Denn die Kriterienlage, d. h. die Eigenschaften der Datenbank und die Übersetzung dieser Rationale in Metriken sind in der digitalen Interaktionsoberfläche nicht mehr erfassbar. Wer ist alles in der Datenbank? Mit wem wird die bewerbende Person verglichen? Mit den Mitarbeiter*innen des Unternehmens? Mit den anderen Bewerber*innen? Je nach Antwort ergeben sich unterschiedliche Zahlen. Oder wie APPADURAI & KIFT (2020) es formulieren:

> „But metrics are not just numbers. Metrics are the numbers that matter. Numbers are empirical, metrics are normative. They are the magic that turns numbers into meanings. But for the magic of metrics to work we need to believe in them."

Problematisch wird es allerdings auch, wenn der Fokus auf die „Korrektur" von Bias den Kontext und Einbettung in menschlichen Prozessen und Verhaltensmuster zugunsten von technischen Korrekturen außen vorlässt. Eine Reflexion über die Angemessenheit der Übersetzung eines psychologischen Konzepts in algorithmische und informatische Sprachen genügt nicht, um Gender-Diskriminierung in einem algorithmischen Verfahren rein konzeptuell zu evaluieren. Es bedarf zugleich einer Reflexion der Konzepte und Methoden, die digitalisiert werden sollen, und zwar 1) an sich selbst und 2) hinsichtlich des Kontextes, in dem es evaluiert werden soll.

Das OCEAN-Modell ist methodologisch entwickelt worden, um möglichst verallgemeinernd Aussagen über emotionale Verhaltensweisen in einer Gruppe zu treffen. Dabei geht es um Beobachtungen im Durchschnitt. Diese Aussagen ergeben sich aus der Zusammenführung aller Daten von allen Individuen in der Gruppe. Eine Modifikation in der Gruppierung führt zu anderen Ergebnissen. Die Methode kann Aussagen treffen über Verhaltensweisen in Kollektiven, verliert aber an statistischer Genauigkeit, wenn sie dafür benutzt wird, Bewerber*innen personalisiert zu evaluieren. Denn nicht das Individuum ist dabei von Bedeutung, sondern seine relative Position in der Gruppe liegt am Fokus. Das wiederum bedeutet, dass die Gruppe der entscheidende Faktor wird, um die Position der Bewerber*in zu bestimmen. Damit lässt sich viel mehr über die Gruppe (die

Datenbank) sagen als über die einzeln analysierten Bewerber*innen selbst. Dabei sind die Persönlichkeitsdimensionen keineswegs weltweit allgemeingültig. Forschung in China (CHEUNG 2011), Südafrika (SAUM-ALDEHOFF 2012) und in bolivianischen Dörfern (GURVEN ET AL. 2012) zeigen, dass in anderen Kulturen andere Dimensionen, die im OCEAN-Modell nicht vorhanden sind, mehr Bedeutung haben, wie etwa Industriousness (Fleiß) oder Prosociality – oder Elternliebe in China, die anders als in vielen westlichen Ländern eine aktive Fürsorge für die Eltern impliziert. Die Befragungen von Fanny Cheung zeigen, dass dieses Verständnis auch bei US-Amerikaner*innen mit asiatischen Wurzeln vorhanden ist. Diese Methode wird also kaum heterogene Gesellschaften abbilden können. Auch in Deutschland besteht diese Heterogenität. Laut einer Erhebung des Statistischen Bundesamts von 2019 hatte jede vierte Person in Deutschland 2018 einen Migrationshintergrund, insgesamt also ca. 20,8 Millionen Menschen, davon 24,6 % Frauen (DESTATIS 2018).

Die Entwicklungen von digitalen algorithmischen Systemen im Bereich Recruiting entsprechen in ihrem Ansatz sehr stark einer Rationalisierung der Welt im europäischen Sinne der Aufklärung: die Natur und die Dinge in Frage zu stellen, herauszufordern und zu optimieren. Es geht darum, die Rationalitätslücken des Menschen zu schließen, die Mechanik an Orten einzusetzen, wo die menschliche Natur nicht rational genug ist oder angeblich in ihrem Wesen reproduziert werden kann. KI ist eine Entwicklung innerhalb der philosophischen Tradition um das Wesen der Dinge und Wesenheiten. Da die derzeitigen KI-Technologien auf Statistiken basieren, offenbaren sie bei der Verwendung relationaler Methoden zur Beantwortung von Fragen nach dem Wesen von Dingen – und damit einem universalistischen Anspruch – zugleich grundlegende konzeptionelle Widersprüche. Die ebenfalls mechanistische und essentialistische Denkweise im rechtlichen Bereich fokussiert sich stark an technischen Vorgaben und, rechtstechnisch gesehen, an kategoriale Lösungen und kreieren damit Marktanreize, die diese Dynamiken perpetuieren.

5. Empfehlungen

5.1. Empfehlungen zur Governance von ASR-Systemen

Eine ex-ante dokumentierte Folgenabschätzung oder ein Test der verschiedenen konkreten Softwareprodukte in ihrem Zusammenwirken (Aufzeigen von Wechselwirkungen mit dem technischen System und der Arbeitsumgebung an der sie angeschlossen sind) ist notwendig. Ebenfalls vonnöten wäre eine begleitende Governancestruktur mit standardisierten Controllingverfahren für Recruiting.
Um Muster zu erkennen, die auf eine potenzielle Diskriminierung hinweisen, wäre eine statistische Dokumentation der Gesamtheit der Ergebnisse eines ASR-Systems hilfreich. Die Statistik sollte die Ergebnisse im Kontext ihres Zusammenwirkens mit den im Unternehmen entwickelten Arbeitsprozessen sowie der einzelnen Recruitingschritte und verwendeten Softwareprodukte dokumentieren.

5.2. Empfehlung zu zusätzlicher Forschung

5.2.1. Studien zur Entwicklung von interdisziplinären Methodologien und wissenschaftlicher Sprache

Am Beispiel von ASR wird die Vielfalt der Disziplinen deutlich, die notwendig sind, um sozio-technische Systeme und deren Effekte zu analysieren; von Psychologie über Wirtschaftswissenschaften, Gender Studies, Arbeitssoziologie, Mathematik und Rechtswissenschaften bis zur Ethik und Organisationswissenschaften. Doch all diese Wissenschaften rezipieren wie im Kapitel 1 und 3 aufgeführt zu wenig voneinander, um ein ganzheitliches Bild im Hinblick auf ASR und Geschlechtergerechtigkeit in Deutschland abgeben zu können.
Aus wissenschaftstheoretischer Sicht ist es notwendig, Theorien und Fachsprachen zu entwickeln, die interdisziplinäre Methoden und Forschungsfragen für die Konzeptualisierung und Evaluation von algorithmischen Systemen im HR-Bereich systematisieren.

5.2.2. Studien zum Nutzungsverhalten

Bisher fehlen umfassende, öffentlich zugängliche Daten dazu, welche Personengruppen auf welchen Plattformen vertreten sind und welche Daten sie dort über sich preisgeben, insbesondere, wenn die Plattformen nicht direkt für die Arbeitssuche ausgelegt ist (z. B.

Facebook, aber auch spezialisierte *Slack*-Instanzen, z. B. bei IT-Jobs). Geben Personengruppen weniger Daten über sich preis, ist es möglich, dass sie von den Suchalgorithmen und Filtern der Recruiter*innen weniger oft gefunden werden. Auch fehlen Daten darüber, welches Spektrum an Stellen – Niedriglohntätigkeiten, prekäre oder reguläre Anstellungen, Leitungspositionen – über welche Plattformen angeboten werden. Um das Diskriminierungspotenzial, das durch die zunehmende Plattformisierung entsteht, differenziert betrachten zu können, ist es notwendig, detaillierte Daten zum individuellen Nutzungsverhalten von Plattformen im Kontext von Recruiting zu erheben bzw. Zugang dazu zu bekommen.

