



Künstliche Intelligenz in der Personalauswahl

Projektleitung AMS:
Claudia Felix

Projektleitung Universität Graz:
Stefan Thalmann



Wien, Juli 2021

Impressum

Arbeitsmarktservice

Dienstleistungsunternehmen des öffentlichen Rechts

Treustraße 35-43

1200 Wien

Telefon: +43 50 904 199

Künstliche Intelligenz in der Personalauswahl

Stefan Thalmann, Christine Malin, Cordula Kupfer, Jürgen Fleiß, Martin Griesbacher, Bettina Kubicek

Juli 2021

Inhaltsverzeichnis

1	Motivation und Überblick	4
2	KI in der Personalauswahl	6
2.1	Künstliche Intelligenz	6
2.2	Traditioneller Personalauswahlprozess	7
2.3	KI-Unterstützung im Personalauswahlprozess	8
3	Perspektive Personalist*innen	11
3.1	Personalist*innenperspektive in der Literatur	11
3.2	Interviewstudie mit Personalist*innen in Österreich	11
3.2.1	Methodisches Vorgehen	11
3.2.2	Status Quo und Barrieren	13
3.2.3	Einsatzgebiete von KI in der Personalauswahl	16
3.2.4	Trends	18
4	Perspektive Bewerber*innen	23
4.1	Bewerber*innenperspektive in der Literatur	23
4.2	Allgemeine Perspektive von Bewerber*innen in Österreich	25
4.3	Die Perspektive von AMS Klient*innen	25
5	Dashboarddesign	33
5.1	Erstellung eines Dashboards	33
5.1.1	Definition von Anforderungen	33
5.1.2	Der Dashboard Prototyp	35
5.1.3	Konzeptevaluation	36
5.2	Experimentelle Untersuchung des Dashboards	39
5.2.1	Methodisches Vorgehen	39
5.2.2	Ergebnisse	41
5.2.3	Explorative Analysen	49
6	Anforderungskatalog und Empfehlungen	50
7	Ausblick	52
8	Anhang	57

Executive Summary

Künstliche Intelligenz (KI) wird zunehmend als eines der wesentlichsten Elemente der digitalen Transformation der Arbeitswelt gesehen. Der vorliegende Forschungsbericht verfolgt das Ziel, die aktuelle Nutzung und die Anforderungen für den Einsatz von KI in der Personalauswahl in Österreich zu erheben. Ein Hauptgrund für den verstärkten Bedarf an digitalen Lösungen in der Personalauswahl sind steigende Bewerbungszahlen durch mobile Anwendungen und niedrigere Hemmschwellen. Der damit einhergehende Automatisierungsbedarf kann potentiell auch durch KI-Lösungen abgedeckt werden. Das Thema wird zwar in der Literatur diskutiert, aber belastbare Ergebnisse aus der Perspektive von Personalist*innen und Bewerber*innen allgemein und insbesondere für Österreich fehlen bisher.

Im **2. Kapitel** wird der Begriff KI abgegrenzt und es werden mögliche Einsatzgebiete im Kontext traditioneller Personalauswahlprozesse diskutiert. Besonders Potential verspricht dabei die Informationssammlung und -aufbereitung, die Vorselektion von Kandidat*innen und die Ansprache und Informationsbereitstellung für Bewerber*innen.

Im **3. Kapitel** wird aus Sicht der Personalist*innen untersucht, wie KI aktuell in Österreich für die Personalauswahl eingesetzt wird, welche Auswirkungen der Einsatz (potentiell) auf die Arbeit von Personalist*innen hat und welche zukünftigen Einsatzgebiete diese für die KI erwarten. Befragt wurden 25 Personalverantwortliche in 21 österreichischen Unternehmen in unterschiedlichen Betriebsgrößen und Branchen. Insgesamt sind KI-gestützte Systeme derzeit in Österreich nur vereinzelt im Einsatz. Als Gründe wurden vier Barrieren identifiziert: 1.) ein unklares bzw. ungünstiges Nutzen-Aufwand-Verhältnis, 2.) die mangelnde Reife der Technologie, 3.) die Angst vor dem unfreiwilligen Verlust von Bewerber*innen und 4.) die Angst von Personalist*innen interessante Aufgaben an die Technologie zu verlieren. Insgesamt ist den Personalverantwortlichen wichtig, dass Menschen die Kontrolle über den Prozess haben und dass durch die Automatisierung von Routineaufgaben mehr Zeit für wichtige und kreative Aufgaben bleibt.

Im **4. Kapitel** wird die Perspektive der Bewerber*innen beim Einsatz von KI im Personalauswahlprozess untersucht. Im Rahmen einer repräsentativen Fragebogenstudie wird insbesondere die Perspektive diskriminierter Gruppen betrachtet. An der Befragung nahmen 1335 beim AMS gemeldete Personen teil, welche zu ihrem möglichen Vertrauen in den die Entscheidungen von menschlichen Personalist*innen und verschiedenen KI-Varianten befragt wurden. Während insgesamt Entscheidungen von Menschen eher vertraut wird, konnte eine steigende Fairnesswahrnehmung von KI mit Erklärungskomponente gefunden werden, insbesondere bei Befragten, die vermehrt von Alltagsdiskriminierung betroffen sind.

Im **5. Kapitel** wird eine Designstudie beschrieben und die Anforderungen von KI in der Personalauswahl in Bezug auf Designstandards dargestellt. Anforderungen wie Interaktivität, Integration von Schlüsselindikatoren zur Berufseignung, Möglichkeiten menschlicher Intervention, Informationsbereitstellung und den Abgleich zwischen Bewerber*innen- und Anforderungsprofil werden diskutiert. Mit einer experimentellen Studie wird der Einfluss von Dashboards auf das Entscheidungsverhalten von Personalist*innen untersucht.

Im **6. Kapitel** resümieren wir die Ergebnisse der Untersuchungen in Form eines Anforderungskatalogs. Es werden fünf zentrale Empfehlungen für den Einsatz von KI in der Personalauswahl gegeben:

1. KI sollte Personalist*innen vordergründig in ihren Aufgaben unterstützen und nicht automatisieren.
2. KI sollte Bewerber*innen einen Mehrwert bieten und nicht abschrecken.
3. KI sollte nur mit einer Erklärungskomponente eingesetzt werden.
4. KI sollte gezielt für mehr Vielfalt und gegen Diskriminierung eingesetzt werden.
5. KI braucht Rechtssicherheit für den Einsatz.

Der Bericht schließt mit einem kompakten **Ausblick** zu dem aktuellen und zu erwartenden Einsatz von KI in der Personalauswahl.

1 Motivation und Überblick

Die Digitalisierung hatte und hat weiterhin einen sehr großen Effekt auf Konsument*innen, Unternehmen und die Gesellschaft insgesamt (Vial, 2019). Diese Effekte wurden durch die COVID-19 Pandemie und die vermehrte virtuelle Zusammenarbeit noch verstärkt (Waizenegger u.a., 2020). In diesem Zusammenhang hat sich auch der Einsatz von Künstlicher Intelligenz (KI) weiter verbreitet. Durch den Einsatz von KI konnten Unternehmen ihren Nutzer*innen viele neue oder verbesserte Services mit einer neuartigen Interaktionserfahrung bieten (Alfaro u.a., 2019). Digitale und mobile Personalauswahl ist eine Antwort aus dem Personalwesen auf diese Trends. Durch bessere Verfügbarkeit und einfachere Prozesse steigen die Bewerbungszahlen rasant an (Black und Esch, 2020). So langten bei internationalen Unternehmen durchschnittlich 250 Bewerbungen ein, von denen ca. 80 Prozent mobil oder per Social Media erstellt werden (Black und Esch, 2021). Für Personalabteilungen bedeutet dies einen erhöhten Aufwand und damit steigt auch die Notwendigkeit von IT-Unterstützung im Bewerbungsprozess.

KI, insbesondere in Kombination mit Chatbots, nimmt diese Trends auf und verspricht einerseits eine Unterstützung für Personalist*innen und andererseits eine verbesserte Interaktion für Bewerber*innen. Doch wie wird KI aktuell in Österreich für die Personalauswahl eingesetzt bzw. welche Anforderungen für diesen Einsatz gibt es in der Praxis? Mit Hilfe der bestehenden Literatur können diese Fragen aktuell nicht beantwortet werden. Aus internationaler Perspektive gibt es einige Arbeiten, die auf einzelne Anwendungsfelder fokussieren, aber keine umfassende Untersuchung vornehmen. Daher soll aufbauend auf einer Literaturanalyse die Frage mit Hilfe empirischer Studien beantwortet werden. Es ergeben sich dabei zwei zentrale Perspektiven: (1) die Sichtweise der Personalist*innen, die den Personalauswahlprozess für Organisationen durchführen und (2) die Sichtweise der Bewerber*innen, die den Prozess durchlaufen.

Abbildung 1 gibt einen Überblick über das Forschungsvorhaben und den Aufbau dieses Berichts. Einführend wird zunächst der Begriff KI abgegrenzt und definiert (Abschnitt 3). Aufbauend darauf wird auf Basis der bestehenden Literatur untersucht, für welche Bereiche bzw. Teilaufgaben der Personalauswahl KI eingesetzt werden kann. Anschließend werden die beiden Perspektiven (Personalist*innen und Bewerber*innen) getrennt untersucht. Für die Untersuchung der Perspektive der Personalist*innen wurde eine Interviewstudie mit 25 Teilnehmer*innen aus Österreich durchgeführt. Dabei wurde insbesondere der aktuelle Einsatz von KI im Personalauswahlprozess, Anforderungen für den Einsatz, und Auswirkungen des Einsatzes untersucht. Dabei hat sich herausgestellt, dass die Entscheidungsunterstützung für Personalist*innen eine vielversprechende Anwendung mit noch zu lösenden Herausforderungen ist. Aus diesem Grund wurde in einer Designstudie untersucht, wie große Datenmengen dargestellt und aggregiert werden können, um den Personalist*innen eine rechtssichere und angemessene Analyse zu ermöglichen. Zu diesem Zweck wurde ein prototypisches Dashboard erstellt und in einem Experiment mit 60 Teilnehmer*innen evaluiert.

Aus Sicht der Bewerber*innen wurde insbesondere die Frage der Akzeptanz näher betrachtet. Denn nur wenn Nutzer*innen einer Technologie vertrauen, wird sie erfolgreich sein. Im Rahmen einer Vignettenstudie wurde daher der Einsatz von Chatbots in einem Bewerbungsprozess mit und ohne Erklärungskomponente untersucht. Ein vielfach vorgebrachter Vorbehalt gegenüber KI ist, dass KI diskriminierend sei. Die Literatur zeigt jedoch, dass KI-Systeme für die Personalauswahl derart gestaltet sein können, dass einer Diskriminierung entgegengewirkt wird. Im Rahmen einer zweiten Vignettenstudie mit Arbeitslosen in Österreich wurde hierzu die Bewerber*innenperspektive auf KI in der Personalauswahl untersucht.

Abschließend werden die Erkenntnisse aus der Literaturanalyse sowie aus den empirischen Arbeiten zur Personalist*innenperspektive und zur Bewerber*innenperspektive zusammengeführt. Als Ergebnis werden Anforderungen für den Einsatz von KI in der Personalauswahl formuliert und Empfehlungen für den Einsatz gegeben.



Abbildung 1.1: Struktur des Berichtes

Die vorliegende Studie wurde an der Universität Graz im Auftrag des AMS Österreich unter der Leitung von Univ.-Prof. Dr. Stephan Thalmann (Business Analytics and Data Science Center, BANDAS-Center) in Kooperation mit Univ.-Prof. Dr. Bettina Kubicek (Institut für Psychologie, Arbeitsbereich Arbeits- und Organisationspsychologie) durchgeführt. Am Projekt haben Christine Malin, Jürgen Fleiß und Robert Andréé

(BANDAS-Center), Cordula Kupfer und Rita Prassl (Psychologie) sowie Martin Griesbacher (Forschungsnetzwerk Human Factor in Digital Transformation) mitgearbeitet. Besonderer Dank gilt dem AMS Österreich für die erfolgreiche Kooperation, den Interviewpartner*innen und allen Teilnehmer*innen an der Online-Befragung.

2 KI in der Personalauswahl

Als Einführung in das Thema KI in der Personalauswahl folgt (a) eine Definition von KI, (b) eine Darstellung des traditionellen Personalauswahlprozesses und (c) eine Beschreibung potentieller KI-Unterstützung im Personalauswahlprozess.

2.1 Künstliche Intelligenz

Der Begriff "Künstliche Intelligenz" und KI als Disziplin wurden 1956 auf der Dartmouth Conference vom amerikanischen Informatiker John McCarthy geprägt (Moor, 2006). John McCarthy definiert KI wie folgt: "The science and engineering of making intelligent machines, especially intelligent computer programs" (McCarthy, 2007). Aufgrund der vielfältigen Teilgebiete von KI und der diversen Ziele und Perspektiven, die von den Wissenschaftler*innen verfolgt werden, ist es eine Herausforderung, eine allgemeingültige Definition festzulegen. Grundsätzlich soll KI menschliche Eigenschaften und Fähigkeiten wie Problemlösen, Erklären, Lernen, Sprachverständnis und flexibles Reagieren simulieren (Russel und Norvig, 2016).

Die folgende Definition des Merriam-Webster Dictionary wird für den Zweck dieses Forschungsberichts als am geeignetsten angesehen: Künstliche Intelligenz ist "1) ein Zweig der Informatik, der sich mit der Simulation von intelligentem Verhalten in Computern beschäftigt, 2) die Fähigkeit einer Maschine, intelligentes menschliches Verhalten zu imitieren" (Dictionary, 2021).

In den Anfängen von KI wurden Systeme mit vorprogrammierten Regeln inklusive passender Reaktionen entworfen, die durch Schlüsselbegriffe und Muster ausgelöst werden (Weizenbaum, 1966). In den letzten Jahrzehnten wurden allerdings enorme Fortschritte im Bereich von KI-Technologien erzielt, der wohl bedeutendste ist das maschinelle Lernen (ML). Es ist die Fähigkeit eines Systems, selbständig zu lernen, sich zu verbessern und Aufgaben auszuführen. Ein System, das ML verwendet, ist nicht auf ein bestimmtes Ergebnis programmiert. Es lernt Regeln und Modelle aus

großen Mengen an Beispiel-Daten (Raub, 2018) und nutzt strukturiertes Feedback, um Probleme selbständig zu lösen, wie z. B. das Erkennen von Gesichtern oder Tieren. Es gibt verschiedene Kategorien von ML, die jedoch am weitesten verbreiteten ML-Methoden sind überwachte Lernmethoden. Das bedeutet, dass ein System Beispiele mit korrekten Ergebnissen erhält, z.B. verschiedene Bilder von Tieren mit den entsprechenden Bezeichnungen wie Hund, Katze etc. Nachdem es mit einer großen Menge an Beispielen trainiert wurde, kann ein solches System mit hoher Treffergenauigkeit Ergebnisse für neue Daten liefern (Brynjolfsson und McAfee, 2017). Eine andere Kategorie sind unüberwachte Lernmethoden, bei denen dem System das Ergebnis nicht vorgegeben wird und das System eine Datenmenge strukturieren soll. Das Ziel ist, dass die Maschine anhand der Daten selbst lernt und Strukturen entwickelt, oft mit Methoden, die vom menschlichen Gehirn inspiriert sind (Cherkassky und Mulier, 2007).

KI kann unter anderem zur Mustererkennung, Texterkennung, Bilderkennung, Sprach- und Emotionserkennung in vielen verschiedenen Kontexten eingesetzt werden (Aggarwal und Zhai, 2012; Ripley, 2007; El Ayadi, Kamel und Karray, 2011; Russel und Norvig, 2016; DiRomualdo, El-Khoury und Girimonte, 2018; Upadhyay und Khandelwal, 2018). Die potentiellen Anwendungsbereiche in der Arbeitswelt reichen dabei vom Marketing über Logistik bis hin zu Frage-Antwort Interaktionen in Service Centern (Rathi, 2018).

Ein weiteres vielversprechendes Einsatzgebiet stellt die Personalauswahl dar. In der Literatur werden dabei vor allem zwei große Einsatzbereiche hervorgehoben: (1) die Sichtung von Bewerbungsunterlagen durch Dashboards, und (2) Bewerbungsgespräche durch Chatbots. Im ersten Fall können Daten in Dokumenten automatisiert erfasst oder aus Datenbanken (die Daten wurden zuvor per Formular erfasst) ausgelesen werden. Die KI schlägt dann Bewerber*innen für ein Gespräch vor und vergibt möglicherweise einen Matching-Score zur Visualisierung der Passgenauigkeit für die Stelle. Im zweiten Fall werden die Bewerber*innen gar nicht aufgefordert, ihre Daten strukturiert einzugeben, vielmehr führt ein Chatbot ein Gespräch und erfragt alle relevanten Informationen. Am

Ende trifft die KI eine Entscheidung bzw. liefert einen Entscheidungsvorschlag. Aus Kostengründen wurde bisher vorwiegend das statische Verfahren mit Hilfe von Bewerbungsunterlagen verwendet, da Gespräche für alle Kandidat*innen einen zu großen Aufwand bedeuteten. Mit Hilfe von Chatbots lässt sich dies jedoch nun viel breiter aufstellen und es kann bereits in der Vorauswahl individueller auf die Kandidat*innen eingegangen werden. Grundsätzlich ist KI in der Lage, nahezu jeden Personalauswahlschritt zu unterstützen (Sekhri und Cheema, 2019), wie im Folgenden dargestellt werden soll.

2.2 Traditioneller Personalauswahlprozess

Im Rahmen dieses Forschungsberichts bezieht sich der Begriff "traditioneller Personalauswahlprozess" auf Prozesse und Ansätze, die vor Mitte der 1990er Jahre eingesetzt wurden. Bis zu diesem Zeitpunkt erfolgte der Personalauswahlprozess vorwiegend mithilfe analoger Medien (Black und Esch, 2020). Die traditionelle Personalauswahl wird häufig als "Post & Pray" - Strategie beschrieben (Ullah und Witt, 2018). Unternehmen veröffentlichten ihre Stellenanzeigen in analogen Medien wie Printmedien und Jobbörsen ("Post"). "Pray" deutet darauf hin, dass Unternehmen nahezu keinen Einfluss auf den weiteren Prozess nach der Veröffentlichung der Stellenanzeige hatten und auf Bewerber*innen hofften. Erst später wurde auf den weiteren Verlauf durch gezieltere Zielgruppenansprache Einfluss genommen.

Eine Herausforderung der traditionellen Personalauswahl mit ihrer "Post & Pray" - Strategie sind die begrenzte Reichweite und der begrenzte Informationsgehalt einer analogen Personalsuche (Black und Esch, 2020). Unternehmen hatten drei Optionen der Personalsuche: Stellenanzeigen in Printmedien, Vermittlungen durch bestehende Mitarbeiter*innen und durch

Personalvermittlungsagenturen. Die Reichweite einer Stellenanzeige konnte durch zusätzliche Anzeigen in Printmedien, der Informationsgehalt durch größere, umfangreichere Anzeigen erhöht werden. Beides bedeutete jedoch erhebliche Mehrkosten für das Unternehmen. Hohen Informationsgehalt hatten mündliche Vermittlungen durch bestehende Mitarbeiter*innen, hier beschränkte sich die Reichweite jedoch auf ihr unmittelbares Umfeld. Hohe Reichweite und hoher Informationsgehalt konnten durch Personalvermittlungsagenturen erreicht werden, die allerdings kostspielig waren (Black und Esch, 2020).