5.2.3. Studie zum Einsatz von algorithmischen Assistenzsystemen

Es mangelt an interdisziplinären Studien und Daten darüber, welche Technologien von den Unternehmen im Detail eingesetzt werden, wenn diese von „KI im Recruiting“ schreiben. Aus den Produktbeschreibungen lässt sich nur vage erahnen, welche Verfahren Anwendung finden und welche Daten dafür herangezogen werden. Aktuelle Analysen überprüfen die eingesetzten Technologie häufig in einem artifiziellen Setting. Es fehlt an Studien mit anthropologischer und interdisziplinärer Sicht, die den tatsächlichen Einsatz der Technologien (auch in der Zusammenwirkung an andere Technologien) unter Einbezug von Faktoren wie den involvierten Personen, Abteilungen etc. betrachten.

5.2.4. Methodologische Einordnung der eingesetzten Verfahren

Wie unter 2.5.2 und 3.2.2. aufgezeigt, bergen die Verkaufsversprechen von *Skill Matching* und *Culture Fitting* ein großes Potenzial an Diskriminierung, da ihre Qualitätsmaße, die konkreten technischen Verfahren zur Datenerhebung und die zugrundeliegenden psychologischen Persönlichkeitsmodelle, mit denen die Produkte arbeiten, intransparent sind und entgegen ihrer Verkaufsversprechen keine Aussagen über einzelne Individuen in heterogenen Gesellschaften zulassen. Es steht zu befürchten, dass die Verfahren aus den oben genannten Gründen diskriminierend wirken, ohne dass dies im Einzelfall belegbar wäre. Hier braucht es neben einer Erforschung der eingesetzten Qualitätsmaße und technischen Verfahren eine Einordnung der Funktionsweise und Belastbarkeit der psychologischen Modelle.

5.3.5. Sprachwissenschaftliche Erforschung

Analysen aus dem englischsprachigen Raum zeigen, dass beispielsweise Verfahren aus dem Bereich des NLP dazu tendieren, afroamerikanisch klingende Namen weniger positiv zu konnotieren als europäisch-amerikanische Namen, und männliche Pronomen mit „maths and sciences“ verbinden, weibliche Pronomen dagegen stärker mit dem Begriff „arts“ (SUTTON ET AL. 2018, 4). Auch hier zeigen sich die möglichen Effekte intersektioneller Diskriminierung. Für Sprachen mit grammatischem Geschlecht wie dem Deutschen ergeben sich weitere Interferenzen zwischen grammatikalischem Bias und Sachbias (*topical bias*), die bei der Bewertung einzelner NLP-Modelle untersucht werden müssen (MCCURDY & SERBETÇI 2020, 2–3). Studien, die sich konkret auf den deutschsprachigen Raum und den Kontext von Recruiting beziehen, fehlen bisher.

6. Ausblick und Fazit

Recruiting ist ein komplexer Vorgang, der aus verschiedenen Elementen besteht und zugleich selbst Teil eines größeren Verfahrens im Bereich Personalmanagement ist. Die Digitalisierung dieser Prozesse mit der Nutzung komplexer algorithmischer Systeme steht in Deutschland noch am Anfang. Für eine Analyse der digitalen Produkte im Recruiting ist nicht nur ein punktueller Blick auf einzelne Schritte und Verfahren innerhalb des Recruitingvorgangs notwendig, sondern es bedarf auch eines systemischen Blicks auf die Einbettung der technischen Systeme in ihren Verwendungskontext.
Durch den Einzug digitaler Dienstleistungen in Deutschland befassen sich Ingenieurswissenschaften und andere technische Disziplinen mit Sektoren, die den technischen Wissenschaften fremd sind. Ein interdisziplinärer Austausch findet kaum statt. Das führt zum Teil zu digitalen Dienstleistungen, die auf szientistischen Annahmen basieren und Genderdiskriminierung amplifizieren.
Die rechtliche Sicht zur intersektionalen Gendergerechtigkeit konstruiert in algorithmischer Denkmanier Kategorien und Anforderungen, die die prozedurale und relationale Natur von intersektionaler Genderdiskriminierung nicht abbilden, sondern eine starre, verzerrte algorithmische Sicht perpetuieren. Durch Kategorienlisten bleiben neue vulnerablen Gruppen und Kontexte unberücksichtigt. Eine Rechtsentwicklung, die mehr auf gerichtliche Analogie-Fallbildung, ähnlich wie im Äußerungsrecht setzt, wäre ein gangbarer Weg, um den Fokus weg von Kategorienlisten hin zu Kontexten und der Relationalität von intersektionaler Genderdiskriminierung in algorithmischen Systemen zu lenken.
Die derzeitig auf dem Markt angebotenen Produkte werden mehrheitlich von mittleren und größeren Unternehmen eingesetzt. Softwaredienstleistungen, insbesondere von größeren Unternehmen, werden primär in englischer Sprache entwickelt. Die gendergerechten sprachlichen Konventionen sind in dem Kontext anders als auf Deutsch. Die deutsche Umsetzung ist in manchen Fällen eine Übertragung angelsächsischer sprachlicher Konventionen und kann bspw. zu Problemen bei der Erkennung von generischem Maskulinum und der Entwicklung von gendergerechten Sprachkorrekturen oder Suchalgorithmen (im Falle von Suchplattformen) führen. Die Identifikation solcher Unterschiede und die Entwicklung von Lösungsansätzen stehen noch am Anfang. Ein Grund hierfür ist, dass die meisten ASR-Produkte als Teilsysteme oder -anwendungen entwickelt werden, die an weitere digitale Infrastrukturen angeschlossen werden sollen (*Slack*, *SAP Success Factors* oder andere HR-Software). Durch diese Vernetzung mit anderen proprietären Infrastrukturen entstehen Kollisionen zwischen den Logiken und Metriken der involvierten Systeme und deren Nutzer*innen. Es bedarf mehr Forschung, die sich mit

der tatsächlichen Implementierung beschäftigt und einen interdisziplinären, systemischen Blick trägt.
Neben der Entwicklung von Recruitingsystemen, die nicht als alleinstehendes Produkt funktionieren, sondern als Teil eines Produktsystems und als anschlussfähige Software an andere digitale Infrastrukturen, sind es primär zwei weitere Trends, die auf dem deutschen Markt aus der Sicht intersektionaler Genderdiskriminierung problematisch erscheinen.

6.1. Fehlende Reflexion und Berücksichtigung der bestehenden Strukturen und Asymmetrien im Arbeitsmarkt

Viele der untersuchten Produkte haben als Benchmark den Vergleich von Leistungen bereits angestellter Mitarbeiter. Die Fokussierung auf vergangene Strukturen als normativer Maßstab setzt auf eine Fortführung bereits bestehender Strukturen und somit Asymmetrien. Diese Art von Vergangenheit als Proxy operiert zudem mit sozial-homophilen Annahmen. Solche Annahmen in algorithmischen Systemen gehen davon aus, dass es wünschenswert ist, Bewerber*innen einzustellen, deren Profile denen der anderen Angestellten ähneln. Auch dies verschärft die Perpetuierung bereits bestehender Genderasymmetrien.