Erst seit Ende der 1990er Jahre, mit dem Einsetzen digitaler Möglichkeiten in der Personalauswahl, können Unternehmen ohne hohe Kosten eine hohe Reichweite und hohen Informationsgehalt ihrer Stellenanzeigen erreichen (Black und Esch, 2020). Stellenanzeigen können nun relativ günstig im Internet veröffentlicht werden, und Job-Plattformen und soziale Netzwerke ermöglichen die Suche nach geeigneten Kandidat*innen.

Der traditionelle Personalauswahlprozess kann üblicherweise durch fünf Schritte charakterisiert werden (Ullah und Witt, 2018), die in Abbildung 2.1 dargestellt sind. Die einzelnen Schritte werden nachfolgend näher beschrieben.

Stellenausschreibung

Um eine Stellenausschreibung vorzubereiten, müssen zunächst detaillierte Stellenanforderungen erstellt werden, die auf der Personalbedarfsplanung und einer Bedarfsanalyse basieren. Die Stellenanforderung wird gemeinsam von der jeweiligen Fachabteilung und der Personalabteilung definiert. Dabei werden erfolgsrelevante Anforderungskriterien - einerseits fachliche Anforderungskriterien (Hard Skills), andererseits Persönlichkeitsmerkmale (Soft Skills) - identifiziert. Für



Abbildung 2.1: Personalauswahlprozess nach Ullah und Witt (Ullah und Witt, 2018)

die Veröffentlichung wird die Stellenbeschreibung zudem in einer zielgruppengerechten Sprache verfasst (Ullah und Witt, 2018; Schulz, 2014).

Zielgruppenansprache

Bei der Zielgruppenansprache werden die Kanäle, über die die Stellenanzeige veröffentlicht werden soll, festgelegt. Analoge Personalauswahl beschränkte sich hier hauptsächlich auf Printmedien wie Zeitungen, und auf Jobbörsen bei Arbeitsagenturen. Mittlerweile häufig genutzte Kanäle sind die unternehmenseigene Internetseite, Personalvermittlungsagenturen, Empfehlungsprogramme, Jobbörsen sowie soziale Netzwerkseiten (z.B. LinkedIn) (Armstrong und Taylor, 2014). Es können auch unternehmensinterne Stellenvermittlungen in Betracht gezogen oder gezielt ehemalige Mitarbeiter*innen angesprochen werden. Grundsätzlich sollte jener Kanal genutzt werden, den die Zielgruppe präferiert.

Sichtung

In diesem Schritt werden die eingegangenen Bewerbungsunterlagen gesichtet. Dabei werden die Informationen aus den Bewerbungsunterlagen mit den formulierten Stellenanforderungen von Personalist*innen abgeglichen. Das Ziel ist die Erstellung einer Liste mit nach Eignung gereihten Bewerber*innen (Armstrong und Taylor, 2014). Anhand dieser Liste können, oft in Zusammenarbeit mit dem jeweiligen Fachbereich, Bewerber*innen für persönliche Bewerbungsgespräche ausgewählt werden (Ullah und Witt, 2018).

Auswahl

"Das Ziel der Auswahl ist es, die Eignung der Kandidat*innen zu beurteilen, indem man das Ausmaß vorhersagt, in dem sie in der Lage sein werden, eine Rolle erfolgreich auszuführen" (Armstrong und Taylor, 2014). Die Bewerber*innen werden hinsichtlich der Anforderungen an Hard- und Soft-Skills verglichen und weitere relevante Beurteilungskriterien werden berücksichtigt. Die drei traditionellen Auswahlmethoden sind das Bewerbungsgespräch, Referenzen und

Bewerbungsformulare. Darüber hinaus können zusätzlich Eignungstests und Assessment Center herangezogen werden. Auch hier ist meist der jeweilige Fachbereich involviert (Ullah und Witt, 2018; Armstrong und Taylor, 2014).

Vertrag

Im letzten Schritt finden die Vertragsverhandlungen sowie das Onboarding - also die administrative Aufnahme und Integration in die Strukturen und Abläufe des Unternehmens - der ausgewählten Bewerber*innen statt. Darüber hinaus beinhaltet diese letzte Phase auch die Probezeit der neuen Mitarbeiter*innen (Ullah und Witt, 2018).

2.3 KI-Unterstützung im Personalauswahlprozess

Der traditionelle Personalauswahlprozess bietet vielfältige Einsatzmöglichkeiten für KI-basierte Lösungen. In der Literatur werden vor allem die Zielgruppenansprache, die Sichtung der Bewerbungsunterlagen und die Auswahl von Bewerber*innen als Bereiche diskutiert, in denen KI-Unterstützung einen Mehrwert bieten kann. Der Einsatz von KI-Systemen hat das Potential, die Personalauswahl nachhaltig zu verändern (Upadhyay und Khandelwal, 2018). Insbesondere repetitive Tätigkeiten können nun von Chatbots oder Dashboards übernommen werden. Beispielsweise werden die Zielgruppenansprache und die Sichtung der Bewerbungsunterlagen an Bewerbungsmanagement-Systeme delegiert und somit vollautomatisch ausgeführt (Upadhyay und Khandelwal, 2018; Verhoeven, 2020).

Zielgruppenansprache

Die Zielgruppenansprache beschränkt sich gegenwärtig nicht mehr ausschließlich auf die Auswahl geeigneter Kanäle. Mit dem Fortschreiten der Digitalisierung und dem Aufkommen von sozialen Netzwerken sowie großen Job-Plattformen ist es nun möglich, spezifische Zielgruppen zu identifizieren. Dabei kann zwischen di-

rekter und breit angelegter Zielgruppenansprache unterschieden werden. Bei der direkten Zielgruppenansprache wird zunächst ein Stellenprofil erstellt, anhand dessen Datenbanken nach geeigneten Kandidat*innen durchsucht werden. Im Rahmen der breit angelegten Zielgruppenansprache werden Nutzer*innenprofile aus Online-Kanälen identifiziert und erst im weiteren Verlauf ein Abgleich mit einem Stellenprofil durchgeführt. Bei beiden Varianten kommen KI-Systeme zum Einsatz, die entweder unterstützend oder völlig autonom genutzt werden können (Verhoeven, 2020).

Eine direkte Zielgruppenansprache erfolgt im Rahmen der aktiven Suche von Kandidat*innen (Active Sourcing), dem Gegenstück zur bereits erwähnten "Post & Pray" - Strategie. Personalist*innen definieren zunächst die Stellenbezeichnung sowie Fähigkeiten und Qualifikationen. Darauf basierend durchsucht ein unterstützendes KI-System die Datenbanken und schlägt geeignete Kandidat*innen vor. Diese können nun gezielt von den Personalist*innen kontaktiert werden. Die KI lernt aus der tatsächlichen Auswahl der Personalist*innen und verbessert so laufend den Algorithmus für die Suche. Ein völlig autonomes KI-System würde zudem noch die eigentliche Vorauswahl aus der Datenbank sowie die anschließende Kommunikation mit den Kandidat*innen durch den Einsatz von Chatbots übernehmen (Verlinden, 2019). Die Leistung eines solchen autonomen KI-Systems wird im Vergleich zu menschlichen Experten*innen nur geringfügig schlechter eingestuft. Jedoch benötigen KI-Systeme für eine aktive Suche lediglich 3,2 Sekunden, während dieser Arbeitsschritt bei menschlichen Expert*innen vier bis 25 Stunden dauern kann (Eubanks, 2018).

Eine breit angelegte Zielgruppenansprache erfolgt durch neuere KI-Anwendungen, die über Daten aus LinkedIn, XING, Instagram, Twitter, Facebook, Job-Plattformen und internen Datenbanken einen riesigen Kandidat*innen-Pool generieren. Basierend auf diesen großen Datenmengen werden Nutzer*innenprofile erstellt (Black und Esch, 2020). Das System macht hierbei jedoch keinen Unterschied, ob eine Person aktiv nach einem Job sucht oder nicht. In einem nächsten Schritt gleicht die KI passende Kandidat*innen mit dem erwünschten Stellenprofil ab. Basierend auf den erstellten Nutzer*innenprofilen erkennt das KI System

auch, über welchen Kanal bestimmte Kandidaten*innengruppen am besten angesprochen werden sollten. Dabei gibt es verschiedene Möglichkeiten wie Anzeigen, Banner, E-Mails oder Push-Nachrichten, um die Erfolgsaussichten zu erhöhen. Einige KI Anwendungen personalisieren sogar den Wortlaut der Anzeigen und Nachrichten (Black und Esch, 2020).

Auch Chatbots werden zunehmend für die Zielgruppenansprache eingesetzt. In der Praxis kann dies beispielsweise bedeuten, dass der Chatbot eine Pushnachricht sendet, wenn potentielle Kandidat*innen eine Stellenanzeige auf der Homepage des Unternehmens lesen. Der Chatbot übernimmt den Erstkontakt und das Einholen von Bewerbungsunterlagen und Informationen über die potentiellen Kandidat*innen. Ein Chatbot kann auch Fragen zu den Bewerbungsunterlagen stellen. KI-basierte Chatbots können daher als erste interaktive Kontaktstelle für die Personalauswahl gesehen werden (Schikora, Galster und Högerl, 2020; Upadhyay und Khandelwal, 2018).

Sichtung

Neben der Zielgruppenansprache hat sich auch die Sichtung der Bewerbungsunterlagen in den letzten Jahren stark verändert. Eine Sichtung und Vorauswahl erfolgt oft bereits im Rahmen der Zielgruppenansprache, da viele KI-Systeme nur Kandidat*innen einer bestimmten Zielgruppe ansprechen. Diese Kandidat*innen sind somit bereits zu einem gewissen Grad aus einem Pool von Personen vorselektiert. Bei der Vorauswahl spielen Bewerbungsmanagement-Systeme eine große Rolle, die dann die eingehenden Bewerbungen mithilfe von KI-Matching-Tools sortieren (Ullah und Witt, 2018). Solche Tools können Chatbots sein, die Lebensläufe und andere Bewerbungsunterlagen analysieren. Die in den Dokumenten verwendeten Satzstrukturen und Wörter dienen als Grundlage, um den gesamten Inhalt der Bewerbung zu bewerten und mit dem Anforderungsprofil zu vergleichen. Darüber hinaus können Chatbots Tests durchführen und auswerten, die auch in die Bewertung der Kandidat*innen durch das KI Matching-Tool einfließen (Verhoeven, 2020; Black und Esch, 2020). Ein möglicher KI- und Chatbot-Matching Prozess wird im Folgenden dargestellt (Schikora, Galster und Högerl, 2020; Ullah und Witt, 2018): 1. Ein Chatbot analysiert die gesammelten

Informationen aus Bewerbungsunterlagen und Social Media-Profilen. 2. Die KI versucht eine Übereinstimmung zwischen den Kandidat*innen und dem Anforderungsprofil zu finden. 3. Bei fehlenden Informationen tritt der Chatbot erneut mit den Kandidat*innen in Kontakt. 4. Wenn alle Informationen vorhanden sind, können die Kandidat*innen nun nach einem Matching-Score gereiht werden. Der Chatbot arrangiert dann weitere Bewerbungsgespräche mit Video oder übermittelt Personen mit einem zu niedrigen Score eine Absage. 5. Die KI analysiert die Sprache und Mikroausdrücke des Videointerviews und aktualisiert den Matching-Score auf Basis der neu generierten Informationen.

Wie dieser beispielhafte Sichtungungs-Prozess nahelegt, verfügen KI-Systeme über das Potential, menschliche Personalist*innen bereits in diesen Phasen vollständig zu ersetzen. Ein Grund, der für den Einsatz von KI spricht, ist dass Algorithmen im Vergleich zu Menschen potentiell gerechtere Urteile fällen. Menschliche Fehleinschätzungen, die durch Aussehen, Ähnlichkeitseffekte oder Vorurteile gegenüber Ethnie, Geschlecht oder Alter geprägt sind, können minimiert werden (Herrmann, 2016). Wie bei der Zielgruppenansprache sollen KI-Anwendungen die Arbeitsbelastung der Personalist*innen reduzieren und zusätzliche Funktionen bereitstellen. Die erste Flut an Kandidat*innen wird heutzutage mit Hilfe von KI-Matching-Tools gesichtet und zu einer kompakten Auswahl vorsortiert. Ein bekanntes Beispiel ist das Unternehmen Unilever, das ein Experiment in 68 Ländern mit insgesamt 250.000 Kandidat*innen durchführte. Der Input für das KI-Matching System war das LinkedIn-Profil der Kandidat*innen, mehrere in Chatbots integrierte neurowissenschaftliche Spiele und ein Videointerview. Nachdem die besten Kandidat*innen mithilfe eines Algorithmus vorselektiert wurden, wurden Personalist*innen nur noch für die endgültige Auswahl involviert (Feloni, 2017).

Auswahl

Im traditionellen Personalauswahlprozess werden im Rahmen der Auswahlphase Bewerbungsgespräche sowie diverse Tests mit Kandidat*innen durchgeführt. Durch den Einsatz von KI bzw. von Chatbots ist kein

deutlicher Unterschied zwischen der Sichtung und dem Auswahlprozess erkennbar. In der Regel werden alle verfügbaren Daten bereits im Sichtungs-Prozess ausgewertet und analysiert, wodurch diese Phase entbehrlich wird. Je nach eingesetztem Bewerbungsmanagement-Tool werden bereits Assessments und Bewerbungsgespräche im Vorfeld der eigentlichen Auswahl durchgeführt und teilweise in den Sichtungs-Prozess einbezogen. Mittlerweile gibt es auch Tools, die Persönlichkeitsprofile auf Basis des Klick- und Like-Verhaltens auf Social-Media-Plattformen erstellen. Infolgedessen könnte das Führen von Bewerbungsgesprächen mit Kandidat*innen ersetzt werden (Ullah und Witt, 2018; crystalknows.com, 2020). Auch das Beispiel von Unilever zeigt, dass wesentliche Schritte der traditionellen Auswahl in den Sichtungungs-Prozess integriert sind, wodurch ein fundiertes Matching geschaffen wird. In welchem Umfang und in welcher Phase Personalist*innen involviert sind, kann variieren. Bei Unilever wurden beispielsweise die vielversprechendsten Kandidat*innen nach dem Sichtungungs-Prozess eingeladen, einen Tag mit den Personalist*innen an ihrem möglichen neuen Arbeitsplatz zu verbringen (Feloni, 2017). Das zeigt, dass durch freiwerdende Ressourcen der Personalist*innen ein persönlicheres Kennenlernen der Kandidat*innen im Auswahlprozess möglich wird. Aktuell gehören Bewerbungsgespräche, Tests und Assessment Center jedoch noch zum beruflichen Alltag der Personalist*innen. In der Praxis verfügt laut Ullah und Witt (Ullah und Witt, 2018) kaum ein Bewerbungsmanagement-Tool über den notwendigen Reifegrad, um effektiv im Personalauswahlprozess eingesetzt zu werden.

Die genannten Personalauswahlsschritte lassen sich üblicherweise gut strukturieren und typischerweise sind ausreichend Daten von historischen Auswahlentscheidungen vorhanden, um KI einzusetzen. Es stellt sich jedoch auch die Frage, inwieweit die Bewerber*innen und die Personalist*innen Vertrauen in ein solches System haben und welche Anforderungen beide Interessensgruppen an die KI stellen. Sowohl Personalist*innen als auch Bewerber*innen sind gleichermaßen von KI in der Personalauswahl betroffen. Im Rahmen des Projekts wurde versucht, beide Sichtweisen einzufangen. Im nächsten Kapitel werden zunächst die Einstellungen der Personalist*innen bzw. der Unternehmen, für die sie arbeiten, näher erläutert.

3 Perspektive Personalist*innen

Die Perspektive der Personalist*innen zum Einsatz von KI-Systemen in der Personalauswahl wird (a) durch Erkenntnisse bestehender wissenschaftlicher Studien und (b) durch eine Interviewstudie mit Personalverantwortlichen in österreichischen Unternehmen näher beleuchtet.

3.1 Personalist*innenperspektive in der Literatur

Die Sicht der Personalist*innen zum Thema KI wurde in bisherigen Studien in der Personalauswahl noch wenig berücksichtigt. Beispielsweise fand eine Umfrage mit Personalist*innen, dass die Bereitschaft für die Einführung von KI in der Personalauswahl steigt, wenn hohe Kosteneffizienz, Mehrwert, Wettbewerbsdruck, Ressourcen und Fähigkeiten in der Personalabteilung, und Unterstützung durch das Top-Management sowie KI-Anbieter gegeben sind. Bedenken bezüglich Sicherheit und Datenschutz hemmen unter anderem den Einsatz von KI in der Personalauswahl (Pillai und Sivathanu, 2020). Eine Studie von Langer et al. (Langer, König und Busch, 2020) untersuchte die konkrete Gestaltung des Einsatzes einer KI im Personalauswahlprozess. Sie zeigten, dass der Zeitpunkt des Einsatzes von KI im Entscheidungsprozess Auswirkungen auf die Zufriedenheit mit der Entscheidung sowie die wahrgenommene Selbstwirksamkeit durch die Personalist*innen hat. Wird die Vorselektion der KI erst nach der eigenständigen Bearbeitung durch Personalist*innen für einen Abgleich eingeblendet, und nicht bereits von Anfang an vorgegeben, sind Personalist*innen mit der Entscheidung zufriedener und erleben eine höhere Selbstwirksamkeit. Eine Umfrage in Deutschland zum Einsatz von Chatbots in der Personalauswahl verdeutlichte, dass Personalist*innen Chatbots vorwiegend positiv wahrnehmen und geringe Bedenken hinsichtlich des Verlustes ihrer Aufgaben sowie des Ersatzes durch die Technologie hegen (Eißer, Torrini und Böhm, 2020).

3.2 Interviewstudie mit Personalist*innen in Österreich

Um den Einsatz von KI in der Personalauswahl und die Einschätzung der Personalist*innen speziell in Österreich einzufangen, wurde im Rahmen dieses Projektes eine Interviewstudie durchgeführt. Dabei wurden der Status Quo sowie die Chancen und Risiken von KI in der Personalauswahl erhoben (siehe Interviewleitfaden im Anhang). Im Folgenden werden die methodische Durchführung und Ergebnisse der Interviewstudie erläutert.