6.2. Die Ökonomisierung von Emotionen und Charakter

Die Ökonomisierung von Emotionen, d. h. die Normierung von Emotionen und Charakteren im Beruflichen und deren Kapitalisierung erfolgte anfangs des 20. Jahrhunderts. Die Entwicklung von allgemeingültigen, mathematisierbaren Metriken wurde in der Wissenschaft strittig diskutiert und vorangetrieben. Doch die Umsetzung solcher Metriken (wie etwa das *OCEAN Model*) im beruflichen Kontext wurde durch die Erweiterung maschineller Speicherkapazitäten, der Basis für den Betrieb von dafür notwendigen großen Datenbanken, erst im großen Stil möglich. Die Modelle arbeiten in algorithmischen Systemen mit stark gegenderten Emotionen und Charaktereigenschaften. Sie dienen der Beurteilung nicht fachlicher Fähigkeiten, sondern der emotionalen Verfasstheit und Ähnlichkeit mit der Firmenkultur. Diese Charakter- und Emotionenmärkte psychologisieren auf codierte Art und Weise intersektionale Genderasymmetrien und nehmen gesellschaftliche Asymmetrien und Vorurteile als normativer Maßstab für individuelle Evaluationen der Bewerber*innen. Unternehmen, die Technologien basierend auf diesen Modellen nutzen, sind derzeit noch in der Minderheit (siehe Abschnitt 3.).

Viele diese Technologien werden sich womöglich nicht durchsetzen, andere werden sich, solange sie nicht hinterfragt werden und den Anschein von Effektivität erwecken, auf dem Markt ausbreiten. Auch die Entwicklung neuer Recruitingansätze, die in rechtliche Lücken stoßen, wie etwa das *Crowdworking*, stehen in Deutschland noch am Anfang. Es besteht ein dringender Bedarf für begleitende, interdisziplinäre Forschung und eine sektorale Aufsicht solcher Systeme, die in der Lage ist, jenseits der isolierten Software-Evaluation kontextuelle, systemische Evaluationsmethoden, Sprachen und normative Verfahren zu entwickeln.

Executive Summary

Recruiting ist ein komplexes Verfahren, das aus verschiedenen Elementen besteht und zugleich selbst Teil des größeren Bereichs Personalmanagement ist. Die hierfür entwickelten Anwendungen rangieren von kleinen, alleinstehenden Anwendungen wie Chatbots über Anwendungen, die in größere Softwareprodukte integriert werden, bis hin zu Produktfamilien mit einer Vielzahl von Anwendungsbereichen. Durch den Einzug dieser digitalen HR-Anwendungen in Deutschland befassen sich Ingenieurwissenschaften und andere technische Disziplinen mit einem Bereich, der ihnen eigentlich fremd ist. Ein interdisziplinärer Austausch zwischen den techniknahen und technikferneren Bereichen findet kaum statt. Das führt zum Teil zu digitalen Produkten und Dienstleistungen, die auf szientistischen Annahmen basieren und Diskriminierung verstärken, bspw. durch die Ökonomisierung von Emotionen und Charaktereigenschaften und deren Digitalisierung unter den Stichwörtern Skills, Culture Fit oder emotionale Kompetenz. Sie dienen weniger der Beurteilung fachlicher Fähigkeiten, sondern der Erfassung der emotionalen Verfasstheit und dem Abgleich mit Ähnlichkeiten in Bezug auf die Firmenkultur. Die so entwickelten Recruitingprodukte nehmen gesellschaftliche Asymmetrien und Vorurteile als normativen Maßstab für die individuelle Bewertung von Bewerber*innen.
Die Anwendungen berücksichtigen kaum potenzielle Widersprüche zwischen den Rechtssystemen der Länder, in denen sie entwickelt oder eingesetzt werden.
Für eine Analyse der digitalen Produkte im Bereich Recruiting hinsichtlich Gendergerechtigkeit bedarf es eines ganzheitlichen Blicks, der über die Betrachtung der bloßen Softwarefunktionen hinausgeht und die soziale und rechtliche Einbettung mit aufgreift. Die vorliegende Expertise stellt mögliche Ansätze dafür vor.

Literaturverzeichnis

ALLPORT 1937: Gordon W. Allport. *Personality: A psychological interpretation.* New York: Henry Holt. 1937.

ANDREWS 2017: T.L. Andrews. *Just a few words can increase female and minority job applicants by more than 20%:* https://qz.com/1023518/just-a-few-words-can-increase-female-and-minority-job-applicants-by-over-20/ (zuletzt abgerufen am 01.07.2020).

APPADURAI & KIFT 2020: Arjun Appadurai & Paula Kift. *Beware the magic of metrics.* In: Eurozine, 27. Juni 2020: https://www.eurozine.com/beware-the-magic-of-metrics/?fbclid=IwAR3Rbr97mvdJQtSpM0XyfhVHAeYL4awMAwgdjDyVumJTAq97iP2rU6S8_TI (zuletzt abgerufen am 01.07.2020).

ASIF 2015: Talha Asif. *Here Are Tennis Top 10 Worst Falls on Grass Courts in Recent Times*. In: Tennisworldusa.org: https://www.tennisworldusa.org/tennis/news/Editors_Thoughts/24495/here-are-tennis-top-10-worst-falls-on-grass-courts-in-recent-times-/ (zuletzt abgerufen am 21.07.2020).

ATHANAS & BARISLOVITS 2020: Christoph Athanas & Stefan Barislovits. Recruiting in der Coronakrise. Eine Studie zu Krisenreaktionen und Auswirkungen auf die Personalgewinnung von stellenanzeigen.de und der meta HR Unternehmensberatung: https://recruitingnews.stellenanzeigen.de/recruitingindercoronakrise (zuletzt abgerufen am 01.07.2020).

BAER ET AL. 2010: Susanne Baer, Melanie Bittner & Anna Lena Göttsche. *Mehrdimensionale Diskriminierung – Begriffe, Theorien und juristische Analyse:* https://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/publikationen/Expertisen/Expertise_Mehrdimensionale_Diskriminierung_jur_Analyse.pdf?__blob=publicationFile (zuletzt abgerufen am 01.07.2020).

BARABAS ET AL. 2020: Chelsea Barabas, Colin Doyle, JB Rubinovitz & Karthik Dinakar. *Studying up: reorienting the study of algorithmic fairness around issues of power.* In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* ’20). Association for Computing Machinery, New York, NY, USA, 167–176. https://doi.org/10.1145/3351095.3372859.

BEHAGHEL ET AL. 2015: Luc Behaghel, Bruno Crépon & Thomas Le Barbanchon. *Unintended Effects of Anonymous Résumés.* American Economic Journal: Applied Economics, 7 (3): 1-27.

BIERNAT & FUEGEN 2001: Monica Biernat & Kathleen Fuegen. *Shifting Standards and the Evaluation of Competence: Complexity in Gender-Based Judgment and Decision Making.* Journal of Social Issues, 57: 707-724. https://doi.org/10.1111/0022-4537.00237.

BINDER & HESS 2011: Beate Binder & Sabine Hess. “Intersektionalität aus der Perspektive der Europäischen Ethnologie”. In: Sabine Hess et al. (Hg.). Intersektionalität revisited Empirische, theoretische und methodische Erkundungen. Bielefeld, 2011: S. 15-54.

BOGEN & RIEKE 2018: Miranda Bogen & Aaron Rieke. *Help Wanted. An Examination of Hiring Algorithms, Equity, and Bias:* https://www.upturn.org/reports/2018/hiring-algorithms/ (zuletzt abgerufen am 01.07.2020).