3.2.1 Methodisches Vorgehen

Es wurden semi-strukturierte Interviews in deutscher Sprache mit österreichischen Personalverantwortlichen durchgeführt. Alle Interviewpartner*innen erhielten vorab einen Interviewleitfaden, welcher basierend auf einer vorangegangenen Literaturanalyse erstellt wurde. Der Leitfaden wurde in einem Pretest-Interview überprüft und angepasst. Um eine Diskussionsgrundlage zu schaffen, wurden den Interviewpartner*innen fiktive Fallbeispiele zu einem Chatbot sowie zu einem Dashboard präsentiert (siehe Erläuterung der Fallbeispiele in den Ergebnissen). Im Zeitraum von Juni 2020 bis März 2021 wurden 25 Personalverantwortliche von 21 österreichischen Unternehmen befragt. Alle Interviewpartner*innen mussten in ihrer Funktion mit der Personalauswahl im Unternehmen betraut sein. Die Mehrheit der Interviewpartner*innen verfügte über mehrjährige Berufserfahrung in der Personalauswahl und war zum Zeitpunkt des Interviews in einer Führungsposition tätig (z.B. Leitung der Personalabteilung oder Leitung eines Arbeitsbereichs mit Entscheidungsfunktion bezüglich Stellenbesetzung). Die Interviews dauerten im Durchschnitt 60 Minuten und wurden virtuell durchgeführt. Nahmen mehrere Personalverantwortliche aus demselben Unternehmen teil, wurde die Zeit entsprechend verlängert, um den Interviewpartner*innen genügend Raum für ihre Erfahrungen und Meinungen zu geben. Es wurden Personalverantwortliche aus Unternehmen mit und ohne Erfahrung mit KI in der Personalauswahl rekrutiert. Bei der Auswahl der Unternehmen wurde

zudem auf eine große Bandbreite in Bezug auf Unternehmensgröße und Branche geachtet, um einen bestmöglichen Überblick über den Status Quo in Österreich zu erlangen. Personalverantwortliche aus Unternehmen mit einer Anzahl von zehn bis 18.000 Mitarbeiter*innen konnten befragt werden. Die Branchen umfassten unter anderem die Medien, Automobilindustrie und Finanzdienstleistung. Tabelle 3.1 zeigt eine Übersicht über die Branchen und Unternehmensgrößen der rekrutierten Unternehmen.

Aufgrund von COVID-19 wurden die Interviews als Videokonferenzen durchgeführt. Sowohl die Aufzeichnung, als auch die Verschriftlichung der Aufnahmen erfolgten elektronisch. Im Rahmen einer wörtlichen Transkription wurden Dialekte ins Hochdeutsche übersetzt sowie auf Lückenfüller und Wortdoppelungen verzichtet. Die Interpunktion wurde sinngemäß vorgenommen. Für die Auswertung und Codierung der

Interviews wurde die Software MAXQDA verwendet. Die Kategorien wurden deduktiv auf Basis des Interviewleitfadens und der zuvor durchgeführten Literaturanalyse erstellt und während des Auswertungsprozesses induktiv erweitert, d.h. die entsprechenden Zitate aus den Transkripten wurden den vordefinierten Kategorien zugeordnet, wenn nötig wurden neue Kategorien ergänzt.

Nachfolgend werden die Ergebnisse der Interviewstudie dargestellt. Dazu wird zunächst der Status Quo von KI in der Personalauswahl in österreichischen Unternehmen präsentiert sowie die Barrieren, die Unternehmen und Personalverantwortlichen den Einsatz von KI erschweren, erläutert. Anschließend werden die identifizierten Einsatzgebiete von KI-Technologie in der Personalauswahl diskutiert und Zukunftstrends aufgezeigt.

Tabelle 3.1: Übersicht Unternehmen

Unternehmens-ID	Branche	Unternehmensgröße (Standort Österreich)
E1	Personalberatung	10
E2	Forschung und Entwicklung	150
E3	Medien	3.000
E4	Bauwesen, Beschaffung, Druckzentrum, Facility Management & Reinigung, IT	450
E5	Finanzdienstleistung	350
E6	Automobilindustrie	4.300
E7	Unternehmensberatung, Wirtschaftsprüfung, Steuerberatung	1.400
E8	Elektronikindustrie	1.350
E9	Anlagenbau, Intralogistik	2.300
E10	Papier-, Wellpappe- und Verpackungsbranche	650
E11	Automobilindustrie	10.000
E12	Metallindustrie, Maschinen- und Anlagenbau	500
E13	Gesundheitswesen	18.000
E14	Öffentlicher Dienst, Interessenvertretung	1.000
E15	Telekommunikation, Informationstechnik und Mobilfunk	8.000
E16	Forschung	4.325
E17	Lebensmittelproduktion und -handel	1.400
E18	Unternehmens- und Technologieberatung	530
E19	Personaldienstleistung	6.100
E20	IT-Dienstleistung	6.100
E21	Versicherung	1.000

3.2.2 Status Quo und Barrieren

Der digitale Wandel ist in österreichischen Unternehmen deutlich spürbar. Nahezu alle Interviewpartner*innen berichteten vom Einsatz digitaler Technologien in der Personalauswahl in ihrem Unternehmen bzw. von Änderungen und Neuerungen in den letzten Jahren. Zur Unterstützung der Personalauswahlprozesse werden diverse Informationstechniken eingesetzt, vorwiegend jedoch Bewerbungsmanagement-Systeme, die Bewerbungen, Anmerkungen der Personalverantwortlichen und Videoaufnahmen in Datenbanken archivieren (siehe Abb. 3.1 bis 3.3). Neben der Text- und Datenverarbeitung ermöglichen einige Systeme eine unternehmensinterne Koordination und Kommunikation zu Bewerbungen sowie die direkte Freigabe von Inseratschaltungen. Weiters bieten manche Bewerbungsmanagement-Systeme eine Schnittstelle für Bewerber*innen, die dort in Form eines Bewerbungskontos den Status der Bewerbung einsehen, neue Bewerbungen durchführen oder vorherige Bewerbungen bearbeiten können. Diese Funktionen basieren jedoch nicht auf KI.

Ergänzend zu den intern verwendeten Bewerbungsmanagement-Systemen nutzen Unternehmen Karriereseiten zum Schalten von Inseraten sowie zur Suche und Identifikation von potentiellen Bewerber*innen (siehe Abb. 3.4). In Österreich sind vorwiegend karriere.at, XING und LinkedIn in Verwendung. Auf Basis von Algorithmen ermöglichen diese Seiten eine gezielte Kandidat*innensuche und erstellen Vorschläge mit passenden Kandidat*innen, die daraufhin von den Personalverantwortlichen kontaktiert werden können. Es handelt sich teilweise um intelligente Algorithmen, was jedoch einer großen Mehrheit der Interviewpartner*innen nicht bewusst war.

Gerade in Zeiten von COVID-19 setzen Unternehmen vermehrt gängige Videokonferenzsysteme für das Führen von Bewerbungsgesprächen ein. In einem der Unternehmen wurden zusätzlich die Programme Cammio, Cut-e sowie HireVue zur automatischen Analyse von Bewerbungsvideos durch einen Algorithmus angegeben (siehe Abb. 3.5). Es werden Gesichtsausdrücke, die Wortwahl sowie die Stimmlage der Bewerber*innen ausgewertet, darauf basierend können unter anderem Persönlichkeitsprofile der Bewerber*innen erstellt werden.

ber*innen ausgewertet, darauf basierend können unter anderem Persönlichkeitsprofile der Bewerber*innen erstellt werden.

Die genannten Systeme bieten eine Vielzahl von Funktionen an, darunter zum Teil auch automatisierte und KI-basierte Lösungen. Dennoch schätzen viele Personalverantwortliche die eingesetzten Systeme nicht als KI ein. Es kann angenommen werden, dass Personalverantwortliche oft nicht über den vollständigen Funktionsumfang sowie die Funktionsweise der eingesetzten Systeme in Kenntnis sind. Teilweise wird der Einsatz automatisierter bzw. KI-basierter Funktionen eines bereits verwendeten Systems bewusst abgelehnt, weil die Personalverantwortlichen keinen Mehrwert für ihre Arbeit erkennen können.

Obwohl KI in der Personalauswahl von den Interviewpartner*innen durchwegs als vielversprechender Trend wahrgenommen wird, nehmen österreichische Unternehmen vom praktischen Einsatz bisher Abstand. Dies zeigt sich in den Ergebnissen der Interviewstudie: Trotz gezielter Suche nach österreichischen Unternehmen, die bereits KI in der Personalauswahl einsetzen, konnten nur wenige Unternehmen identifiziert werden und nur zwei Unternehmen tatsächlich für Interviews gewonnen werden. Beide Firmen verwenden Dashboards, wobei ein Unternehmen dieses ausschließlich für die interne Personalbesetzung nutzt. Bei einem dritten Unternehmen befinden sich aktuell zwei Chatbots in der Testphase, wobei einer für die Kommunikation mit bestehenden Mitarbeiter*innen und der andere für die Kommunikation mit Bewerber*innen eingesetzt werden soll. Einige weitere Unternehmen beobachten die Entwicklungen rund um das Thema KI in der Personalauswahl aktiv, setzen derzeit aber noch keine entsprechende Technologie ein.

Der zurückhaltende Einsatz von KI in der Personalauswahl kann auf vier Barrieren zurückgeführt werden, welche im Rahmen der Interviewstudie identifiziert wurden: (1) Nutzen-Aufwand-Verhältnis; (2) Reife der Technologie; (3) Angst vor dem unfreiwilligen Verlust von Bewerber*innen; (4) Angst interessante Aufgaben zu verlieren. Die Barrieren werden im Folgenden näher erläutert.

Künstliche Intelligenz in der Personalauswahl

	Kurzbeschreibung	Automatisierte und KI-Funktionen
	<ul style="list-style-type: none"> • Bewerbungsmanagement-System • Talentpool • Employer Branding durch Karriereportal • Videokonferenzen für Interviews • Unternehmensinterne Kommunikation 	<ul style="list-style-type: none"> • Automatisiertes CV-Parsing • Intelligentes Multiposting • Matching
	<ul style="list-style-type: none"> • Personalmanagement-System inkl. Bewerbungsmanagement • Talentpool • Bewertung der Bewerber*innen • Unternehmensinterne Kommunikation 	<ul style="list-style-type: none"> • Automatisierte Stellenausschreibungen • Gehaltsempfehlungen basierend auf maschinellem Lernen • Automatisierung von Prozessschritten
	<ul style="list-style-type: none"> • Personalmanagement-System inkl. Bewerbungsmanagement • Multiposting • Karriereportal und Anbindung an Social Media • Talentpool & Active Sourcing • Berufseignungstest 	<ul style="list-style-type: none"> • Automatisiertes CV-Parsing • Automatisierte Vorauswahl mit Bewerber*innenranking (Matching)

Abbildung 3.1: Bewerbungsmanagement-Systeme – Teil 1

	Kurzbeschreibung	Automatisierte und KI-Funktionen
	<ul style="list-style-type: none"> • HR-Software inkl. Bewerbungsmanagement • Erstellen von Ausschreibungen • Anbindung an Jobbörsen • Unternehmensinterne Kommunikation 	<ul style="list-style-type: none"> • Dataming und Big-Data für Personalorganisation und Controlling
	<ul style="list-style-type: none"> • System für Personalentwicklung & interne Besetzungen 	<ul style="list-style-type: none"> • KI-basiertes Vorhersage zukünftig benötigter Fähigkeiten • Automatisierte und personalisierte Karriereberatung von Mitarbeiter*innen • Dashboard für Mitarbeiter*innen und Personalist*innen • Matching für interne Besetzungen • Automatisiertes CV-Parsing
	<ul style="list-style-type: none"> • Bewerbungsmanagement-System • Stellenausschreibungen • Anbindung an Jobbörsen & Social Media • Karriereportal • Talentpool • Unternehmensinterne Kommunikation 	<ul style="list-style-type: none"> • Automatisiertes CV-Parsing

Abbildung 3.2: Bewerbungsmanagement-Systeme – Teil 2




	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Bewerbungsmanagement-System • Talentpool • Videonachrichten der Bewerber*innen • Assessments & Fragebögen • Unternehmensinterne Kommunikation • Multiposting 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Matching - Algorithmus • Automatisiertes CV-Parsing • Teil- und vollautomatisierte Kommunikation mit Bewerber*innen
	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Bewerbungsmanagement-System • Talentpool • Erstellen von Ausschreibungen • Anbindung an Jobbörsen • Unternehmensinterne Kommunikation • Videokonferenzen für Interviews 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Automatisiertes Multiposting • Intelligente Suche im Talentpool
	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Talentmanagement- und Bewerbungsmanagement-System • Talentpool 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Matching • KI-basiertes Management bestehender Mitarbeiter*innen-Fähigkeiten • Automatisiertes CV-Parsing

Abbildung 3.3: Bewerbungsmanagement-Systeme – Teil 3




	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Karriereseite • Ausschreibung von Stellen • Suche potentieller Kandidat*innen 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Automatisierte Analyse von Kandidat*innen und Inseraten • Ontologiebasiertes Matching
	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Soziales Netzwerk • Active Sourcing in Talentpools • Ausschreibung von Stellen 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Automatisierte Vorschläge von Kandidat*innen
	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • Soziales Netzwerk • Active Sourcing in Talentpools • Ausschreibung von Stellen 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Intelligente Algorithmen für Kandidat*innen-Vorschläge
	<p>Kurzbeschreibung</p> <ul style="list-style-type: none"> • "Mitarbeiter*innen werben Mitarbeiter*innen" - Plattform 	<p>Automatisierte und KI-Funktionen</p> <ul style="list-style-type: none"> • Automatisiertes Matching

Abbildung 3.4: Karriereseiten

Nutzen-Aufwand-Verhältnis

Der Einsatz von KI wird in vielen Fällen mit einem hohen Programmier- sowie Kostenaufwand assoziiert. Im Zusammenspiel mit einer geringen Anzahl an Bewerbungen und einer kleinen Unternehmensgröße scheint der Verbrauch dieser Ressourcen als nicht gerechtfertigt bzw. die Anwendung von KI in der Personalauswahl als nicht ausreichend profitabel. Allerdings hatten die wenigsten Personalverantwortlichen konkrete Vorstellungen darüber, wie viel Zeit- und Kostenaufwand für die Implementierung und Wartung eines KI-Systems erforderlich ist.

Reife der Technologie

Es gibt aktuell die Wahrnehmung, dass KI-Systeme noch nicht ausgereift genug sind, um tatsächlich als Unterstützung eingesetzt zu werden. Personalverantwortliche haben die Befürchtung, dass Fehler in der Datenerhebung zu Problemen und Mehraufwand im gesamten Personalauswahlprozess führen. Hier sind sie vor allem von negativen Vorerfahrungen mit Chatbots geprägt, die Fragen nur unzureichend beantworten konnten.

Angst vor dem unfreiwilligen Verlust von Bewerber*innen

Personalverantwortliche haben Angst, dass aufgrund von Algorithmen vielversprechende Bewerber*innen aussortiert werden könnten, wenn die Bewerbungsunterlagen bestimmte Stichwörter nicht enthalten oder Bewerber*innen nicht dem typischen Profil für eine Stelle entsprechen. Darüber hinaus besteht die Befürchtung, dass der Einsatz von KI vor allem auf ältere oder wenig technikaffine Bewerber*innengruppen abschreckend wirken könnte, was zu einem Verlust des Interesses an der ausgeschriebenen Stelle oder zu einem frühzeitigen Abbruch der Bewerbung führen könnte.

Angst, interessante Aufgaben zu verlieren

Die Erleichterung der Arbeit für Personalverantwortliche gilt als einer der wesentlichen Vorteile, welcher

durch den Einsatz von KI in der Personalauswahl erzielt werden soll. Tatsächlich wird KI von Personalverantwortlichen jedoch teilweise als Bedrohung wahrgenommen. Es gibt die Befürchtung, dass Personalverantwortliche interessante Aufgabenfelder sowie Entscheidungsfreiheiten an KI verlieren könnten. Die Angst, in der Zukunft vollständig von einer KI ersetzt zu werden, haben hingegen nur wenige Personalverantwortliche.

Trotz der Barrieren, die den Einsatz von KI in der Personalauswahl in österreichischen Unternehmen aktuell noch hemmen, wird die Technologie als vielversprechend und etwas, das weiter beobachtet werden sollte, wahrgenommen. Es besteht grundsätzlich eine hohe Bereitschaft, KI zukünftig für bestimmte Einsatzgebiete zu verwenden. Diese Einsatzgebiete werden im nächsten Kapitel näher beleuchtet.

3.2.3 Einsatzgebiete von KI in der Personalauswahl

Trotz der geringen Erfahrungen mit KI konnten die befragten Personalverantwortlichen klare Vorstellungen nennen, wofür diese Technologie in der Personalauswahl eingesetzt werden soll und wofür nicht. In den Interviews wurden den Personalverantwortlichen Fallbeispiele zu einem Chatbot und einem Dashboard präsentiert (siehe Interviewleitfaden im Anhang), um konkretere Vorstellungen, Erwartungen und Befürchtungen erörtern zu können. Diese Fallbeispiele sowie die in den Interviews ermittelten Einsatzmöglichkeiten von KI in der Personalauswahl werden im Folgenden dargestellt.

Chatbot

Fallbeispiel: Eine Bewerberin interessiert sich für eine ausgeschriebene Professur an der Karl-Franzens-Universität Graz und besucht die offizielle Internetseite. Die ausgeschriebene Stelle wird dort zwar beschrieben, jedoch möchte sie sich weitere Informationen einholen. Dazu klickt sie auf das entsprechende Feld und das Chatbot-Fenster öffnet sich. Die Bewerberin hat nun die Möglichkeit, ihre Fragen einzugeben, welche vom Chatbot in Echtzeit beantwortet werden (siehe Abb. 3.6).

Das Bereitstellen von Informationen durch Chatbots wurde von den Personalverantwortlichen durchgängig als hilfreiche Unterstützung im Personalauswahlprozess gesehen:

“Dort, wo wir es uns durchaus vorstellen können, ist in der Kommunikation mit Bewerberinnen, Bewerbern im Rahmen von einem Chatbot, weil durchaus sehr wiederholende Fragen auch immer wieder auf uns zukommen. Das heißt, quasi in der Betreuung der Kandidaten und Kandidatinnen können wir uns durchaus vorstellen, dass dort Potenzial ist.” (Leitende*r Personalverantwortliche*r, Unternehmensberatung, Wirtschaftsprüfung & Steuerberatung, 1.400 MA)

Viele Fragen im Rahmen des Bewerbungsprozesses werden wiederholt von Bewerber*innen gestellt und könnten daher gut von Chatbots bearbeitet werden. Neben den klassischen Frequently Asked Questions (FAQs - also häufig gestellte Fragen), denken Personalverantwortliche dabei auch an Beratungen durch Chatbots, wie das Aufzeigen anderer passender und interessanter Stellen nach Eingabe des Lebenslaufs und der Interessen. Informationsgespräche sind allerdings nur eine Möglichkeit, wie Chatbots in der Personalauswahl eingesetzt werden können.

Fallbeispiel: Das Informationsgespräch mit dem Chatbot hat die Bewerberin darin bekräftigt, ihre Bewerbung für die ausgeschriebene Professur einzureichen. Sie ruft über das entsprechende Feld das Chatbot-Fenster auf und startet damit ein Bewerbungsgespräch mit dem Chatbot direkt auf der Internetseite der Universität. Analog zu einem herkömmlichen Bewerbungsgespräch stellt der Chatbot zunächst allgemeine Fragen, die im Verlauf des Gesprächs immer gezielter und spezifischer für die ausgeschriebene Stelle werden (siehe Abb. 3.7).