BÖSCHEN ET AL. 2012: Ines Böschen, Ramona Alt, Annabelle Krause, Ulf Rinne, Klaus F. Zimmermann. *Pilotprojekt„Anonymisierte Bewerbungsverfahren“ – Abschlussbericht:* https://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/publikationen/AnonymBewerbung/Abschlussbericht-anonymisierte-bewerbungsverfahren-20120417.pdf;jsessionid=ADB6E92BB795FF9168858312398EFCBE.2_cid351?__blob=publicationFile&v=3 (zuletzt abgerufen am 21.07.2020).

BOWER ET AL. 2017: Amanda Bower, Sarah N. Kitchen, Laura Niss, Martin J. Strauss, Alexander Vargo & Suresh Venkatasubramanian. *Fair Pipelines*. FAT/ML 2017, arXiv:1707.00391v1 [cs.CY], Juli 2017.

BUNDESKRIMINALAMT 2019: Partnerschaftsgewalt, *Kriminalstatistische Auswertung. 2019:* https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Partnerschaftsgewalt/Partnerschaftsgewalt_2019.html;jsessionid=416BE750C309508BFFD4AB21EE75F8C8.live0612?nn=63476 (zuletzt abgerufen am 01.07.2020)

BUOLAMWINI 2017: Joy Adowaa Buolamwini. *Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers* (PhD thesis). MIT. hdl:1721.1/114068

BUREL ET AL. 2019: Simone Burel, Daniel Spitzer & Simon Tschürtz. Deutsche Stellenausschreibungen unterscheiden zwischen Mann und Frau: Wie geschlechtsspezifische Sprache die Ungleichheit zwischen den Geschlechtern verfestigt. https://www.lub-mannheim.de/wp-content/uploads/Artikel_Genderbias-1.pdf (zuletzt abgerufen am 30.06.2020).

BVMW o.J.: Webseite des Bundesverbands Mittelständische Wirtschaft Unternehmerverband Deutschlands e.V.: https://www.bvmw.de/themen/mittelstand/zahlen-fakten/ (zuletzt abgerufen am 01.07.2020).

CHEFSACHE REPORT 2018: Chefsache Report. August 2018, 2: https://initiative-chefsache.de/content/uploads/2018/08/Chefsache-Report-2018.pdf (zuletzt abgerufen am 01.07.2020).

CHEUNG 2011: Fanny M. Cheung, Fons J. R. van de Vijver & Frederick T. L. Leong. *Toward a new approach to the study of personality in culture.* American Psychologist 66 (7), 2011: S. 593–603.

CHOWDHURY 2005: Chowdhury, Gobinda. *Natural language processing*. ARIST. 37. 51-89. 10.1002/aris.1440370103.

CHUN 2018: Wendy Hui Kyong Chun. *Queering Homophily: Muster der Netzwerkanalyse*. In: Zeitschrift für Medienwissenschaften (2018) Nr. 18, 131-148. https://doi.org/10.25595/502.

COARD 2020: Callum Coard. *Facebook Targeting Options Explained*: https://www.business2community.com/facebook/facebook-targeting-options-explained-02291947 (zuletzt abgerufen am 01.07.2020).

COLLIER & ZHANG 2016: Danielle Collier & Charlotte Zhang. Can We Reduce Bias in the Recruiting Process and Diversify Pools of Candidates by Using Different Types of Words in Job of Candidates by Using Different Types of Words in Job Descriptions? Zuletzt abgerufen am 01.07.2020 von Cornell University, ILR School site: http://digitalcommons.ilr.cornell.edu/student/140

COSTA & MCCRAE 1985: Paul T. Costa Jr. & Robert. R. McCrae. *The NEO Personality Inventory Manual.* Odessa, FL: Psychological Assessment Resources, 1985.

CRAWFORD & PAGLEN 2019: Kate Crawford & Trevor Paglen, *Excavating AI: The Politics of Training Sets for Machine Learning* (19. September 2019): https://www.excavating.ai/ (zuletzt abgerufen am 01.07.2020).

DAVENPORT & RONANKI 2018: Thomas H. Davenport & Rajeev Ronanki. *Artificial Intelligence for the Real World.* Harvard Business Review, January–February 2018, 108–116.

DESTATIS 2018: Statistisches Bundesamt (Destatis). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund – Ergebnisse des Mikrozensus 2018.* Statistisches Bundesamt Fachserie 1 Reihe 2.2, 2018: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-2010220187004.pdf?__blob=publicationFile (zuletzt abgerufen am 01.07.2020).

DESTATIS 2020: Statistisches Bundesamt (Destatis). *Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikations-technologien.* Statistisches Bundesamt Fachserie 15 Reihe 4, 2020: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publikationen/Downloads-IT-Nutzung/private-haushalte-ikt-2150400197004.pdf;jsessionid=407CB4947ECA4C6931699810A5FF96E6.internet8722?__blob=publicationFile (zuletzt abgerufen am 21.07.2020).

FOBES 2016: The Muse. 3 Things You Should Know About The Robots Reading Your Resume, But Don't. https://www.forbes.com/sites/dailymuse/2016/10/20/3-things-you-should-know-about-the-robots-reading-your-resume-but-dont/?sh=5630816f2562#4747e12e2562 (zuletzt abgerufen am 21.07.2020).

FORD 2004: Donna Y. Ford. *Intelligence Testing and Cultural Diversity: Concerns, Cautions and Considerations,* The National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT. 2004.

GADE ET AL. 2018: Christel D. Gade, Kai H. Helfritz & Jörg Murmann. *Recruiting zwischen Executive Search und Digitalisierung. Eine Kooperationsstudie von iubh, BDU und DGFPIUBH.* Discussion Papers, Reihe: Business & Management, Vol. 1, Issue 3 (Nov. 2018): https://www.bdu.de/media/352926/studie_digitalisierung-in-der-personalberatung_2018.pdf. (zuletzt abgerufen am 01.07.2020).

GIOLLA & KAJONIUS 2019: Eric Mac Giolla & Petri J. Kajonius. (2019), Sex differences in personality are larger in gender equal countries: Replicating and extending a surprising finding. Int J Psychol, 54: 705-711.

GOLDBERG 1981: Lewis R. Goldberg. *Language and individual differences: The search for universals in personality lexicons.* In: Wheeler (Ed.), Review of Personality and social psychology*,* 1981, Vol. 1, 141–165. Beverly Hills, CA: Sage.

GRANT 2013: Adam Grant. *Goodbye to MBTI, the Fad That Won't Die:* https://www.psychologytoday.com/intl/blog/give-and-take/201309/goodbye-mbti-the-fad-won-t-die (zuletzt abgerufen am 01.07.2020).

GURVEN ET AL. 2012: Michael Gurven, Christopher von Rueden, Maxim Massenkoff, Hillard Kaplan & Marino Vie. How universal is the big five? Testing the five-factor model of personality variation among forager-farmers in the bolivian amazon Journal of Personality and Social Psychology 104(2), S. 354–370.

HARRIS 2016: Jensen Harris. *Everything Textio is new again:* https://textio.com/blog/everything-textio-is-new-again/13034792954#.40z32cks9 (zuletzt abgerufen am 01.07.2020).

HEISE 2000: Elke Heise. *Sind Frauen mitgemeint? Eine empirische Untersuchung zum Verständnis des generischen Maskulinums und seiner Alternativen*. Sprache & Kognition, 19(1-2), 3–13. https://doi.org/10.1024//0253-4533.19.12.3

HERMANN 1971: Donald H. Hermann III. *Privacy, the Prospective Employee, and Employment Testing: The Need to Restrict Polygraph and Personality Testing*. 47 Wash.L. Rev.73 (1971): https://digitalcommons.law.uw.edu/wlr/vol47/iss1/4 (zuletzt abgerufen am 01.07.2020).

HOFFMAN ET AL. 1954: Paul J. Hoffman, Leon Festinger, & Douglas H. Lawrence. (1954). *Tendencies toward Group Comparability in Competitive Bargaining*. Human Relations, 7(2), 141–159.

HUCZYNSKI 1993: Andrzej Huczynski. *Management Gurus*. Routledge, 1993.

ILLOUZ 2009a: Eva Illouz. *Capitalism with feelings.* Vortrag von Eva Illouz am 6. Juli 2009 in der Volksbühne im Prater: https://ifg.rosalux.de/2009/07/08/kapitalism-with-feelings-vortrag-von-eva-illouz-soziologin-der-hebrew-university-am-6-juli-2009-in-der-volksbuhne-im-prater/ (zuletzt abgerufen am 01.07.2020).

JÄGER & MEURER 2018: Wolfgang Jäger & Sebastian Meurer. *Recruiting Strategien 2018. Erfolgreiche Instrumente zur Bewerbersuche.* Wolters Kluwer Deutschland GmbH, Köln 2018: https://www.personalwirtschaft.de/assets/documents/Downloads/Studienband-Recruiting-Strategien-2018.pdf (zuletzt abgerufen am 01.07.2020).

JONASON ET AL. 2018: Peter K. Jonason, Monica A. Koehn, Ceylan Okan & Peter J. O'Connor. *The role of personality in individual differences in yearly earnings.* Personality and Individual Differences, Volume 121, 2018, S. 170-172.

JUNG 2017: Sybille Jung. *Active Sourcing:* https://softgarden.de/ressourcen/glossar/active-sourcing-2/ (zuletzt abgerufen am 01.07.2020).

KARWOWSKI ET AL. 2013: Maciej, Karwowski, Izabela Lebuda, Ewa Wisniewska & Jacek Gralewski. *Big Five Personality Traits as the Predictors of Creative Self-Efficacy and Creative Personal Identity: Does Gender Matter?* J Creat Behav, 47: 215-232. doi:10.1002/jocb.32

KNOBLOCH & HUSTEDT 2019: Tobias Knobloch & Carla Hustedt. *Der maschinelle Weg zum passenden Personal. Zur Rolle algorithmischer Systeme in der Personalauswahl:* https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/SNV_Robo_Recruiting_final.pdf (zuletzt abgerufen am 01.07.2020).

KOSCHICK 2020: Anne Koschick. *Warum digitale Bewerbungsprozesse an Fahrt gewinnen:* https://www.karriere.de/mein-naechster-job/coronakrise-und-recruiting-warum-digitale-bewerbungsprozesse-an-fahrt-gewinnen/25857526.html (zuletzt abgerufen am 01.07.2020).

KRUEGER & LISCHKA 2018: Julia Krüger & Konrad Lischka. *Damit Maschinen den Menschen dienen:* https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Algorithmenethik-Loesungspanorama.pdf (zuletzt abgerufen am 30.06.2020).

LI 2016: Peng Lai "Perry" Li, *Natural Language Processing*. 1 GEO. L. TECH. REV. 98: https://www.georgetownlawtechreview.org/natural-language-processing/GLTR-11-2016 (zuletzt abgerufen am 01.07.2020).

LIPPA 1995: Richard Lippa. (1995). *Gender-related individual differences and psychological adjustment in terms of the Big Five and circumplex models.* Journal of Personality and Social Psychology, 69(6), 1184–1202. https://doi.org/10.1037/0022-3514.69.6.1184.

MANPOWER 2017: Manpower Group. *Swipe Right: Candidate Technology Preferences During the Job Search*. https://talentsolutions.manpowergroup.com/wcm/connect/f19c8642-ca3f-4edf-bd3c-9937bde7ae39/MPGS_SwipeRight_lo.pdf?MOD=AJPERES&CVID=l.pW4wJ (zuletzt abgerufen am 01.07.2020).

MCCURDY & SERBETÇI 2020: Katherine McCurdy & Oğuz Serbetçi. *Grammatical gender associations outweigh topical gender bias in crosslinguistic word embeddings.* arXiv:2005.08864v1 [cs.CL], Mai 2020.

NORMAN 1963: Warren T. Norman. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. Journal of Abnormal and Social Psychology, 1963, 66, 574– 583.

OKOLIE & IRABOR 2017: Ugo Chuks Okolie & Ikechukwu Emmanuel Irabor. *E-Recruitment: Practices, Opportunities and Challenges.* European Journal of Business and Management Vol.9, No.11, 116–122.

ORWAT 2019: Carsten Orwat. Studie *Diskriminierungsrisiken durch Verwendung von Algorithmen*, Antidiskriminierungsstelle des Bundes. Nomos Verlag. 2019: https://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/publikationen/Expertisen/Studie_Diskriminierungsrisiken_durch_Verwendung_von_Algorithmen.pdf?__blob=publicationFile&v=5 (zuletzt abgerufen am 01.07.2020).

ONLINE-RECRUITING.NET 2015: *Social Media Recruiting Studie 2015 Deutschland*: https://www.online-recruiting.net/ressource/social-media-recruiting-studie-2015-deutschland/ (zuletzt abgerufen am 21.07.2020).

PESCH 2018: Ulrich Pesch. *Machine Learning.* In: Personalwirtschaft Recruiting Guide 2018, S. 25-27.

PLENKE 2015*:* Max Plenke. *The Reason This "Racist Soap Dispenser" Doesn't Work on Black Skin*: https://www.mic.com/articles/124899/the-reason-this-racist-soap-dispenser-doesn-t-work-on-black-skin (zuletzt abgerufen am 22.07.2020)

PRINCE & SCHWARCZ 2020: Anya Prince & Daniel B. Schwarcz. *Proxy Discrimination in the Age of Artificial Intelligence and Big Data* (August 5, 2019). 105 Iowa Law Review 1257–1318 (2020). Abrufbar unter SSRN: https://ssrn.com/abstract=3347959.

RECRUITING 2019/2020: Recruiting Trends 2019 und 2020 des Centre of Human Resources Information Systems (CHRIS) der Universität Bamberg und der Friedrich-Alexander-Universität Erlangen-Nürnberg im Auftrag der Monster Worldwide Deutschland GmbH: https://www.uni-bamberg.de/isdl/chris/recruiting-trends/recruiting-trends-2019/ (zuletzt abgerufen am 01.07.2020).

RENFRO 2019: Kim Renfro. *The incredible life and career of tennis prodigy Naomi Osaka, who just won $2.9 million at the Australian Open.* In: Insider.com: https://www.insider.com/naomi-osaka-career-biography-photos-2019-1 (zuletzt abgerufen am 21.07.2020).

ROPOHL 2009: Günter Ropohl. *Allgemeine Technologie. Eine Systemtheorie der Technik.* 3., überarbeitete Auflage. Karlsruhe 2009.
ROTH 2009: Lorna Roth. *Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity*. Canadian Journal of Communication, 34(1). https://doi.org/10.22230/cjc.2009v34n1a2196.

SANCHEZ-MONEDERO ET AL. 2020: Javier Sánchez-Monedero, Lina Dencik & Lilian Edwards. What does it mean to „solve“ the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* ’20). Association for Computing Machinery, New York, NY, USA, 458–468. DOI:https://doi.org/10.1145/3351095.3372849.

SAP 2019: SAP Success Factors: Learning Language Pack Availability. Administration guide, Document Version Q3 2019 – 2019-08-10: https://help.sap.com/doc/e0400a6228d54188875ba4adde13c83e/1908/en-US/SF_LRN_Lang_Pack_Avail.pdf (zuletzt abgerufen am 21.07.2020).