Die Übernahme von Bewerbungsgesprächen durch Chatbots wird von österreichischen Personalverantwortlichen sehr unterschiedlich beurteilt. Die fehlende “menschliche Komponente”, also möglicherweise fehlende Flexibilität sowie emotionale Intelligenz bei der Durchführung dieser Tätigkeit, führt vorwiegend zu Bedenken. Nur sehr wenige können sich vorstellen, dass diese sehr komplexe Tätigkeit zur Zufriedenheit der Personalverantwortlichen, aber auch Bewerber*innen, von Chatbots übernommen werden könnte. Jedoch wird es als hilfreich gesehen, wenn ein Chatbot

Bewerbungsunterlagen über die Möglichkeit des Hochladens sammelt, eventuell Persönlichkeitstests durchführt und die Bewerber*innen nach der Bewerbung weiter betreut. So können Bewerber*innen direkt vom Chatbot den Status ihrer Bewerbungen erfragen bzw. können automatische Updates zum Bewerbungsstatus von ihm erhalten. Bei einer Zusage könnte ein Chatbot die ersten Schritte des Onboarding, wie Fragen zum ersten Arbeitstag und erforderliche Dokumente, mit den Bewerber*innen klären.

Dashboard (graphische Datenvisualisierung)

*Fallbeispiel: Die Personalverantwortlichen sehen auf einer Übersichtsseite alle Bewerber*innen für eine Stelle aufgelistet. Diese enthält einen Matching-Score mit der Stelle, der automatisiert von der KI berechnet wurde (siehe Abb. 3.8). Die Daten dazu stammen beispielsweise aus einem Chatbotverlauf, den Bewerbungsunterlagen oder Profilen auf Sozialen Medien. Die Personalverantwortlichen haben die Möglichkeit, sich mehr Informationen zu den Bewerber*innen einzuholen und eine detailliertere Ansicht aufzurufen (siehe Abb. 3.9). Hier wird der Matching-Score auf die drei Bereiche Bildung, Befähigung und Persönlichkeit aufgegliedert und relevante Schlüsselwörter zu den Bewerbungen dargestellt. Für alle Bewerber*innen kann eine dritte Detailansicht aufgerufen werden (siehe Abb. 3.10). Hier kann beispielsweise für Persönlichkeit eingesehen werden, welches Profil die Bewerber*innen mitbringen und was für die Stelle erforderlich ist.*

Insbesondere das Matching im Rahmen einer Vorselektion der Bewerber*innen wird als vielversprechendes Einsatzgebiet von Dashboards gesehen. Dabei erfolgt ein Vergleich der Anforderungen einer Stelle mit den Qualifikationen der Bewerber*innen. Grundsätzlich wird das Dashboard von Personalverantwortlichen als hilfreiches Hilfsmittel gesehen, um Ergebnisse zu visualisieren und übersichtlich darzustellen.

“Ich finde ein Dashboard gut als Entscheidungsgrundlage, ein Dashboard sollte aber keine Entscheidungen treffen. Es soll einfach die entsprechenden Informationen liefern, um dann gut und rasch zu Entscheidungen zu kommen.” (Leitende*r Personalverantwortliche*r, Anlagenbau & Intralogistik, 2.300 MA)

Im Rahmen der Gespräche betonten die Personalverantwortlichen, dass sie zwar den Einsatz von Dashboards in Vorselektionen als sinnvoll erachten, sie jedoch lediglich als Entscheidungshilfe dienen sollte. Die endgültigen Entscheidungen sollen weiterhin von menschlichen Personalverantwortlichen getroffen werden. Es bestand Uneinigkeit darüber, ob Dashboards automatisiert Absagen beziehungsweise Zusagen schicken sollen.

Weitere Funktionen, die ein Dashboard übernehmen soll, umfassen die automatische Schaltung von Stelleninseraten, die Suche und Identifikation von potentiellen Kandidat*innen in Datenbanken, die Unterstützung bei der internen Personalentwicklung sowie das Erstellen von Statistiken zur Personalauswahl im Unternehmen. Das Erstellen eines Persönlichkeitsprofils mithilfe eines Dashboards wurde hingegen als kritisch angesehen, weil Personalverantwortliche dafür eine "menschliche Komponente" in der Erhebung und Beurteilung der Persönlichkeit voraussetzen.

Zusammenfassend lässt sich feststellen, dass besonders eine Vorselektion von Bewerber*innen mithilfe eines Matching sowie die Bereitstellung von Informationen durch KI als interessant und unterstützend beurteilt wurden. Das deckt sich mit den Einsatzbereichen, in denen einige österreichische Unternehmen erste Umsetzungsversuche gestartet bzw. angedacht haben.

3.2.4 Trends

Die Offenheit der Personalverantwortlichen gegenüber KI in der Personalauswahl, unter der Prämisse dass die Technologie ausgereift ist, zeigt sich in den Zukunftsvorstellungen zu diesem Thema. Die Personalverantwortlichen wurden gefragt, welche Trends sie sich im Bereich von KI in der Personalauswahl vorstellen können. Es hat sich herausgestellt, dass für die meisten Personalverantwortlichen bzw. Unternehmen in Österreich der Einsatz von KI in oben beschriebener Form noch ein Trend der Zukunft ist, wenn auch bereits der nahen Zukunft. KI in der Personalauswahl z.B. Chatbots zur Terminvereinbarung, Kommunikation und Beantwortung von FAQ, zum eigenständigen Führen von Interviews sowie Dashboards für Matching

müssen erst einsatzfähig werden. Damit ist auch gemeint, dass die Vorschläge und Entscheidungen nicht regelbasiert, sondern tatsächlich auf Basis historischer Daten entstehen. Längerfristig ist es für manche aber im Bereich des Vorstellbaren, dass die ersten Schritte in der Personalauswahl durch KI übernommen werden und Personalist*innen erst später, einmalig Kontakt zu den Kandidat*innen haben. Die Beurteilung von Texten oder Arbeitsproben wird großteils als sinnvoller nächster Schritt von KI gesehen. Anwendungsbeispiele sind einerseits die Beurteilung der Qualität der Arbeitsproben, andererseits Rückschlüsse über die Persönlichkeit aufgrund sprachlicher Muster. Für ersteres muss KI allerdings auch in der Lage sein, "die Bedeutung" von Texten sinngemäß zu verstehen. Weiterer Trend von KI ist die Videoanalyse von z.B. Bewerbungsvideos. KI soll daraus zusätzlich zu den Inhalten auch sprachliche Fähigkeiten und Muster sowie Mimik und Gestik beurteilen können und so ein Persönlichkeitsprofil mit Eignung für bestimmte Stellen erstellen. Weiterführend soll KI Persönlichkeits- oder Eignungsprofile durch die Analyse von Social Media-Profilen erstellen können oder im Rahmen von Assessments mit z.B. Rollenspielen eingesetzt werden. Hier sind auch Tests durch interaktive Spiele, im Sinne von Gamification, denkbar. Das volle Ausmaß der Einsatzmöglichkeiten sehen viele Personalist*innen noch nicht als vollständig abschätzbar, jedoch muss der Einsatz über die Personalauswahl hinausgehend gedacht werden, z.B. im Rahmen des Onboardings neuer Mitarbeiter*innen oder für bestehende Mitarbeiter*innen in der Personalentwicklung, wo Weiterbildungsmöglichkeiten und -interessen ermittelt werden könnten. Auch im administrativen Bereich wird KI zukünftig als Unterstützung gesehen, beispielsweise beim automatischen Erstellen von Arbeitsverträgen und von Stelleninseraten, sowie bei der Auswahl von Jobbörsen auf Basis von errechneten Erfolgchancen.

Skeptischer wurde der Einsatz von Robotern in der Personalauswahl gesehen. Roboter, die beispielsweise das Bewerbungsgespräch führen oder eingestellte Bewerber*innen am ersten Arbeitstag begleiten und onboarden, wurden als überflüssig wahrgenommen. Interessantere Umsetzungsmöglichkeit von KI waren audiobasierte Systeme, also Systeme, die keine schriftlichen Eingaben erfordern, sondern mündlich mit Daten gefüttert werden, da sie mit einer höheren Usability für

Bewerber*innen assoziiert wurden. Außerdem wurde betont, dass es zuallererst wichtig sei, vorurteilsfreie KI hinzubekommen. Einen letzten interessanten Zukunftstrend stellt das persönliche Gespräch mit Personalverantwortlichen dar:

“Ich frage mich allerdings, [...] ob es irgendwann wieder eine Auszeichnung ist, wieder mit einer Person zu sprechen. Im Sinne von, die Stelle ist so oder das Unternehmen ist so. . . Ich bin denen so wichtig, dass tatsächlich ein Mensch mit mir spricht. [...] Also, wenn ich mir jetzt überlege, das Pendel geht Richtung Technik, jetzt probieren wir alles Mögliche zu optimieren und dann, wenn es

*wieder zurück geht, ob das vielleicht dann einmal ein Qualitätskriterium ist. Werden wir beobachten.” (Leitende*r Personalverantwortliche*r, Gesundheitswesen, 18.000 MA)*

Während im Moment versucht wird, Technologien als Unterstützung für Menschen zu optimieren und diese ein Novum sind, sieht diese*r Personalist*in den Trend in der Zukunft zurück Richtung menschliche Interaktion: wenn sich das Unternehmen die menschliche Arbeitskraft und ihre Zeit für Bewerber*innen leisten, könnte das als Wertschätzung den Bewerber*innen gegenüber interpretiert werden.