SATOW 2012: Lars Satow. *Big-Five-Persönlichkeitstest (B5T): Test- und Skalendokumentation*: http://www.drsatow.de. (zuletzt abgerufen am 01.07.2020).

SAUM-ALDEHOFF 2012: Thomas Saum-Aldehoff. *Big Five: Sich selbst und andere erkennen.* Patmos, Ostfildern 2012.

SCHÜNEMANN & LEBERT 2019: Isabel Schünemann & Yannick Lebert. *Algorithmen & Gesellschaft.* Zur Zukunft sozialer Teilhabe in Deutschland: https://www.vodafone-institut.de/wp-content/uploads/2019/10/Algorithmen_und_Gesellschaft.pdf (zuletzt abgerufen am 01.07.2020)

SHOUSE 2005: Eric Shouse, *Feeling, emotion, affect*. M/C journal: A Journal of Media and Culture 8 (2005): www.media-culture.org.au (zuletzt abgerufen am 01.07.2020).

SPILLANE 2012: Robert Spillane. *Why workplaces must resist the cult of personality testing*: https://theconversation.com/why-workplaces-must-resist-the-cult-of-personality-testing-5540 (zuletzt abgerufen am 30.06.2020).

STEPSTONE 2020: Stepstone und Bundesverband der Personalmanager. *Digitales Bewerben und Recruiting im Praxistest*: https://www.stepstone.de/wissen/stepstone-BPM-digitales-recruiting/ (zuletzt abgerufen am 01.07.2020).

SUTTON ET AL. 2018: Adam Sutton, Thomas Lansdall-Welfare & Nello Cristianini. *Biased Embeddings from Wild Data: Measuring, Understanding and Removing:* arXiv:1806.06301v1 [cs.CL], Juni 2018.

VESTERLING 2018: *Pre-Employment-Screening – richtig umgesetzt: https://www.vesterling.com/fr/aktuell/blog/blog-post/2018/1/16/pre-employment-screening-richtig-umgesetzt* (zuletzt abgerufen am 21.07.2020).

WALGENBACH 2012: Katharina Walgenbach. Intersektionalität - eine Einführung: www.portal-intersektionalität.de (zuletzt abgerufen am 20.10.2020).

WEISBERG ET AL. 2011: Yana J. Weisberg, Colin G. DeYoung, Jacob B. Hirsh. *Gender Differences in Personality across the Ten Aspects of the Big Five.* Front Psychol. 2011;2:178. Published 2011 Aug 1. doi:10.3389/fpsyg.2011.00178

WIDDOWSO 2018: Niki Widdowso. *Gender pay gap—personality affects income*. In: Medical Xpress. 18. April 2018: https://medicalxpress.com/news/2018-04-gender-gappersonality-affects-income.html (zuletzt abgerufen am 01.07.2020).

Über die Autorinnen

Lorena Jaume-Palasi (The Ethical Tech Society)

Lorena Jaume-Palasí ist Gründerin von The Ethical Tech Society, einer gemeinnützigen Initiative, die das Ziel verfolgt, Prozesse der Automatisierung und Digitalisierung zu erforschen und in Bezug auf ihre gesellschaftliche Relevanz normativ einzuordnen. Lorena forscht zur Ethik der Digitalisierung und Automatisierung. Sie befasst sich in diesem Zusammenhang auch mit rechtsphilosophischen Fragen. 2017 wurde sie von der Regierung Spaniens in den Weisenrat zu Künstlicher Intelligenz und Datenpolitik berufen und ist ehem. Mitglied der High Level Expert Group on Artificial Intelligence der EU Kommission. Sie ist Fellow der Bucerius Stiftung und Mitglied des Beirats Bildung und Diskurse des Goethe Instituts. Lorena war mehrfach Sachverständige für den Bundestag, das Europäische Parlament und die Europäische Kommission zur künstlichen Intelligenz und Ethik. Lorena wird regelmäßig von internationalen Organisationen, Verbände und Regierungen konsultiert. Sie ist hat diverse Publikationen zu Internet Governance mitverfasst und herausgegeben und schreibt regelmäßig zu Datenschutz, Privatheit und Öffentlichkeit, öffentliche Güter und Diskriminierung.

2018 erhielt sie für die von ihr mitgegründeten Initiative AlgorithmWatch die Theodor Heuss Medaille „für ihren Beitrag zu einer differenzierten Betrachtung von Algorithmen und deren Wirkmechanismen“.

Elisa Lindinger ist Mit-Gründerin von Superrr – einem feministischen Think Tank mit Sitz in Berlin. Sie arbeitet an der Schnittstelle von Technologie- und Geisteswissenschaften und setzt sich in ihrer derzeitigen Forschung mit der Analyse soziotechnischer Systeme auseinander, beispielsweise der Erforschung von Praktiken, Normen und Werten, die die Entwicklung digitaler Infrastrukturen beeinflussen. Darüber hinaus berät sie verschiedene Einrichtungen wie z.B. die Helmholtz Gemeinschaft und die EU-Kommission, sowie große internationale Konferenzen (z. B. die Konferenzreihe re:publica) zu Digitalthemen und Technologiefragen im gesellschaftlichen Kontext.

Vor ihrer Tätigkeit bei Superrr war Elisa unter anderem bei der Open Knowledge Foundation Deutschland, der Humboldt Universität zu Berlin und der HTW Berlin in diversen Projekten zur Wissenschaftskommunikation, Technologieförderung, Machine Learning, Statistik und Datenverarbeitung tätig. Sie war geladene Expertin zum Thema "Öffentliche Daten" in der Enquete-Kommission "Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale" des Deutschen Bundestags.

Julia Kloiber ist Digitalexpertin und Gründerin. In den letzten zehn Jahren hat sie zahlreiche Initiativen und Programme an der Schnittstelle von Technologie und Gesellschaft gestartet. Sie ist Mit-Gründerin des Prototype Fund, einem 8 Millionen Euro Förderprogramm für Open Source Innovationen und der gemeinnützigen Organisation Superrr. Sie arbeitet und forscht schwerpunktmäßig zu den Themen Data-Governance und Diversität. Der Fokus ihrer Arbeit liegt auf den Chancen und Potenzialen von Technologie für Gesellschaft, Demokratie und nachhaltiger Wirtschaft. Als Partner bei Ashoka ist sie für den Aufbau eines internationalen Schwerpunkts zu Tech Policy verantwortlich. Als Digitalexpertin berät sie Politik, Verwaltung und Wirtschaft zu den gesellschaftlichen Auswirkungen von Technologie, u.a. die Deutsche Bahn AG, die OECD, die Österreichische Bundesregierung und die Daimler AG. Gemeinsam mit Elisa Lindinger hat sie in 2019 die gemeinnützige Organisation Superrr gegründet.

Julia Kloiber ist Senior Fellow der Mozilla Foundation, sie ist Beiratsmitglied für Offene Wissenschaft bei Wikimedia Deutschland und der Deutschen Postcode Lotterie. Sie war im Beirat des Weizenbaum Instituts für Internet und Gesellschaft. Ihre Arbeit wurde in nationalen und internationalen Medien publiziert. Am 02.03. ist sie als Expertin zum Thema “Frauen in der KI” bei der Enquete-Kommission "Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale" des Deutschen Bundestags.

Anhang

Produktname	Produkt	Firma	Link/Website	Recruitingschritte	Use Case	Kunden	Status	Versprechen	Land
AllyORecruit, AI Recruiting Automation Package and Business Texting for Recruiting & HR Solution	Recruiting Automation Package (Advertisement, Job-Matching, Pre-screen, Assessment, Management, Schedule	AllyO	https://www.allyo.com	Sourcing, Kandidatenmanagement, Screening	HR-Abteilungen	Randstad, G4S, Staples, Marathon Petroleum Corporation	produktiv	Conversational Recruiting über verschiedene Kanäle	USA
Arya Pulse / Arya Quantum	Sourcing und Bewerbermanagement-Plattform	Leoforce	https://goarya.com/	Sourcing, Screening	HR-Abteilungen	Lenovo, Parsons, Korn Ferry	produktiv	Beschleunigung der Einstellungsverfahren für mittelgröße und große Unternehmen	USA
Cammio	Automatisierte Video-Interview Plattform	Cammio	https://carerix.com/de/product/videopitch-automatedinterviews/ https://cammio.com/our-platform/	Screening, Auswahlverfahren	HR-Abteilungen	EU Careers (Europäisches Amt für Personalauswahl), Metro, AEGO, Randstad	produktiv	Beurteilung der Kandidaten auch anhand von Soft Skills	NL, Deutschland, USA, Singapur, Südafrika
Checkr	Background Checks Plattform	Checkr	https://checkr.com/	Screening	HR-Abteilungen und Recruiter*innen	Uber, Lyft, Partner mit ua SAP Success Factors	produktiv	Schutz und Verbesserung der Markenreputation, Risikominimierung bei Anstellungen, Faire Anstellungen.	USA
Eightfold	Matching Plattform	Eightfold AI Inc.	https://eightfold.ai	Sourcing	HR-Abteilungen	Hertz, PizzaHut, The New York Times, Macys, United, Walgreens	produktiv	Talent Intelligence für Talent -Management, -Akquise, -Diversity und -Erfahrung. "Enterprise Enlightment", debiasing, reduziert Personalfluktuation	USA
Facebook Jobs	Stellenanzeigen	Facebook	https://www.facebook.com/jobs?ref=fbb_jobs#	Sourcing, Kandidatenmanagement	Stellenanzeigen, insb. für KMU	KMU			USA
Facebook Marketplace	Stellenanzeigen	Facebook	https://www.facebook.com/marketplace?ref=fbb_jobs#	Sourcing, Kandidatenmanagement	Stellenanzeigen, insb. für KMU	KMU	produktiv		USA
Greenhouse Recruiting	Applicant Tracking System Plattform	Greenhouse	https://www.greenhouse.io/recruiting	Screening	HR-Abteilungen; für Unternehmen, die schnell ihren Personal erweitern müssen, für KMU als günstige Gesamtlösung	KMU, Airbnb, FitBit, Pinterest, Snapchat, Instacart, TripAdvisor, Reddit	produktiv	Höhere Effizienz mit KI	USA, Irland
Hiretual AI Screening	Screening & Shortlisting, Bewerber Markt-Analytics	HireTeamMate, Inc.	https://hiretual.com/features/screening/	Screening, Auswahlverfahren	Recruiter*innen, HR-Abteilungen	Cornerstone, Appito	produktiv	Reduktion des Zeitaufwands für die Prüfung von Antragstellern um 90%., effizienteres Sourcing, Bewerbermarkt-Analytics, mehr Diversität und Inklusion.	USA
Hiretual AI Sourcing	Sourcing	HireTeamMate, Inc.	https://hiretual.com/features/sourcing/	Sourcing	Recruiter*innen		produktiv	Ein Wissensnetzwerk, das Daten über Beziehungen, Bedeutungen, Personen und Orte speichert. Selbstlernende und sich selbst erweiternde Infrastruktur.	USA
HR Chatbots	Chatbot & Avatar	living actor	https://corporate.livingactor.com/en/homepage/	Sourcing	Recruiter*innen	Audi, Barclays, Airbus, Nokia, Toshiba, Sopra Steria, Atos	produktiv	Chatbot mit Avatar für personalisierte auf das Unternehmen angepasste Recruiting-Experiences	USA
Human Intelligence Hiring TM	Referenz-Checking und Sourcing Plattform	crosschq Inc	https://crosschq.com/product/	Screening, Auswahlverfahren	Recruiter*innen, HR-Abteilungen	Nerdwallet, Eventbrite, Unilog, billtrust	produktiv	Überprüfung von Referenzen, um Unternehmen bei besseren Einstellungen zu unterstützen.	USA
Intelligent screening software	Screening & Shortlisting	ideal.	https://ideal.com/product/screening/	Screening	HR-Abteilungen	Indigo, Burlington	produktiv	Kandiat*innen werden mit Hilfe von KI bewertet und überprüft	Kanada
Interview Suite	Videobewerbungen	Viasto GmbH	https://www.viasto.com/produkte/videobewerbungen/	Auswahlverfahren	HR-Abteilungen	Axa, ProSieben, Beiersdorf, Telekom, Union Investment, Stadt Essen, Bosch	produktiv	Video-Recruiting Lösungen. KI-basierte, stellenspezifische Interviewleitfäden, basiert auf personalpsychologischen Erkenntnissen. Standardisierte Interviews für mehr Fairness	Deutschland
Manatal Profile Enrichment	Datenanreicherung und Matching Plattform	Manatal	https://www.manatal.com/features/social-media-enrichment/	Screening, Auswahlverfahren	Recruiter*innen		produktiv	So viele Informationen wie möglich über die Kandidat*innen im Web finden, KI-basiertes Matching, Auto-Scheduling, CV-Parsing	Thailand
Metaview	Conversation Analytics	Metaview	https://www.metaview.ai/how/	Auswahlverfahren	HR-Abteilungen		experimentell	Metaview sammelt Daten aus Interviews, generiert Erkenntnisse und liefert Ratschläge, wie sich Interviewer*innen verbessern können.	UK