Abbildung 3.5: Videoanalysetools

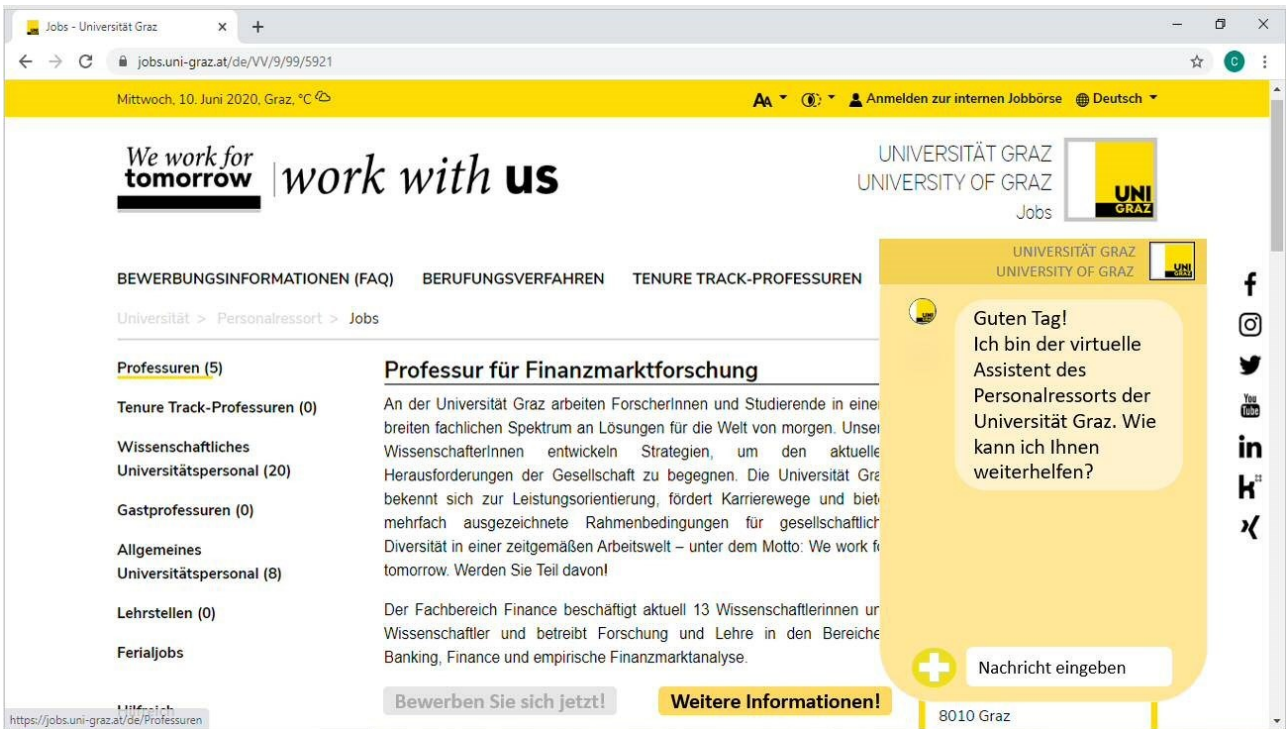


Abbildung 3.6: Fallbeispiel Chatbot – Informationsgespräch

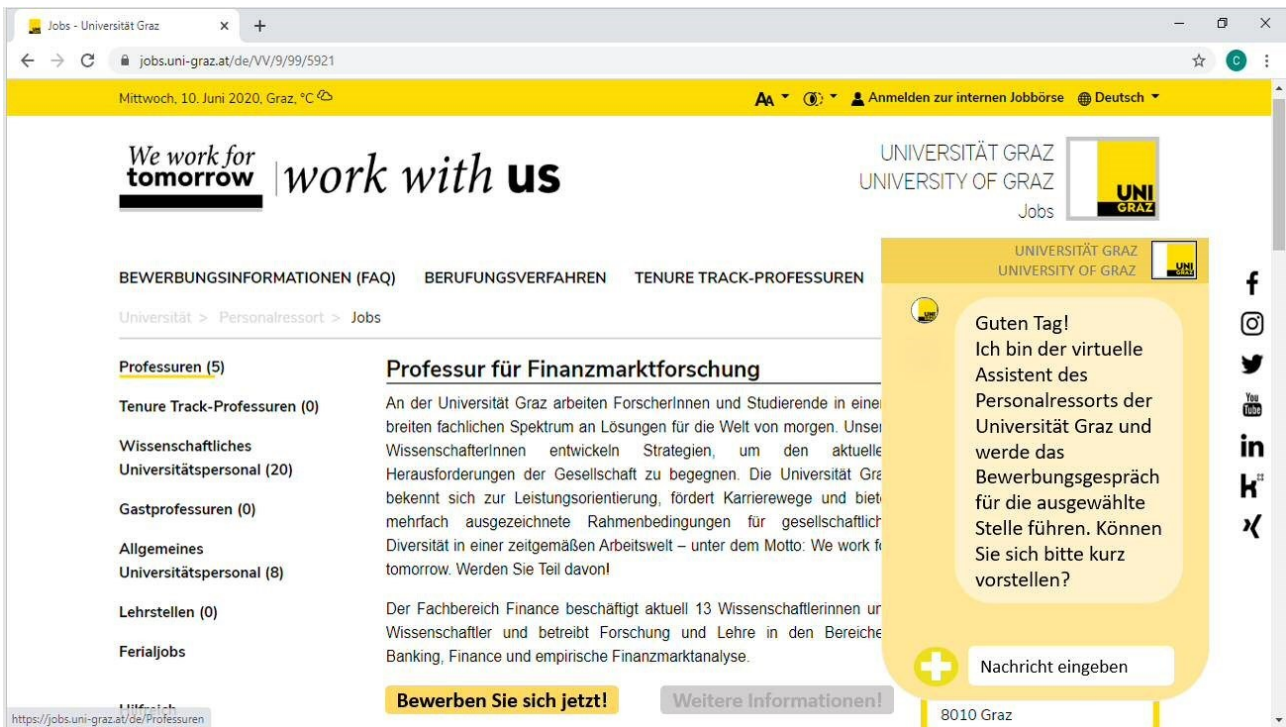


Abbildung 3.7: Fallbeispiel Chatbot – Bewerbungsgespräch

Künstliche Intelligenz in der Personalauswahl

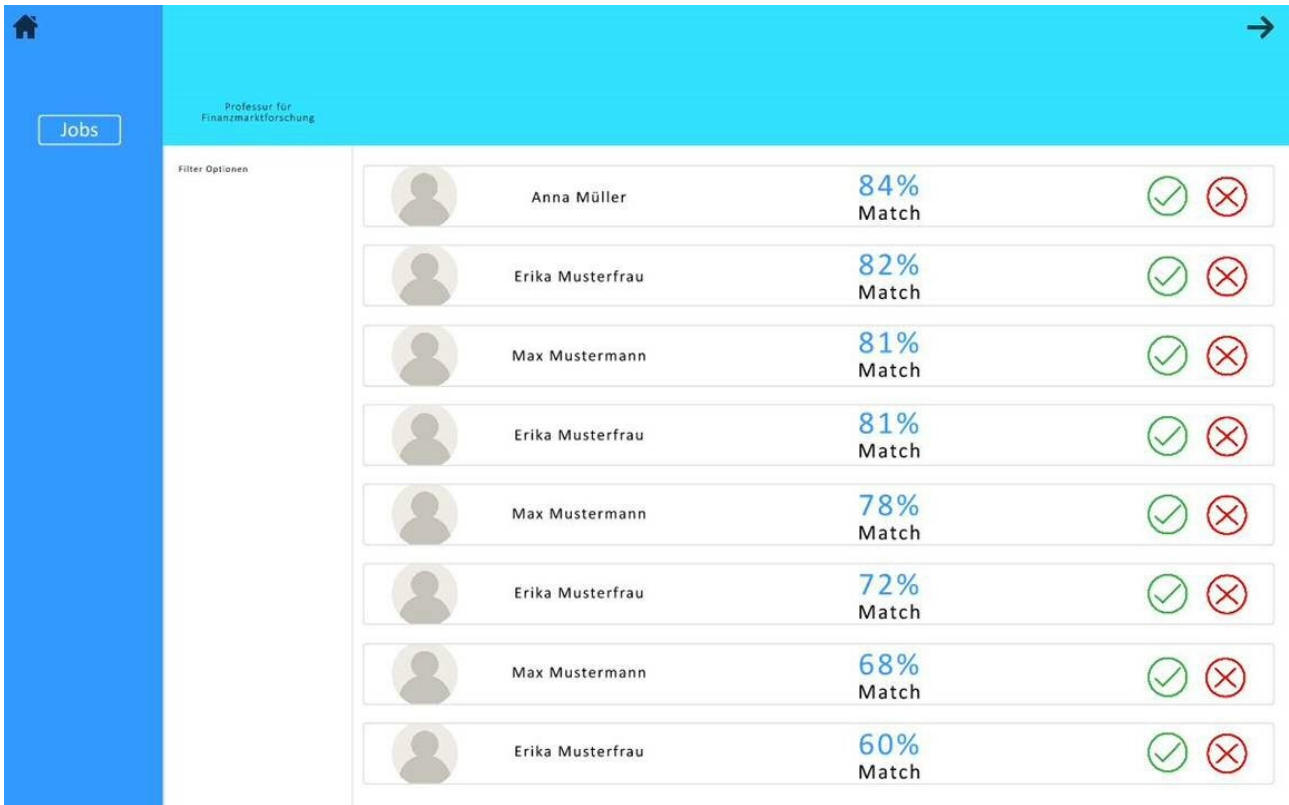


Abbildung 3.8: Fallbeispiel Dashboard – Ebene 1

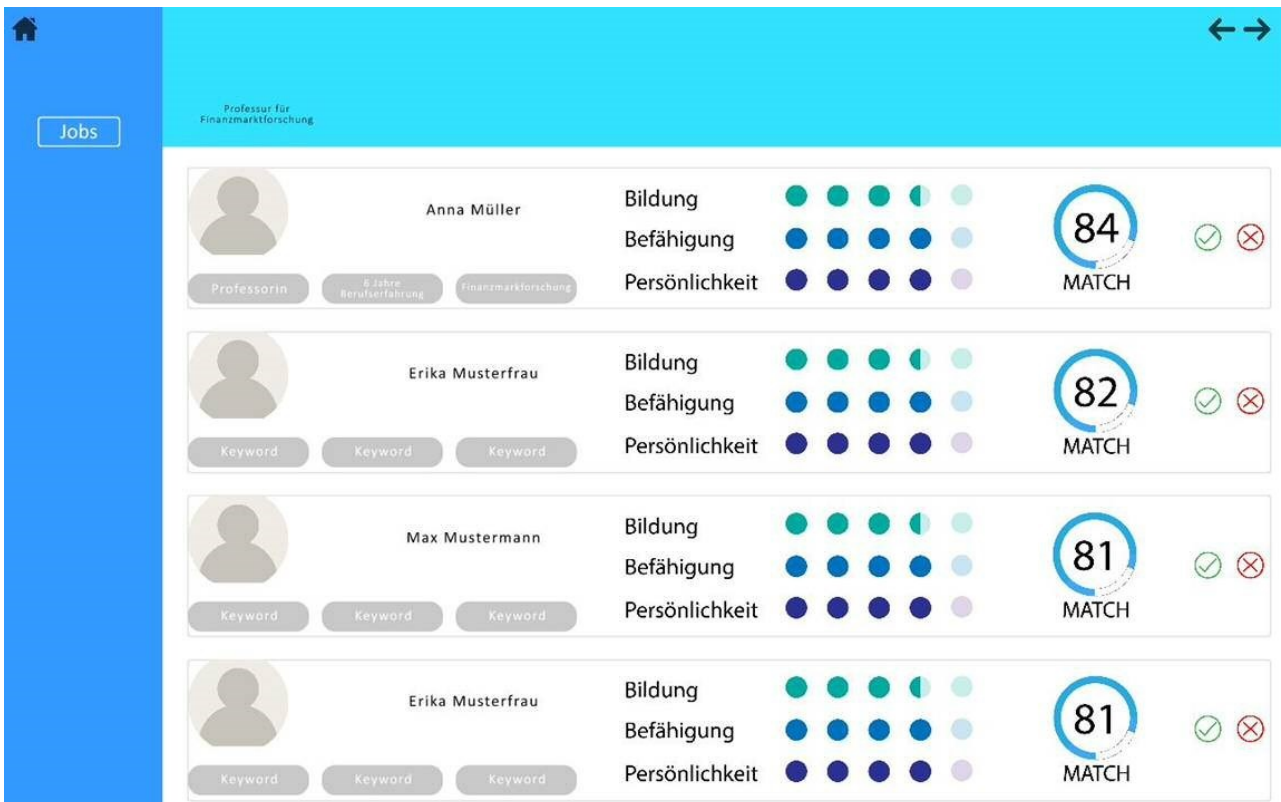


Abbildung 3.9: Fallbeispiel Dashboard – Ebene 2

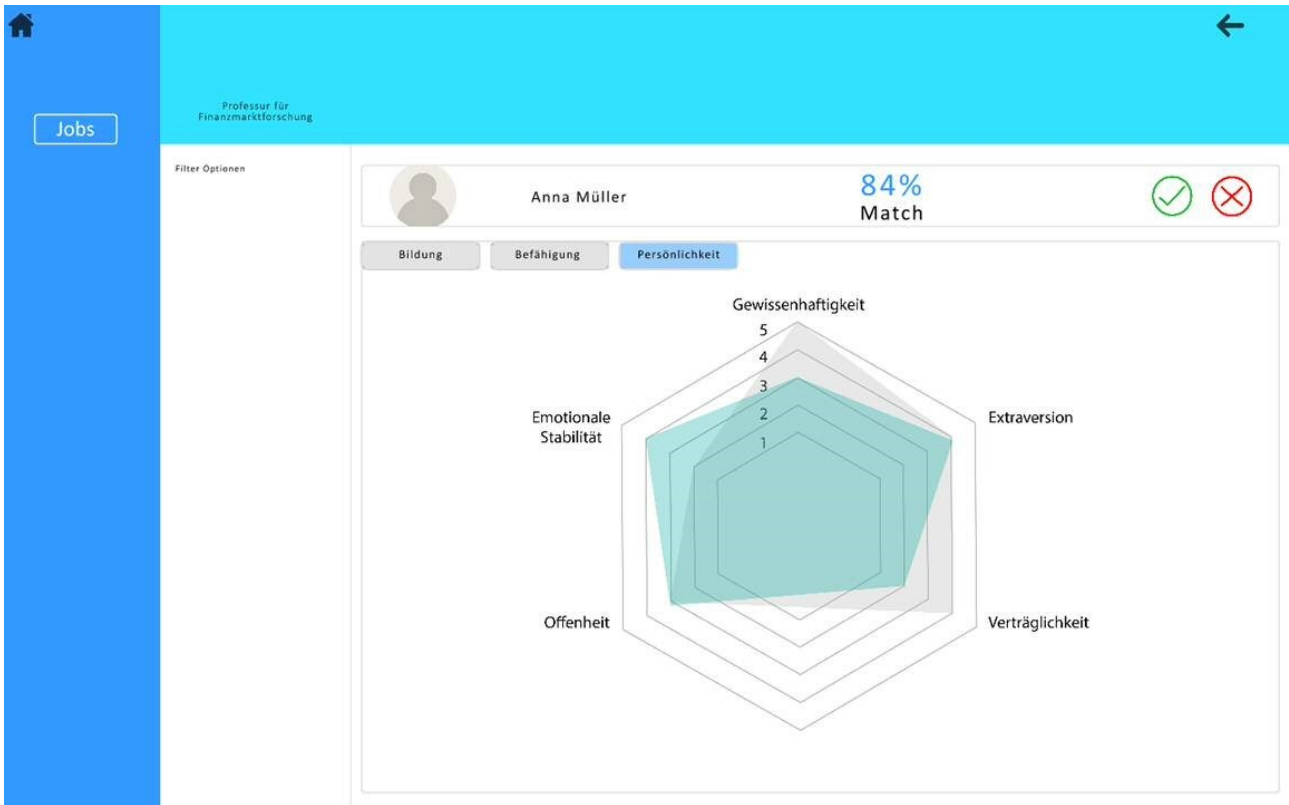


Abbildung 3.10: Fallbeispiel Dashboard – Ebene 3

4 Perspektive Bewerber*innen

Um die Perspektive der Bewerber*innen zur Nutzung von KI-Systemen zu erfassen wird (a) auf Erkenntnisse aus bestehenden wissenschaftlichen Studien zurückgegriffen und (b) eine Sekundäranalyse von zwischen Januar und Februar 2020 erhobenen Umfragedaten vorgenommen. Schließlich (c) werden Ergebnisse einer im Zuge des Projektes durchgeführten quotenrepräsentativen, standardisierten Umfrage präsentiert.

4.1 Bewerber*innenperspektive in der Literatur

KI-Systeme sind bereits jetzt oft klassischen Computersystemen aber auch menschlichen Expert*innen überlegen (Wanner et al. 2020). Wenig überraschend werden daher KI-Systeme in verschiedenen Bereichen verwendet. Aber es gibt auch viele Bereiche, in denen KI-Systeme ein hohes Potential bieten, jedoch nicht eingesetzt werden. Denn eine hohe technische Leistungsfähigkeit von KI-Systemen (wie bspw. eine hohe Prognosegenauigkeit oder eine gute Erkennungsrate bei Bildern) ist alleine nicht immer ausreichend für deren erfolgreichen Praxiseinsatz. Die Einführung von KI-Systemen im Speziellen und IT-Systemen im Allgemeinen ist keine ausschließlich technische Frage (Hansen et al. 2019).

Auch bei der Anwendung von KI in der Personalauswahl zeigt sich, dass vielfältige Faktoren beeinflussen, wie KI von Bewerber*innen tatsächlich an- und wahrgenommen wird. Ähnlich wie bei der Diskussion um die Akzeptanz von KI aus der Sicht von Personalist*innen (siehe Abschnitt 3.1) zeigt eine Studie, dass Bewerber*innen aktuell noch eine höhere Bewerbungsabsicht haben, wenn ihre Bewerbung von Menschen und nicht von KI bewertet wird (Mirowska, 2020). Allerdings können zum Beispiel eine positive Wahrnehmung vom Unternehmen sowie die Neuartigkeit der KI-Technologie die Absicht, einen Bewerbungsprozess mit KI vollständig zu durchlaufen, positiv beeinflussen, während Ängste gegenüber der KI-Technologie diese

Absicht vermindern (Van Esch, Black und Ferolie, 2019; Esch, Black und Arli, 2020). Bewerber*innen, die häufig soziale Medien nutzen, KI-Technologie als modern und angesagt empfinden, und zudem das Gefühl haben, von der KI fair behandelt zu werden, haben eine höhere Bereitschaft sich im Rahmen eines KI-gestützten Personalprozesses zu bewerben und diesen vollständig abzuschließen (Van Esch und Black, 2019).

Wahrgenommene Fairness wurde in mehreren Studien als wichtiger Faktor für die Akzeptanz von KI in der Personalauswahl gefunden. Aktuell werden menschliche Bewerbungsgespräche und Personalentscheidungen fairer beurteilt als KI-gestützte Personalauswahl (Acikgoz u. a., 2020; Dineen, Noe und Wang, 2004; Gonzalez u. a., 2019; Nørskov u. a., 2020). Das Gefühl geringerer Fairness wirkt sich auch negativ auf die Akzeptanz hochautomatisierter Interviews aus, in denen Bewerber*innen zudem das Gefühl von Kontrollverlust erleben (Langer, König und Papathanasiou, 2019). Andererseits vermitteln hochautomatisierte Interviews den Eindruck, dass Entscheidungsprozesse einheitlich für alle Bewerber*innen durchgeführt werden und damit fair sind (Langer u. a., 2019).

Um Akzeptanz gegenüber der KI bei Bewerber*innen zu fördern, muss das Gefühl der Fairness entstehen und Unsicherheiten sowie falsche Annahmen über die KI-Systeme abgebaut werden. Dabei ist es wenig förderlich, dass viele KI-Systeme eine Black-Box darstellen, d.h. Menschen haben keinen direkten Einblick, wie die Daten verarbeitet werden und warum eine bestimmte Entscheidung getroffen wurde. Darauf soll im Folgenden näher eingegangen werden.

Das-Black Box Problem von KI

Einer der besonders kritischen Gründe für die Nichteinführung von KI-Systemen ist der Black-Box Charakter von KI, bei denen die Gründe für die getroffenen Entscheidungen nicht nachvollzogen werden können. Insbesondere zwei Probleme entstehen durch den Black Box Charakter von KI:

1. Die Intransparenz der von durch KI-Systeme getroffenen Entscheidungen kann eine geringe Akzeptanz oder ein geringes Vertrauen durch nachgestellte menschliche Entscheider*innen sowie

durch Personen, die von diesen Entscheidungen betroffen sind, zur Folge haben. Diese geringe Akzeptanz bzw. das geringe Vertrauen kann dann wieder den Einsatz dieser KI-Systeme erschweren oder sogar unmöglich machen.

2. Ein weiteres Problem aus Sicht der Bewerber*innen sind die zu erfüllenden rechtlichen Regelungen für den Einsatz von KI-Systemen, die von Black-Box Algorithmen nicht ohne zusätzliche Anstrengungen erfüllt werden können. Eine Erklärbarkeit der Entscheidung kann hier von Bewerber*innen verlangt werden (Königstorfer and Thalmann 2020).

Besonders relevant für die Perspektive der Bewerber*innen ist der Fall der Akzeptanz bzw. des Vertrauens in durch KI getroffene Entscheidungen. Betrachten wir dazu zuerst zwei Fälle außerhalb des Personalauswahl-Bereichs. So werden KI-Systeme bspw. eingesetzt, um (a) die Kreditwürdigkeit von Personen einzuschätzen und dies inkludiert unter anderem auch schon Daten aus Facebook-Profilen (Rothmann et al. 2014). Die Entscheidung über die Kreditwürdigkeit hat natürlich starke Auswirkungen für die Betroffenen. Potentiell noch schwerwiegender können die Konsequenzen automatisierter Entscheidungen bspw. bei (b) durch KI gesteuerten autonomen Fahrzeugen sein. Diese waren dabei auch bereits in tödliche Unfälle verwickelt (Marshall 2018). In beiden Fällen ist intuitiv klar, dass wir Erklärung verlangen und wissen möchten, warum (a) kein Kredit gewährt wurde und (b) die KI das Fahrzeug nicht gestoppt hat. Die Beantwortung der letzten Frage dauerte schließlich aufgrund der Komplexität des KI-Modells mehrere Wochen (Gilpin et al. 2019).

Auch der Bereich der Personalauswahl ist aus Sicht der Bewerber*innen mit weitreichenden Konsequenzen verbunden und entsprechend erwarten wir auch hier intuitiv nach (zumindest der Möglichkeit) einer Erklärung. Diese für uns intuitive Erwartung einer Erklärung wird auch durch empirische Studien gestützt. So wurde beispielsweise gezeigt, dass sowohl Personalist*innen als auch Bewerber*innen von KI-Systemen in der Personalauswahl eine transparente und begründbare Entscheidung erwarten und Black-Box Systemen nicht vertrauen (Ochmann und Laumer, 2019).

Es kann somit gesagt werden, dass die unfaire und intransparente Wahrnehmung von KI, die Akzeptanz im Allgemeinen (Guidotti et al. 2018; Ribeiro et al. 2016; Ye and Johnson 1995) und speziell bei sensiblen Entscheidungen negativ beeinflusst (Binns u. a., 2018; Lee, 2018; Schumann u. a., May 2020; Zerilli u. a., 2019). Damit variiert auch die Dringlichkeit, mit der eine Erklärung erwartet wird. Während die konkrete Auswahl einer uns auf Facebook gezeigten Werbung in den meisten Fällen wohl keiner Erklärung bedarf sieht es bei dem oben genannten Beispiel des Kreditantrages schon anders aus.

Aus rechtlicher wie auch aus Nutzer*innensicht ist die Erklärung von Entscheidungen also zentral. Transparenz durch Erklärungen hilft den erfolgreichen Einsatz zu realisieren. Daher ist es wenig verwunderlich, dass aktuell auf technischer Seite große Anstrengungen im Bereich der Erforschung von Erklärungskomponenten für KI-Entscheidungen unternommen werden.

In der Vergangenheit kam es zu zahlreichen Kontroversen hinsichtlich KI-Systemen, die verzerrte und diskriminierende Entscheidungen getroffen haben (Adadi and Berrada, 2018). Diesbezügliche Zweifel wurden unter anderem durch medienwirksame Vorfälle verstärkt. Ein bekanntes Beispiel ist der automatisierte Bewerbungsprozess des amerikanischen Onlineversandhändlers Amazon, wobei die KI weibliche Bewerberinnen bei der Personalauswahl systematisch benachteiligte (Dastin, 2018). Wie im vorherigen Teil bereits gezeigt wurde, besteht seitens der Nutzer*innen zunehmender Bedarf, Begründungen und Rechtfertigungen für KI-Ergebnisse sicherzustellen. Der verstärkte Wunsch nach erklärbarer KI zeigt sich vor allem in sensiblen Anwendungsbereichen, wie beispielsweise der Personalauswahl (Gade et al., 2020; Ochmann et al., 2020; Schumann et al., 2020). Mithilfe von KI können Bewerbungsgespräche online durchgeführt werden sowie auf Grundlage der eingereichten Lebensläufe und Antworten eine Vorauswahl der geeignetsten Bewerber*innen getroffen werden. Da KI entscheidet, ob Kandidat*innen für eine Stelle geeignet sind oder nicht, nimmt die KI einen direkten Einfluss auf das Leben eines Menschen (Yarger et al., 2019), weshalb Bewerber*innen stichhaltige Erklärungen für die getroffenen Entscheidungen erwarten (Gade et al., 2020; Kim et al., 2020).

4.2 Allgemeine Perspektive von Bewerber*innen in Österreich

Die in diesem Teil präsentierten Ergebnisse basieren auf den Daten einer im Jänner und Februar 2020 durchgeführten quotenrepräsentativen Umfrage in der österreichischen Bevölkerung.

Insgesamt wurden 190 Personen aus Österreich mit einem Durchschnittsalter von 47,0 Jahren (Standardabweichung = 15,5 Jahre) befragt. 49% waren männlich, 37,4% verfügten über mindestens einen Matura als höchsten Bildungsabschluss.

Den befragten Personen wurden eine Reihe von Fragen zur Verwendung von Chatbots im Bewerbungsprozess gestellt. Eine Übersicht dieser Fragen sowie der Anteil an befragten Personen, die den entsprechenden Aussagen zustimmten (stimme sehr zu & stimme eher zu) findet sich in Tabelle 4.1. Am meisten Zustimmung finden mit über 60% Aussagen über eine entsprechende Unterstützung zu einem Chatbot, sowohl in Form einer Ansprechperson wie auch einer Anleitung. Dazu passt, dass nur je 21%, 26% bzw. 33% davon ausgehen, dass Interaktionen mit einem Chatbot flexibel sowie klar und verständlich, sowie insgesamt einfach sind. Am wenigsten Zustimmung finden Aussagen, die auf positive Reaktionen von Freund*innen und Bekannten zur Nutzung von Chatbots im Bewerbungsprozess abzielen (13 bzw. 11%). Auch Zustimmungen dazu, dass der Einsatz von Chatbots zu einem faireren und transparenteren Bewerbungsprozess führen, findet sich nur bei zwischen 20 und 28% der Befragten. Immerhin 40% stimmen der Aussage zu, dass der Chatboteinsatz zu einem schnelleren Feedback führt.

In einem weiteren Teil der Umfrage wurde den Teilnehmer*innen ein konkreter Einsatz eines Chatbots vorgestellt. Dieser Chatbot kommuniziert mit den Teilnehmer*innen vor der Bewerbung und beantwortet Fragen sofort. Dazu wurde Abbildung 4.1 gezeigt. Im Anschluss daran wurden die Befragten gebeten, anzugeben, ob sie den in Tabelle 4.2 dargestellten Aussagen zur Verwendung eines Chatbots zur Beantwortung von Fragen zur offenen Stelle zustimmen oder nicht. Hier zeigt sich eine gespaltene Meinung. 50% geben an, dass sie darauf vertrauen, dass der Chatbot gute Antworten auf Fragen gibt; für knapp die Hälfte (48%)

beeinflusst der Einsatz eines solchen Chatbots die Entscheidung sich zu bewerben positiv und 40% halten einen solchen Einsatz generell für positiv. Allerdings würden nur 30% ihre Fragen dem Chatbot stellen, wenn dieser die Alternative zu einer menschlichen Ansprechperson wäre.