Produktname	Produkt	Firma	Link/Website	Recruitingschritte	Use Case	Kunden	Status	Versprechen	Land
MoBerries	Matching Plattform	MoBerries GmbH	https://www.moberries.com/	Sourcing, Screening	HR-Abteilungen	BASF, Tesla	produktiv	Für Bewerber*innen: mehr Privatheit durch individuelle Freigabe des Profils. Für Unternehmen: Relevanz der Ergebnisse, die Bewerber werden vorab von der Plattform gefiltert	Deutschland
Mya	Chatbot	Mya Systems, Inc.	https://www.mya.com/	Sourcing	HR-Abteilungen	L`Oreal, Addeco	produktiv	Mit dialogorientierter KI mehr Geschwindigkeit und Effizienz im Rekrutierungsprozess.	USA
Olivia Capture	Chatbot	Paradox Inc.	https://www.paradox.ai/capture	Kandidatenmanagement	HR-Abteilungen	Delta, Staples, Unilever	produktiv	Schnelleres Scheduling, FAQs, 24/7 Verfügbarkeit, einfacher Bewerbungsprozess für Kandidat*innen, Referenz-Programm	USA
Orange Tree Employmer	Screening von Bewerber*innen	Orange Tree Employment Screening	https://www.orangetreescreening.com/	Screenig	HR-Abteilungen	Diversicare Health, Rain for Rent	produktiv	Hintergrundchecks.	USA
Power Suite	Recruitment Process Outsourcing	Manpowergroup Talent Solutions RPO	https://www.manpowergroupsolutions.com/	Sourcing, Screening	HR-Abteilungen		produktiv	Maßgeschneiderte Recruiting-Lösungen	USA
Pre-screening & digitale Interviews	Assesment, Persönlichkeitsanalyse	Pymetrics	https://www.pymetrics.ai/solutions	Screening, Auswahlverfahren	HR-Abteilungen	JPMorgan, Unilever Plc, Accenture Plc, Tesla, LinkedIn	produktiv	Nutzt Daten und Neurowissenschadten für bessere Vorhersagen über Eignung und Potenzial, verspricht fairere Recruiting-Prozesse.	USA/UK/AUS/ Singapore/
Precire	Assesment, Persönlichkeitsanalyse	Precire	https://precire.com/technologie/?lang=en	Auswahlverfahren	HR-Abteilungen	Talanx AG, Fraport, Randstad n.V.	produktiv	Eine einzigartige Kombination aus Psychologie und künstlicher Intelligenz	Deutschland
Prescreen	Bewerbungsmanagement	Prescreen International GmbH	https://prescreen.io/	Sourcing, Screening	HR-Abteilungen	Greenpeace, Amnesty International, Birkenstock, Beiersdorf	produktiv	Schnelles Schalten von Anzeigen, Talentpool-Management	Österreich
Recruiter chatbots	Chatbot	ideal.	https://ideal.com/product/recruiting-chatbot/	Sourcing	HR-Abteilungen	Indigo, Burlington	produktiv	Rund um die Uhr mit Kandidat*innen kommunizieren	Kanada
Search!	Semantic Search	Carerix	https://carerix.com/de/product/search/		Zeitarbeit (Tool für Personaldienstleister, die Zeitarbeit anbieten), Recruiting für große Unternehmen	Adecco, Randstad, Manpower	produktiv	Viele Datenbanken schnell durchsuchen und die besten Kandidat*innen finden und effizient verwalten. Personalfluktuation verringern, effizienteres Verwalten des Personals	NL
Seedlink	Assesment, Persönlichkeitsanalyse	Seedlink	https://www.seedlinktech.com/	Auswahlverfahren	HR-Abteilungen und Recruiter*innen	L'Oreal Frankreich und Deutschland, Coca Cola, City of Rotterdam	produktiv	Sprachanalyse für schnellere und präzisere Personalentscheidungen. Gender-Assoziationen werden entfernt	Netherlands/China
Skillsurvey	Referenz-Checks Plattform	Skillsurvey	https://www.skillsurvey.com/source/	Screening	HR-Abteilungen und Recruiter*innen	Walmart, Adidas, Adecco, Tufts Medical Center	produktiv	Gesundheitsbackground-Checks, Referenzchecks, Background-Checks.	USA
SuccessFactors Human Experience Management (HXM)	Talent Management	SAP	https://www.sap.com/products/human-resources-hcm/hxm-suite.html	Sourcing, Kandidatenmanagement, Screening	HR-Abteilungen	Motaengil, Terex, Siegwerk	produktiv	Geeignete Kandidat*innen finden, Prozessoptimierung	Deutschland
Talent analytics and culture fit	Assesment, Persönlichkeitsanalyse	humantic	https://humantic.ai/	Auswahlverfahren	Recruiter*innen	Apple, Secu, Reinvent, paired sourcing	produktiv	KI mit emotionaler Intelligenz, hilft dabei die Organisationskapazität um 7-12% zu erhöhen	USA
textio	Recruitment Marketing	textio	https://textio.com/products/recruiting/	Sourcing	HR-Abteilungen	Novartis, Johnson & Johnson, die Weltbank, McDonald's, American Express, Dropbox, Spotify, Nestle, Credite Suisse, BP	produktiv	Hilft dabei die richtigen Worte für die Ausschreinung zu finden. Beseitigt Bias	USA

Produktname	Produkt	Firma	Link/Website	Recruitingschritte	Use Case	Kunden	Status	Versprechen	Land
Truffls	Teil-Anonymisiertes Matching/Bewerbung	truffls	https://truffls.de/	Sourcing	HR-Abteilungen	Vodafone, Paypal, idealo, Allianz	produktiv	Marktführer in Mobile Recruiting für Young Professionals	Deutschland
Unbias	LinkedIn Anonymiser für Recruiter	Jenny Brennan	https://unbias.io/	Sourcing	Browser-Erweiterung für Recruiter*innen, die sie LinkedIn nutzen		produktiv	Bias-Minimierung durch Anonymisierung: Browser-Erweiterung, die Gesichter und Namen im LinkedIn verdeckt, um unbewusstes Bias zu minimieren	UK
unbiasify	LinkedIn, Angelist, Twitter Anonymiser für Recruiter	Martin Hauck/Chrome extension	https://chrome.google.com/webstore/detail/unbiasify/affijhegklbkdinpepgphhlgphnhbenk	Sourcing	Recruiter*innen		produktiv	Namen und Gesichter von Bewerber*innen werden auf LinkedIn, Twitter und Angellist durch eine Browsererweiterung anonymisiert	Kanada
Veremark	Screening von Bewerber*innen	Veremark	https://www.veremark.com/	Screening	HR-Abteilungen	Schneider Electric, Kühne+Nagel, Octopus Deploy	produktiv	Hintergrundchecks.	UK
Video interviewing software & Pre-hire assessments	Assesment, Persönlichkeitsanalyse	HireVue	https://www.hirevue.com/	Auswahlverfahren	HR-Abteilungen	Unilever, Hilton Hotels	produktiv	Schnell und einfach passende Kandidat*inenn finden. Für Unternehmen, die schnell ihren Personal eskalieren müssen. Un-Bias, Meinungsfreiheit, Privacy und Daten-Recycling	USA
Wonderkind	Targeted job advertising	Wonderkind Global B.V.	https://wonderkind.com/	Sourcing	Recruiter*innen, HR-Abteilungen	Heineken, KFC, Booking.com, TATA Steel	produktiv	Ansprache von passiven Kanidat*innen mittels KI, die nicht aktiv auf Jobsuche sind	Niederlande
XOR	Suite mit Chatbot, Screening, Bewerbungs-management	XOR Inc.	https://www.xor.ai	Sourcing, Screening	HR-Abteilungen	IKEA, McDonald's, Exxon, Manpower	produktiv	Vollautomatisiertes Screening, Employee Referrals und Talentpool-Engagement via Chatbot, Analytics sammeln für zukünftige Strategien	USA

Impressum

Lorena Jaume-Palasi
The Ethical Tech Society

Elisa Lindinger, Julia Kloiber
Superrr Lab

Dieses Dokument wurde im Auftrag der Sachverständigenkommission für den Dritten Gleichstellungsbericht der Bundesregierung erstellt. Der Inhalt des Dokuments wird vollständig von den Autorinnen und Autoren verantwortet und spiegelt nicht notwendigerweise die Position der Sachverständigenkommission wider.

Herausgeberin
Geschäftsstelle Dritter Gleichstellungsbericht der Bundesregierung
Institut für Sozialarbeit und Sozialpädagogik e.V.
Lahnstraße 19, 12055 Berlin
www.dritter-gleichstellungsbericht.de

Stand: Juli 2020
Erscheinungsjahr: 2020

Zitierhinweis
Jaume-Palasi, Lorena/Lindinger, Elisa/Kloiber, Julia (2020): AI Powered Recruiting? Wie der Einsatz von algorithmischen Assistenzsystemen die Gleichstellung auf dem Arbeitsmarkt beeinflusst. Expertise für den Dritten Gleichstellungsbericht der Bundesregierung, www.dritter-gleichstellungsbericht.de.

Umschlagsgestaltung
WARENFORM, Berlin | www.warenform.de

DRITTER GLEICHSTELLUNGSBERICHT

www.dritter-gleichstellungsbericht.de