4.3 Die Perspektive von AMS Klient*innen

In diesem Abschnitt präsentieren wir schließlich Ergebnisse aus einer quotenrepräsentativen standardisierten Umfrage unter beim AMS gemeldeten Personen.

Szenarienübersicht

Den Kern der Befragung bildeten drei Szenarien zu einem Bewerbungsprozess, wobei der Entscheidungsträger in der Vorauswahl variiert wurde. Einmal war dies ein Mensch, zwei Szenarien betrafen Vorauswahlentscheidungen durch einen KI-basierten Chatbot.

Abbildung 4.4 zeigt die drei Varianten, mit denen die Szenarien eingeführt wurden. Auf der linken Seite der Abbildung findet sich die Einführung für das Szenario mit dem*der menschlichen Entscheider*in (im Weiteren wird dies mit MENSCH abgekürzt). Die rechte Seite zeigt die Einführung für das Szenario mit einem KI-basierten Chatbot mit Erklärungskomponente, der zudem für faire Entscheidungen optimiert ist (FAT_KI). Das zweite Szenario mit KI-basiertem Chatbot ist nicht abgebildet, da es bis auf den Austausch der Wörter Mitarbeiter*in bzw. Mensch identisch mit dem Szenario MENSCH ist. Da hier keine Erklärung eingesetzt wurde, entspricht dies einer Black-Box KI (BLACK_BOX_KI).

Nach dieser Einführung in das Szenario wird eine dazugehörige Entscheidung gezeigt, in der eine Ablehnung in der Vorauswahl stattfindet. In der Umfrage wurden den Klient*innen zwei der drei Szenarien vorgestellt. Eines davon war immer der menschliche Entscheider sowie einer der beiden KI-basierten Chatbots. Nach jeder der beiden gezeigten Entscheidungen wurden die Teilnehmer*innen gebeten, diese hinsichtlich der wahrgenommenen Fairness zu beurteilen.

Tabelle 4.1: Zustimmung zu Aussagen zur Verwendung von Chatbots im Bewerbungsprozess (in %)

Frage	Zustimmung in %
Es ist mir wichtig, dass ich eine Ansprechperson kontaktieren kann, wenn ich während der Interaktion mit einem Chatbot Hilfe benötige.	64,2
Es ist mir wichtig, dass es eine Anleitung zur Verwendung des Chatbots gibt.	61,1
Durch die Verwendung eines Chatbots erwarte ich ein schnelleres Feedback als in einem gewöhnlichen Bewerbungsprozess durch einen menschlichen Experten.	40,0
Es ist einfach mit einem Chatbot zu interagieren.	33,2
Durch die Verwendung eines Chatbots in einem Bewerbungsprozess, erwarte ich mir eine transparentere Entscheidung als von einem menschlichen Experten.	27,9
Durch den Einsatz eines Chatbots in einem Bewerbungsprozess, erwarte ich mir eine fairere Bewertung als von einem menschlichen Experten.	26,8
Die Interaktion mit einem Chatbot ist klar und verständlich.	26,3
Mit einem Chatbot kann ich flexibel interagieren.	21,1
Die Interaktion mit einem Chatbot ist kompliziert und schwierig zu verstehen.	20,5
Die Verwendung eines Chatbots im Bewerbungsprozess erhöht meine Chancen auf eine faire Bewertung.	20,0
Personen, die mir wichtig sind, würden mich darin bestärken, einen Chatbot zu nutzen.	13,2
Wenn ich in meinem Bekanntenkreis davon erzähle, dass ich einen Chatbot benutzt habe, würde das einen guten Eindruck machen.	10,5

Sie haben bereits vorab noch einige offene Fragen und klicken deshalb auf den Button „Noch Fragen?“. Daraufhin erscheint ein Chatbot namens „Recruiting Bot“. Dieser kommuniziert mit Ihnen per Chat und beantwortet sofort Ihre Fragen. Auf Ihrem Screen würde das folgendermaßen aussehen:

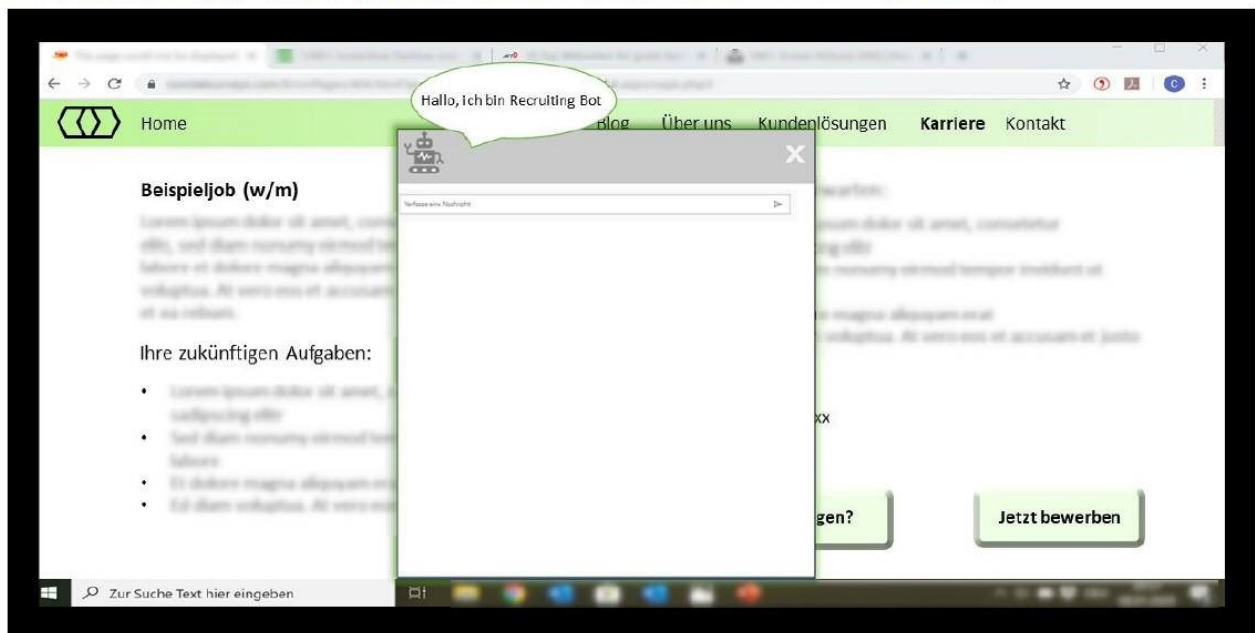


Abbildung 4.1: Beschreibung Chatbot

Tabelle 4.2: Zustimmung zu Aussagen zur Chatbot-Verwendung im Bewerbungsprozess zur Beantwortung von Fragen (in %)

Frage	Zustimmung in %
Ich vertraue darauf, dass der eben beschriebene Chatbot gute Antworten auf meine Fragen zur offenen Stelle gibt.	50,0
Der Einsatz eines Chatbots während einer Bewerbung, der meine Fragen sofort beantwortet, beeinflusst meine Entscheidung mich zu bewerben positiv.	47,9
Ich sehe diesen Einsatz von Chatbots im Bewerbungsprozess positiv.	40,0
Wenn dieser Chatbot bei einer Bewerbung als Alternative zu einer menschlichen Ansprechperson angeboten wird, dann wähle ich den Chatbot.	31,1

Bitte versetzen Sie sich in die folgende Situation:

Sie möchten sich auf eine passende freie Stelle bewerben und klicken deshalb auf den Button „Jetzt bewerben“. Daraufhin erscheint ein **Chatfenster in dem ein Mensch** die Vorauswahl der Kandidaten vornimmt.

Die von diesem Mitarbeiter ausgewählten Bewerber werden dann zu einem persönlichen Gespräch eingeladen.

Der Mitarbeiter kommuniziert mit Ihnen per Chat und stellt Ihnen alle für den Bewerbungsprozess notwendigen Fragen um Ihre Qualifikationen und Ihre Eignung für die Stelle zu bewerten. Auf Ihrem Bildschirm würde das folgendermaßen aussehen:

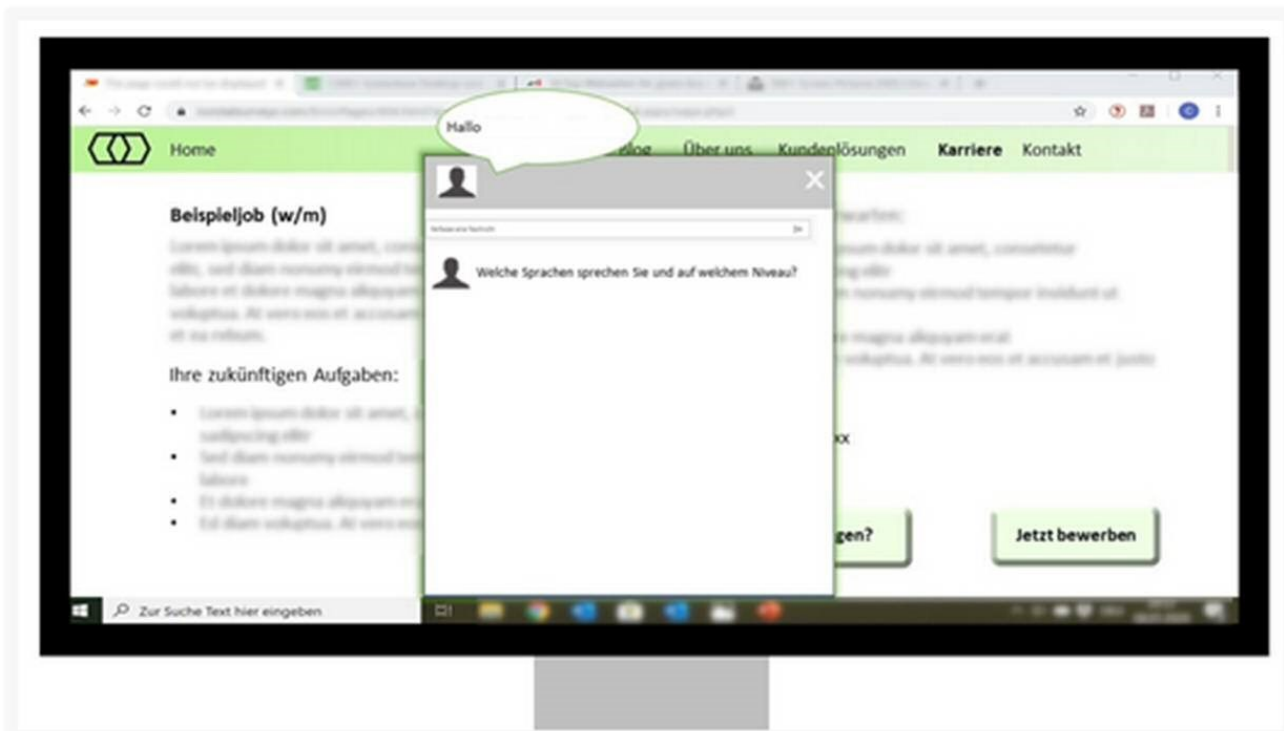


Abbildung 4.2: Szenarienübersicht 1

Die konkrete Situation sieht folgendermaßen aus: Sie möchten sich für die freie Stelle bewerben und klicken deshalb auf den Button „Jetzt bewerben“. Daraufhin erscheint ein Chatbot namens „Recruiting Bot“. Dieser übernimmt anstelle eines Menschen die Vorauswahl der Kandidaten.

Die vom Chatbot ausgewählten Bewerber werden dann zu einem persönlichen Gespräch eingeladen.

Der Chatbot kommuniziert mit Ihnen per Chat und stellt Ihnen alle für den Bewerbungsprozess notwendigen Fragen um Ihre Qualifikationen und Ihre Eignung für die Stelle zu bewerten. Auf Ihrem Bildschirm würde das folgendermaßen aussehen:

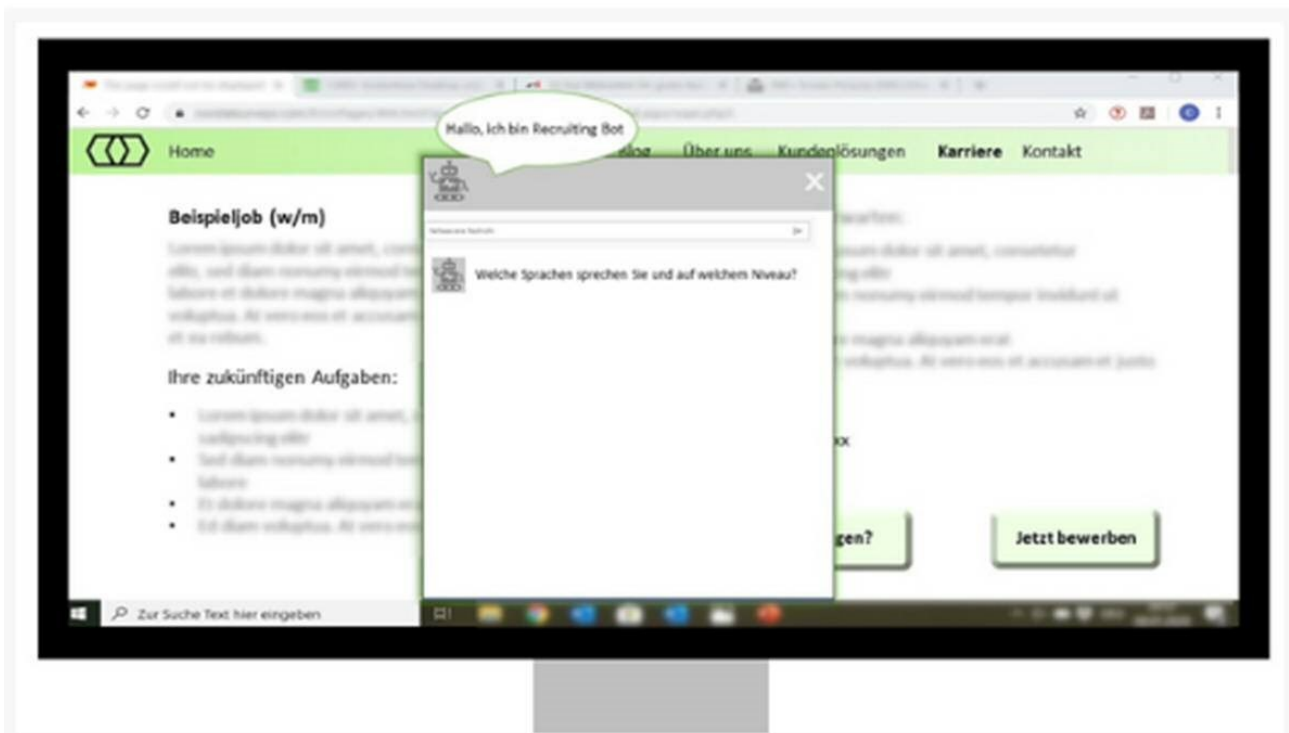


Abbildung 4.3: Szenarienübersicht 2

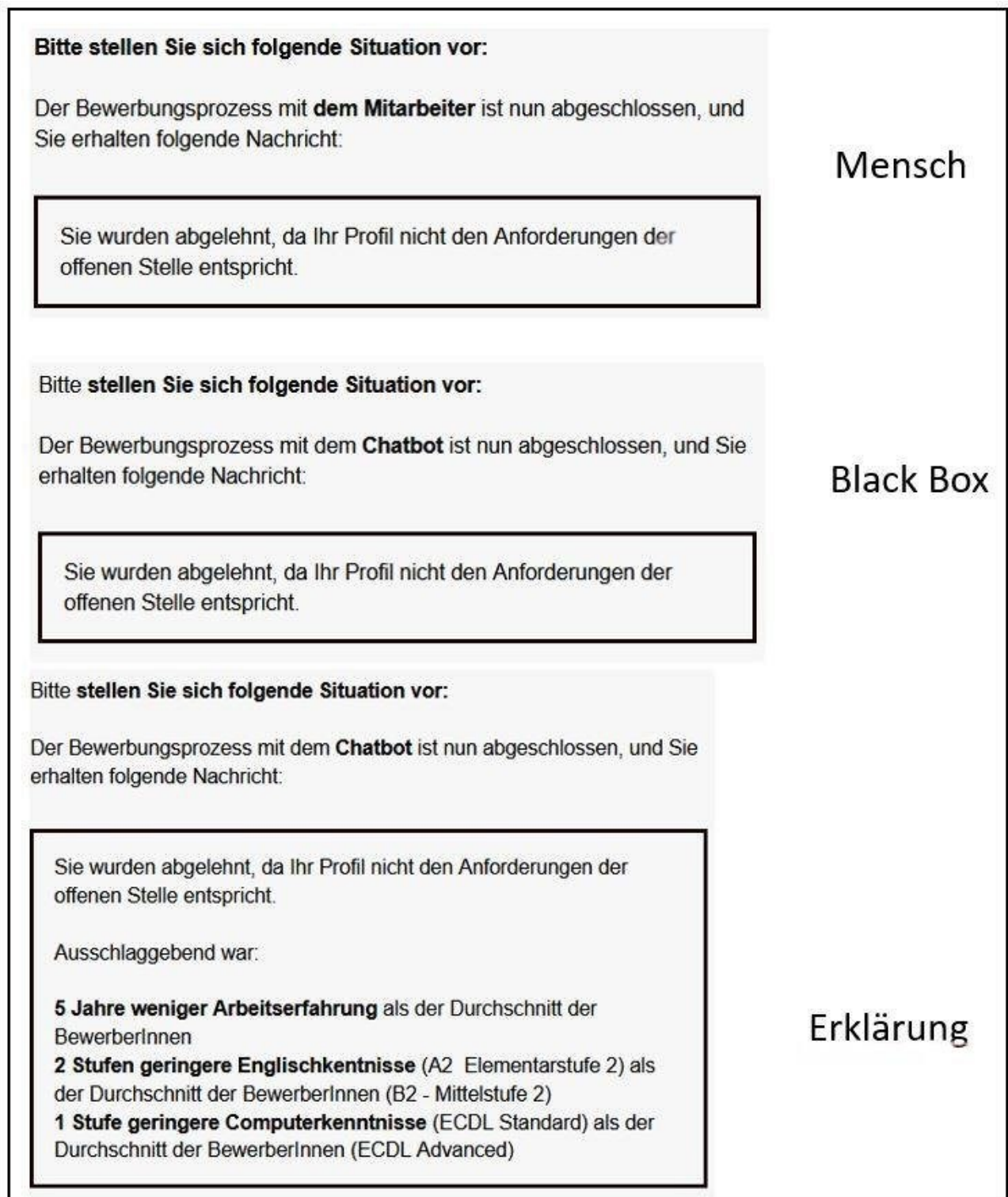


Abbildung 4.4: Szenarienübersicht

Durchführung der Umfrage und Frageblöcke

Die Online-Umfrage wurde in der Software Unipark umgesetzt und durch das AMS an insgesamt 15000 Personen in 3 Wellen ausgesendet. Insgesamt 1336 der kontaktierten Personen beantworteten den Fragebogen, was einer Rücklaufquote von 8,9% entspricht.¹

Unter den beantworteten Fragebögen waren 52,2% von weiblichen Klientinnen, 40,3% verfügten über (mindestens) eine Matura und 82,3% über die österreichische Staatsbürgerschaft. 27,1% fielen in die Altersgruppe von 18-35, 51,7% in die Altersgruppe von 36-55 und die verbleibenden 21,2% waren 56 Jahre oder älter. Abbildungen 4.5 und 4.6 veranschaulichen die Alters- und Ausbildungsverteilung grafisch.

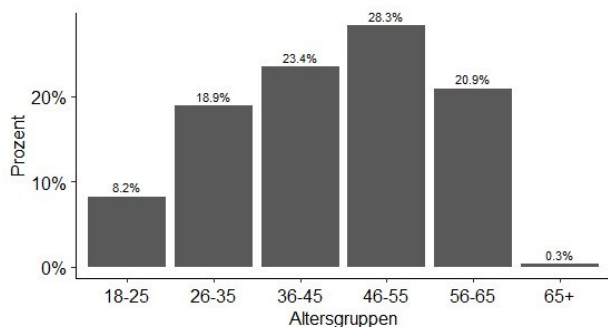


Abbildung 4.5: Verteilung der Altersgruppen

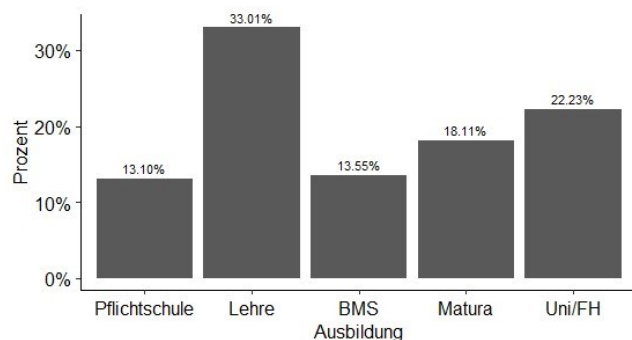


Abbildung 4.6: Verteilung der höchsten abgeschlossenen Ausbildung

Messung der Fairnesseinschätzung der Ablehnungsentscheidung und wahrgenommener Alltagsdiskriminierung

Zur Untersuchung, wie sich die verschiedenen Szenarien sowie die wahrgenommene Diskriminierung im Alltag auf die Fairnesseinschätzung der KI-Entscheidungen auswirken, wurde nach jedem Szenario die Einschätzung der Fairness mit den fünf in Tabelle 4.4 abgefragt, adaptiert von Teilfragen der deutschen Variante der Selection Procedural Justice Scale (SPJS; Bauer et al. 2001) von Konradt et al. Alltagsdiskriminierung wurde mit auf Deutsch übersetzten Fragen von Williams et al. (1997) gemessen, dargestellt in Tabelle 4.3.

Ergebnisse

Die in Tabelle 4.4 angeführten Anteile der Zustimmung zeigen, dass für alle Vertrauenss Aussagen das Mitarbeiter*innen-Szenario durchwegs die höchsten Zustimmungsraten hat. Deutlich abfallend dahinter finden sich die Raten des FAT-KI Szenarios und nochmals niedriger sind jene des Black Box KI-Szenarios.

Die Mittelwerte der Fairnessskala (Mittelwert aus allen Fragen in Tabelle 4.4) werden zur Beantwortung der der ersten der beiden Hauptfragen betrachtet: Gibt es **Unterschiede bei der Fairnessbeurteilung** der Entscheidungen in den Mitarbeiter*innen-, Black Box KI und FAT KI-Szenarien. Abbildung 4.7 zeigt die Mittelwerte der Fairnessbeurteilungen. Wir sehen den niedrigsten Wert von 3.2 für das Black Box Szenario und den höchsten Wert von 4.2 für das Mitarbeiterszenario, welches damit als am fairsten beurteilt wird. Dazwischen findet sich mit einer Fairnessbeurteilung von 3.7 das FAT KI-Szenario. Paarweise Vergleiche der Fairnessbeurteilungen in den Szenarien zeigen, dass sich diese in allen Szenarien voneinander statistisch signifikant unterscheiden (t-tests, alle p-Werte < 0.05 mit Bonferroni Korrektur für 3 paarweise Tests).

¹ Hier ist zu berücksichtigen, dass wenige Stunden nach der Aussendung der zweiten Welle das Umfragesystem für 2 Tage nicht erreichbar war und der Rücklauf dementsprechend niedriger war.

Tabelle 4.3: Wahrgenommene Häufigkeit verschiedener Alltagsdiskriminierungen (in%)

Frage	Fast jeden Tag	Mind. einmal pro Woche	Mehrm. im Monat	Mehrm. im Jahr	Weniger als einmal im Jahr	Nie
Ich werde schlechter behandelt als andere Menschen.	3,1	4,4	6,2	19,1	29,3	37,9
Es wirkt oft, als ob Menschen vor mir Angst hätten.	2,8	1,9	3,0	12,4	20,3	59,6
Menschen tun so, als wäre ich nicht besonders intelligent.	2,2	3,6	2,4	10,9	24,3	56,6
Ich erhalte eine schlechtere Bedienung als andere Menschen in Restaurants oder Geschäften.	1,3	1,9	2,6	7,9	26,2	60,0
Ich werde bedroht oder belästigt.	1,0	1,1	1,9	5,8	16,4	73,7

Tabelle 4.4: Zustimmung zu Vertrauensfragen der Chatbotentscheidung in den drei Szenarien (in %)

Frage	Mensch	FAT KI	Black Box KI
Ich vertraue dem Mitarbeiter/Chatbot, dass er meine Angaben im Bewerbungsgespräch korrekt einstuft.	62,4	26,8	17,4
Der Mitarbeiter/Chatbot trifft eine faire Auswahl.	49,0	22,1	13,4
Auch wenn jemand die Fähigkeiten des Mitarbeiters/Chatbots in der Personalauswahl anzweifelt, würde ich dem Mitarbeiter/Chatbot vertrauen.	42,7	16,2	10,0
Wenn ich entscheiden könnte, würde ich den Mitarbeiter/Chatbot nicht meine Eignung für eine wichtige Arbeitsstelle einstufen lassen.	33,5	12,3	9,4
Ich wünschte, ich hätte eine Möglichkeit, die Auswahlentscheidungen des Mitarbeiters/Chatbots zu kontrollieren.	15,2	9,0	6,0

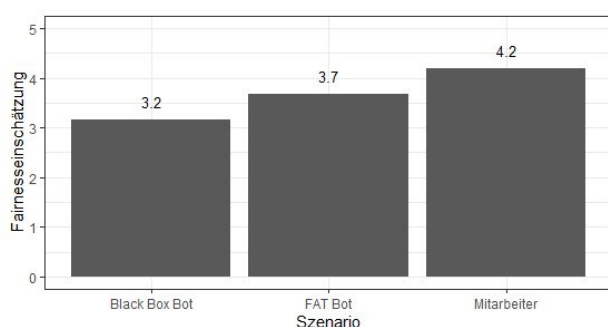


Abbildung 4.7: Fairnessbewertung in den Szenarien

Wenn also auch die Entscheidung durch einen Menschen durchschnittlich als am fairsten beurteilt wird, so ist durch die Verwendung von transparenten KI-Systemen mit Erklärungskomponenten sowie einer auf Nicht-Diskriminierung ausgelegten und kommunizierten Gestaltung möglich, die Lücke gegenüber der Fairnessbeurteilung von intransparenten Black Box KI-Systemen zumindest teilweise zu schließen.

Im nächsten Schritt, wird der **Einfluss wahrgenommener Diskriminierung auf die Fairnesseinschätzung** von KI-Entscheidungen betrachtet. Insgesamt zeigt sich bei den Angaben zur Häufigkeit der im Alltag wahrgenommenen Diskriminierung, dass für alle Bereiche bis auf einen (die allgemeine Aussage "Ich werde schlechter behandelt als andere Menschen") mehr als die Hälfte der Befragten angeben, dass sie noch nie Diskriminierung wahrgenommen haben. Mindestens mehrmals pro Monat nehmen je etwa 10% der Befragten Diskriminierung wahr. Zur Untersuchung des Zusammenhangs zwischen dieser wahrgenommenen Diskriminierung mit der Fairnesseinschätzung

zung in den verschiedenen Szenarien wird der Mittelwert aus allen Diskriminierungsfragen gebildet.² Anhand dieser Skala werden die Teilnehmer*innen nach deren wahrgenommener Diskriminierung in vier Gruppen eingeteilt; die Gruppe mit Werten von 1 bis 1,2 auf der Diskriminierungsskala ist jene, die keine Diskriminierung wahrnimmt; dies sind 36,6% der Befragten. Die Gruppe von 2 bis 6 jene mit der vergleichsweise stärksten Diskriminierungswahrnehmung, was auf 21,9% der Befragten zutrifft. Die beiden mittleren Gruppen weisen dementsprechend eine eher niedrige (1,2 bis 1,6; 22,0%) bzw. eher starke (1,6-2; 19,5%) wahrgenommene Diskriminierung auf.

Abbildung 4.8 zeigt die Mittelwerte der Fairnesseinschätzung der Entscheidungen dieser vier Gruppen jeweils einzeln für die drei Szenarien. Es zeigt sich, dass mit zunehmender wahrgenommener Diskriminierung die Fairnesseinschätzungen sowohl des FAT KI Szenarios wie auch des Black Box KI Szenarios steigen. Demgegenüber sinkt die Fairnesseinschätzung der Entscheidung des Mitarbeiters bzw. der Mitarbeiterin. In Summe nähern sich die Mittelwerte mit zunehmender wahrgenommener Diskriminierung an, die Reihenfolge der Fairnesseinschätzung – Mensch vor FAT KI und diese vor Black Box KI – bleibt aber immer erhalten. Bivariate lineare Regressionsmodelle mit der Fairnesseinschätzung als abhängige Variable und der (hier nicht kategorisierten) wahrgenommenen Diskriminierung

als unabhängige Variable bestätigen die in Abbildung 4.8 ersichtlichen Zusammenhänge: Ein Modell, welches ausschließlich Daten für das Szenario Mitarbeiter*in enthält zeigt einen signifikanten negativen Regressionskoeffizienten von $-0,07$ ($p=0,01$) für die wahrgenommene Diskriminierung. Umgekehrt zeigen Modelle basierend auf Daten für lediglich das Black Box KI Szenario, bzw. das FAT KI Szenario, signifikante positive Regressionskoeffizienten von $0,21$ bzw. $0,12$ ($p<0,01$ bzw. $p=0,02$).

Somit zeigt sich, dass eine wahrgenommene Diskriminierung mit einer höheren Fairnesseinschätzung der Entscheidungen von sowohl Black Box KI Systemen als auch FAT KI Systemen einhergehen, während umgekehrt mit zunehmender wahrgenommener Diskriminierung die Fairnesseinschätzung der menschlichen Entscheidung, wenn auch schwächer, abnimmt.

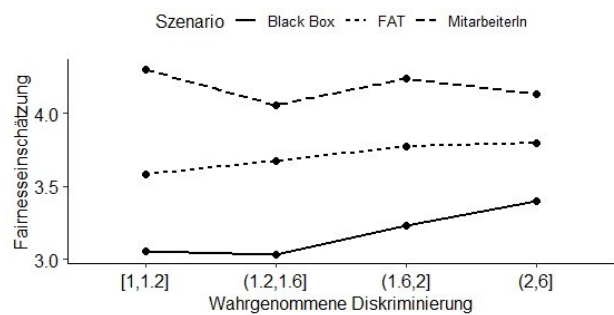


Abbildung 4.8: Einfluss von wahrgenommener Diskriminierung auf die Fairnessbewertung in den Szenarien

² Die Skala wird dabei als ordinal behandelt.

5 Dashboarddesign

Um zu untersuchen, wie KI Personalist*innen in der Personalauswahl unter Einhaltung beispielsweise rechtlicher Anforderungen unterstützen kann, werden im Folgenden (a) das Design eines prototypischen Dashboards sowie (b) dessen experimentelle Überprüfung erläutert.

5.1 Erstellung eines Dashboards

5.1.1 Definition von Anforderungen

Als ersten, vorbereitenden Schritt eines Dashboarddesigns ist es erforderlich, Anforderungen und Kriterien an ein Dashboard für die Personalauswahl zu definieren. Als Grundlage dient eine Literaturrecherche zu den Bereichen Datenschutz-Grundverordnung, Personalauswahl und Visual Analytics, die als relevant für eine rechtlich sichere und praxistaugliche Konzeption eines Dashboards angenommen werden können.

Datenschutz-Grundverordnung

Seit dem 25. Mai 2018 sind die in der Datenschutz-Grundverordnung (DSGVO) enthaltenen Maßgaben verbindlich in den jeweiligen Mitgliedstaaten der EU anzuwenden. Das Ziel der DSGVO ist es, die Verbraucher*innen zu schützen, indem die Verarbeitung personenbezogener Daten strenger reguliert wird. In Hinblick auf die Personalauswahl bzw. auf ein KI-gestütztes Dashboard zur Personalauswahl, ist im Besonderen der Artikel 22 der DSGVO interessant: „Die betroffene Person hat das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung — einschließlich Profiling — beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt“ (...). Von automatisierter Verarbeitung bei der Entscheidungsfindung kann gesprochen werden, wenn diese ohne menschliche Intervention stattfindet. Insbesondere wenn die daraus resultierende Entscheidung für die betroffene Person verbindlich ist, muss folglich sichergestellt werden, dass die Rechte der Person geschützt sind. (Brkan, 2019). Ein

Anwendungsbeispiel, in dem dieser Paragraph zu tragen kommt, ist die KI-gestützte Personalauswahl, die Einfluss auf die Entscheidung über ein Beschäftigungsverhältnis einer Person haben kann.

Die Literatur spricht im Kontext des Artikels 22 auch von der Lesbarkeit im Design, wenn Algorithmen im Entscheidungsprozess eingebunden werden. Die Lesbarkeit kann wie folgt definiert werden: „[...] making data and analytics algorithms both transparent and comprehensible to the people the data and processing concerns“ (Mortier u. a., 2014). Malgieri und Comandé (2017) verdeutlichen diese Definition noch weiter und führen aus: „legibility of data and analytics algorithms is a concept able to combine comprehensibility of the functioning of the algorithm [...] with transparency about the commercial use of that algorithm [...] in an effective way.“ Bei den Darstellungen des Outputs des Algorithmus in einem Dashboard geht es um die Implementierung, also den „commercial use of that algorithm“. In diesem spezifischen Fall kann besonders eine Maßnahme als relevant bezeichnet werden, da sie sich direkt auf die Gestaltung und Design eines Dashboards für die Personalauswahl auswirkt (Malgieri und Comandé, 2017): „The right to obtain human intervention“ – Wodurch und wie können Personalist*innen ermächtigt werden, auf die Entscheidung Einfluss zu nehmen.

Dabei muss auf die folgenden Anforderungen beim Design des Dashboards geachtet werden, damit dem Recht auf menschliche Intervention Folge geleistet wird und Personalist*innen Einfluss auf die Entscheidung haben (Malgieri und Comandé, 2017; Mendoza und Bygrave, 2017):

1. Das Dashboard muss menschliche Intervention für die Vorauswahl ermöglichen.
2. Diese Intervention muss als relevant angesehen werden und somit die Funktion der menschlichen Personalist*innen widerspiegeln.
3. Das Design des Dashboards sollte Informationen enthalten, damit das Urteilsvermögen und die Erfahrung der Personalist*innen für eine Personalentscheidung genutzt werden können.

Personalauswahl

Wenn man die drei Anforderungen, welche sich aus der DSGVO ergeben genauer betrachtet, erkennt man eine enge Verknüpfung zum praktischen Bereich der Personalauswahl. Das Dashboard muss einerseits die relevanten Daten für die Personalentscheidung darstellen und andererseits auch Raum für die Funktion der Personalist*innen bieten. Eine Sichtung und eine Auswahl von Bewerber*innen wie im traditionellen Personalauswahlprozess, muss also auch über ein Dashboard für Personalist*innen ermöglicht werden. Um den Prozess der Sichtung in einem Dashboard abzubilden, muss Personalist*innen die Möglichkeit geboten werden, sich mit den Bewerbungsunterlagen auseinander zu setzen. Somit kann als erste Anforderungen im Bereich der Personalauswahl festgehalten werden, dass es essentiell ist, einen Abgleich zwischen Bewerber*inneninformationen und Anforderungsprofil innerhalb des Dashboards zu ermöglichen. Dies dient nicht nur dem Zweck der Vorauswahl von Bewerber*innen, sondern ist auch im Sinne der DSGVO, da dies der Funktion der Personalist*innen entspricht. Welche Informationen ein solches Anforderungsprofil beinhaltet und wie detailliert dies dargestellt wird, variiert im konkreten Fall. Um jedoch einen Ausgangspunkt für die Konzeption des Dashboards zu schaffen, können die von Schulz (2014) beschriebenen Faktoren des Anforderungsprofils herangezogen werden:

- Erfolgskritische Faktoren – Funktionsaufgaben und Fachkenntnisse
- Erfolgsrelevante Anforderungskriterien – Hard Skills wie Sprachen und Fachkenntnisse, die sich aus dem Berufsbild und Berufserfahrung ergeben. Soft Skills, um die notwendigen persönlichen und sozialen Kompetenzen zu beschreiben.
- Anforderungen an die Persönlichkeit – Persönlichkeitsausprägungen und Verhaltensweisen

Der zweite Prozess, welcher als relevant eingestuft werden kann, ist die Auswahl. In diesem Zusammenhang kann das Konzept der Berufseignung angeführt werden, denn dieses soll die Wahrscheinlichkeit beschreiben, inwieweit eine Person auf Basis der definierten Indikatoren, auf einen Job passt (Schuler, 2013). Zu diesem Zweck werden Schlüsselindikatoren

definiert, welche zum Abgleich dienen. Die grundlegenden Indikatoren bestehen aus: Ausbildung, Fähigkeiten und Fertigkeiten (Hunter, Cushenbery und Friedrich, 2012). Die zweite Anforderung aus dem Bereich der Personalauswahl ist somit die Integration und Definition von Schlüsselindikatoren, die die Personalist*innen ermächtigen und helfen Kandidat*innen zu bewerten. Eine Personalauswahl auf Basis dieser Bewertung ist somit relevant und stellt eine funktionsübliche Intervention der Personalist*innen dar.

Aus dem Bereich Personalauswahl können zwei zentrale Anforderungen für das Dashboard formuliert werden:

1. Den Sichtungs-Prozess integrieren, durch einen Abgleich zwischen Bewerber*innen- und Anforderungsprofil.
2. Die Auswahl und Bewertung von Bewerber*innen durch die Integration von Schlüsselindikatoren, die Rückschlüsse auf die Berufseignung zulassen.

Visual Analytics

Eines der zentralen Anforderungen an die Visualisierung sind Designstandards. Diese können als grundlegende Regeln verstanden werden, welche sich durch die verschiedenen Ebenen und Darstellungen des Dashboards ziehen. Die entscheidenden Designstandards werden nun vorgestellt (Sosulski, 2018):

- Kontext – Damit das Dargestellte von den Nutzer*innen des Dashboards verstanden wird. Um dies zu gewährleisten ist der Einsatz von Legenden oder Beschriftungen notwendig, um den Kontext für eine Darstellung zu schaffen (Sosulski, 2018; Wilke, 2019).
- Farbe – Die grundlegende Aussage dieses Design Standards ist es, dass Farbe bei der Visualisierung von Daten sparsam verwendet werden sollte und lediglich zum Einsatz kommt, wenn beispielsweise eine Differenzierung zwischen zwei Kategorien vorgenommen wird. Es muss daher darauf geachtet werden, dass Farbe nicht als dekoratives Element eingesetzt wird (Sosulski, 2018; Yigitbasiglu und Velcu, 2012).
- Lesbarkeit – Um die Lesbarkeit des Dashboards hoch zu halten, ist es notwendig auf die Schriftart,

-größe,-richtung und -farbe zu achten. Dabei sollte sich das Dashboard auf eine Schriftart beschränken (*Dashboard Design Fundamentals - The Definitive Guide to Dashboard Design*). Jede Beschriftung, Text oder Überschrift steht unter der Prämisse der guten Lesbarkeit für die Anwender*innen des Dashboards (Sosulski, 2018).

- Einfachheit – Jegliche Elemente die eine Darstellung schwerer lesbar oder überladen erscheinen lassen, sind zu entfernen. Solche Elemente werden auch als „Chart Junk“ bezeichnet und sollten vermieden werden (Tuftte, 1990).

Eine weitere Anforderung aus dem Bereich Visual Analytics ist das Konzept des statistischen Storytellings. Durch das Storytelling wird auch eine Interaktion mit dem Dashboard sichergestellt, indem die Daten explorativ auf mehreren Ebenen und unterschiedlichen Aggregationsformen dargestellt werden (Yau, 2013). Dies ist auch im Sinne der Anforderungen aus dem Bereich Datenschutzverordnung notwendig, damit durch die Interaktion menschliche Interventionen und Urteilsvermögen einfließen kann.

Zwei Anforderungen aus dem Bereich Visual Analytics konnten identifiziert werden:

1. Die Designstandards bei der visuellen Gestaltung beachten
2. Ein interaktives Dashboard im Kontext des statistischen Storytellings

5.1.2 Der Dashboard Prototyp

Als erstes muss das grundlegende Layout des Dashboards gestaltet werden. Wie aus den Praxisbeispielen von Dashboards ersichtlich wurde, sind diese nach Vorbild einer Website aufgebaut (*Ideal | AI Powered Screening and Matching; AI for Recruiting Software | SmartAssistant*). Sie verfügen somit über einen Header, Body (Content) und optional eine Sidebar. Aus der Anforderung des Storytellings und den Anforderungen aus der Datenschutzverordnung kann abgeleitet werden, dass mehrere Dashboardebenen konzipiert werden müssen. Erstens, um das Storytelling zu ermöglichen. Zweitens stellt die Interaktion und der Vorgang einzelner Bewerber*innen explizit intensiver zu begutachten, eine Intervention und Funktion dar, welche

von Personalentscheider*innen auch in der traditionellen Personalauswahl ausführt. Durch das Darstellen aller Informationen und Rohdaten auf verschiedenen Ebenen, kommt man der Anforderung drei der Datenschutzverordnung nach. Personalist*innen können so ihre Erfahrung und Urteilsvermögen in den Entscheidungsprozess miteinbringen. Die erste Ebene des Dashboards dient dabei der Gesamtübersicht über alle Bewerber*innen einer ausgeschriebenen Stelle. Die zweite Ebene wird das Personenprofile der Bewerber*innen darstellen. Auf einer weiteren, dritten Ebene können die vom Chatbot erhobenen Rohdaten eingesehen werden.

Abbildung 5.1 zeigt den Entwurf wie ein Dashboard zur Personalauswahl unter Berücksichtigung der Anforderungskriterien auf der ersten Ebene, also der aggregierten Gesamtübersicht aller Bewerber*innen, aussehen könnte. Die wichtigsten Punkte zur Gestaltung und Auswahl der Elemente werden nun kurz zusammengefasst:

1. Es wurde davon abgesehen, einen Matching-Score zu integrieren. Eine Personalentscheidung, die auf einem Score basiert, kann nicht als relevant bezeichnet werden (Malgieri und Comandé, 2017).
2. Stattdessen wurde zur Bewertung ein Fünf-Punkte-System, welches an ein Balkendiagramm angelehnt ist, integriert. In der digitalen Welt ist die Bewertung mit Hilfe von fünf Sternen/Punkte ein übliches Tool – daher liegt eine hohe Praxistauglichkeit und Lesbarkeit vor.
3. Es wurden die drei Schlüsselindikatoren Ausbildung, Fähigkeiten und Persönlichkeit gebildet, welche aus den Konzepten der Berufseignung und Anforderungsprofil entnommen wurden.
4. Keywords sollen personenspezifische Übereinstimmungen hervorzuheben.
5. Demographische Daten und Bewerbungsbild, um sich an den deutschsprachigen Raum anzupassen.

Abbildung 5.2 zeigt den Tab Overview in der zweiten Ebene des Dashboards. Die wichtigsten Punkte des Designs werden wiederum zusammengefasst:

1. Kontext und Konsistenz bei der Gestaltung – Aufschlüsselung der Schlüsselindikatoren aus der ersten Ebene. Hier wird versucht die Bewertung

- erklärbar zu machen, indem die Informationen, welche eingeflossen sind, dargestellt werden.
2. Text als vorrangiges Visualisierungsinstrument. Damit die Informationen in einer Art und Weise kommuniziert werden, wie Personalist*innen es aus dem traditionellen Prozess kennen.
 3. Bewerber*innenauswahl durch Personalist*innen auf Ebene Zwei, um die Interaktion mit jedem Profil durch das Design zu ermöglichen. Des Weiteren wiederum ein Fünf-Punkte-System, um Handlungsspielraum und persönliche Expertise einzubinden.
 4. Durch den niedrigeren Aggregationsgrad der Daten, können Personalist*innen ihre Erfahrung und ihr Urteilsvermögen auf jener Ebene einbringen, wo die Entscheidung getroffen wird.

Der Abgleich zwischen Bewerber*innen und Anforderungsprofil wird durch ein Spinnendiagramm visualisiert, wie dies in Abbildung 5.2 dargestellt ist. Den Personalist*innen werden hier Abweichungen oder Übereinstimmungen mit dem Anforderungsprofil grafisch dargestellt. Der Prototyp sieht dies bei allen drei Schlüsselindikatoren vor.

Die letzte und dritte Ebene des Dashboards wird über einen Button auf Ebene Zwei erreicht. Da die vorliegende Konzeption der Annahme eines Chatbots zugrunde liegt, befindet sich hier der Button „Chat Protocol“. Es könnten auch Lebensläufe, Zeugnisse, Zertifikate oder Motivationsschreiben, die Bewerber*innen im Zuge des Bewerbungsverfahrens hochgeladen haben, verlinkt sein.

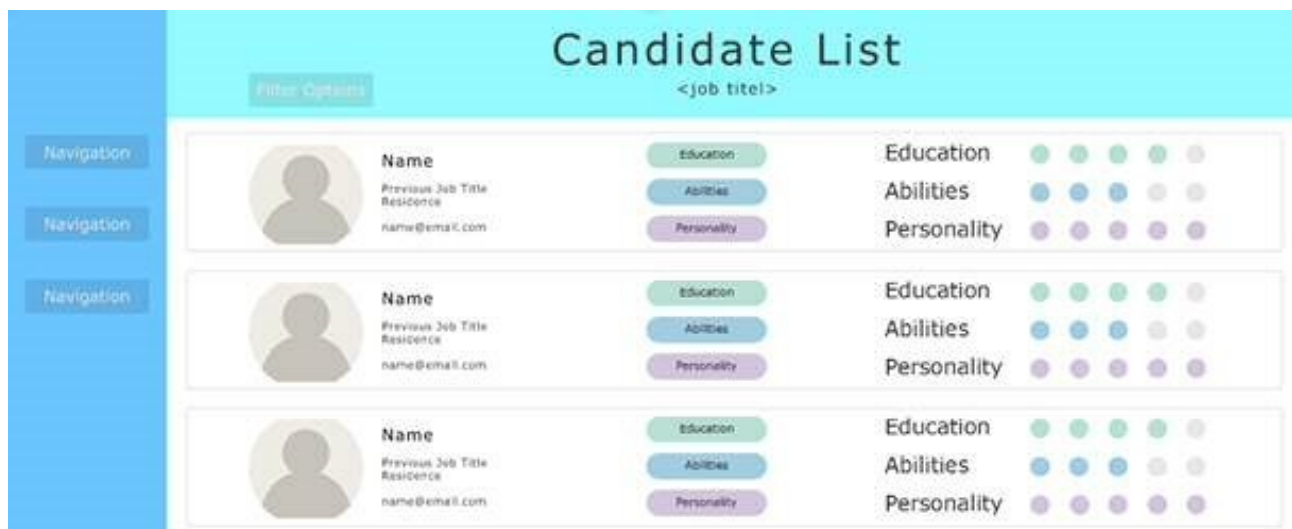


Abbildung 5.1: Bewerber*Innensicht – Ebene 1

5.1.3 Konzeptevaluation

Tabelle 5.1 zeigt eine kurze Übersicht, wie die Anforderungen umgesetzt wurden. Im Zuge der Interviewstudie mit Personalverantwortlichen wurde eine erste Version des Dashboards gezeigt. Die Aussagen in Bezug auf Darstellungsformen und Informationsgehalt sowie die Anforderungen an ein Dashboard wurden mit dem hier konzipierten Dashboard abgeglichen, um es zu evaluieren. Vergleicht man die Aussagen der Personalverantwortlichen mit dem Anforderungskatalog aus der Literaturrecherche, ergeben sich viele Schnittmengen, welche das entworfene Dashboard in seinem

grundlegenden Konzept bestärken. Dabei im Besonderen die Auswahl und Anzahl der Schlüsselindikatoren mit seinen Merkmalen, sowie die demographischen Informationen der Bewerber*innen auf der ersten Ebene. Des Weiteren die intuitive Bedienbarkeit, Übersichtlichkeit und Einfachheit in der Gestaltung. Wichtige Punkte waren überdies die Transparenz und Erklärbarkeit der Bewertung, sowie die Notwendigkeit, dass das Dashboard datenschutzkonform ist. Es gibt jedoch auch Bereiche, wo divergente Meinungen von den Personalverantwortlichen vertreten wurden, beispielsweise die Darstellung eines Bewerbungsfotos.

Auch in Bezug auf die verwendeten Darstellungsformen herrschten unterschiedliche Meinungen.

Aus diesen Implikationen kann eine letzte Designempfehlung für das Dashboard formuliert werden. Ein Dashboard für die Personalauswahl sollte bei der Wahl der Darstellungen und Merkmalausprägungen der Schlüsselindikatoren personenspezifisch und situationsbedingt angepasst werden können. Personenspezi-

fisch in Hinblick auf die Personalist*innen und das Unternehmen, welches das Dashboard nutzt. Die Merkmalausprägungen situationsbedingt in Abhängigkeit des Anforderungsprofil der zu besetzenden Stelle. Die formulierten Anforderungen bilden jedoch einen groben Rahmen und Leitfaden, an welche man sich bei der Gestaltung eines Dashboards für die Personalauswahl orientieren kann.



Abbildung 5.2: Profilübersicht – Ebene 1

Tabelle 5.1: Anforderung und Umsetzung

Anforderungen	Umsetzung
Die Auswahl und Bewertung von Bewerber*innen durch die Integration von Schlüsselindikatoren, die Rückschlüsse auf die Berufseignung zulassen.	Basierend auf den Überlappungen der Konzepte des Anforderungsprofils und der Berufseignung wurden die Indikatoren Ausbildung, Fähigkeiten und Persönlichkeit gebildet. Diese spiegeln die Hard- und Soft Skills, die Persönlichkeitsmerkmale sowie das Fachwissen der Bewerber*innen wieder.
Den Sichtungs-Prozess integrieren, durch einen Abgleich zwischen Bewerber*innen und Anforderungsprofil.	Innerhalb des Personenprofils gibt es drei Tabs für die formulierten Indikatoren. Ein Spinnendiagramm zeigt die einzelnen Indikatoren mit ihren Merkmalswerten. Die Merkmale der einzelnen Bewerber*innen werden hier mit den Anforderungen aus dem Anforderungsprofil abgeglichen. Die Personalist*innen können so auf einen Blick Übereinstimmungen oder größere Abweichungen bei den wichtigsten Merkmalen erkennen.
Ein interaktives Dashboard im Kontext des statistischen Storytellings	Statistisches Storytelling, welches einerseits die Interaktion mit dem Dashboard fördert und die Personalist*innen miteinbezieht und andererseits die Hauptstory – also die Vorauswahl als Gesamtheit. Die erste Ebene stellt dabei den Startpunkt dar und endet mit der Bewertung der Personalist*innen auf der zweiten Ebene. Die zweite Storyline besteht aus den individuellen Personenprofile, welche explorativ durch verschiedene Tabs erkundet werden können.
Die Designstandards bei der visuellen Gestaltung beachten	Bei allen Visualisierungen wurde versucht, durch Beschriftung oder farbliche Unterscheidung einen Kontext zu schaffen. Der Einsatz von Farbe wurde auf das Notwendige beschränkt. Es wurden fünf Farben gewählt, die eine gute und beruhigende Wirkung haben. Außerdem wurde eine Schriftart gewählt, die für Bildschirme geeignet ist. Im Allgemeinen wurde darauf geachtet, unnötige Grafiken oder Text zu vermeiden, um das Dashboard so einfach wie möglich zu gestalten.
Das Dashboard muss menschliche Intervention für die Vorauswahl ermöglichen.	Keine automatisierte oder vorberechnete Entscheidung durch einen Gesamtscore. Eine Intervention zur Bewertung und Vorauswahl ist daher notwendig.
Diese Intervention muss als relevant angesehen werden und somit die Funktion der menschlichen Personalist*innen widerspiegeln.	Der Intervention findet auf dem Personenprofil statt. Hier werden den Personalist*innen die für einen Auswahlprozess typischen Informationen zur Verfügung gestellt. Darüber hinaus wurde ein Abgleich zwischen den Bewerber*innen und dem Anforderungsprofil auf der zweiten Ebene ermöglicht, sowie der Zugang zu Rohdaten. Urteilsvermögen und persönliche Expertise können so einfließen, weshalb die Personalentscheidung als relevant eingestuft werden kann und der Funktion der Personalist*innen entspricht.
Das Design des Dashboards sollte Informationen enthalten, damit das Urteilsvermögen und die Erfahrung von Personalist*innen für eine Personalentscheidung genutzt werden können.	Informationen zu den drei Schlüsselindikatoren Ausbildung, Fähigkeiten und Persönlichkeit werden in Form von Text auf dem Personenprofil dargestellt. Diese Indikatoren enthalten die wichtigsten Faktoren für die Vorauswahl auf der Grundlage des Konzepts der Berufseignung und des Anforderungsprofils.

5.2 Experimentelle Untersuchung des Dashboards

Um den Einfluss des Dashboards auf die Auswahlentscheidung von Personalist*innen zu untersuchen, wurde eine experimentelle Studie mittels eines anhand der zuvor beschriebenen Richtlinien gestalteten Dashboards durchgeführt.

5.2.1 Methodisches Vorgehen

Das Dashboard wurde im Softwareportal Preely erstellt. Es stellt die Übereinstimmung der Bewerber*innen mit dem Anforderungsprofil grafisch dar. Die Versuchspersonen wurden in der Einführung zum Dashboard darüber informiert, dass ein Algorithmus die Übereinstimmung errechnet hat. Das Dashboard besteht aus drei Darstellungsebenen. Die erste Ebene des Dashboards liefert einen Überblick über alle Bewerber*innen und deren Bewertung. In der zweiten Ebene sind genauere Informationen zu den einzelnen Bewerber*innen ersichtlich. Es wird die Eignung der Bewerber*innen in Bezug auf die Schlüsselindikatoren Ausbildung, Fähigkeiten und Persönlichkeit mittels Prozentwerten und Spinnendiagrammen dargestellt. In der dritten Ebene kann der Gesprächsverlauf des Gesprächs der Bewerber*innen mit dem Chatbot eingesehen werden.

Um zu überprüfen, wie Personen mit Hilfe des Dashboards Personalauswahlentscheidungen treffen, wurden zwei Personalvorauswahlaufgaben vorgegeben. Die erste Aufgabe bezog sich auf die Vorauswahl potentieller Führungskräfte für die Marketingabteilung eines großen Unternehmens. Die zweite Aufgabe bezog sich auf potentielle Führungskräfte für die Grazer Zweigstelle einer psychosozialen Einrichtung. Die Versuchspersonen erhielten für beide Positionen zu Beginn ein Anforderungsprofil. Auf Basis dieses Anforderungsprofils und der Darstellung im Dashboard sollten sie entscheiden, welche fünf der jeweils zehn Bewerber*innen für die Position besonders gut geeignet sind

und diese fünf Personen in Abhängigkeit ihrer Eignung reihen. Bei beiden Vorauswahlaufgaben wurden in das Dashboard jeweils drei Fehler eingebaut, um zu ermitteln, wie genau die Teilnehmer*innen die im Dashboard präsentierte Vorauswahl überprüften.

Es gab sechs Experimentalgruppen, welche sich in der Einführung und der Darstellungsform unterschieden. Die Kontrollgruppe erhielt zu Beginn der Untersuchung nur grundlegende Informationen über das Dashboard und dessen Aufbau. Die Informationsgruppe erhielt zusätzlich weitere Informationen darüber, wie die künstliche Intelligenz die Passung der Bewerber*innen errechnet sowie einen expliziten Hinweis darauf, dass vor der Auswahlentscheidung alle Informationen kontrolliert werden sollten, da bei der Berechnung der Übereinstimmungswerte Fehler auftreten könnten (siehe Abbildung 5.4 oben). Die Verantwortungsgruppe erhielt die grundlegende Information und zusätzlich eine Aufklärung darüber, dass nach §22 der Datenschutz-Grundverordnung die Verpflichtung besteht, die Reihung der Bewerber*innen zu kontrollieren (siehe Abbildung 5.4 unten). Zusätzlich wurde das Verantwortungsgefühl dadurch gesteigert, dass in der Instruktion darauf hingewiesen wurde, dass eine Begründung für die getroffene Auswahl geliefert werden muss. Die Darstellungsform war bei einer Gruppe stark zusammenfassend, in Form eines allgemeinen Übereinstimmungswertes und bei der anderen Gruppe differenzierter. Es wurde untersucht, inwiefern sich die sechs Experimentalgruppen in den Variablen Vertrauenswürdigkeit, Vertrauen, Kontrollverhalten, Leistung, prozedurale Gerechtigkeit und Benutzer*innenfreundlichkeit unterscheiden.

Insgesamt nahmen 60 Personen an der Studie teil (40 Frauen, 20 Männer). Das Durchschnittsalter der Teilnehmer*innen betrug 22,6 Jahre ($SD = 3.5$). Ein Großteil der Teilnehmer*innen absolvierte zum Zeitpunkt des Experiments ein Studium (52 Personen), wobei 47 Personen Psychologie studierten. Die meisten Versuchspersonen (53 Personen) hatten keine Vorerfahrung mit der Personalauswahl.

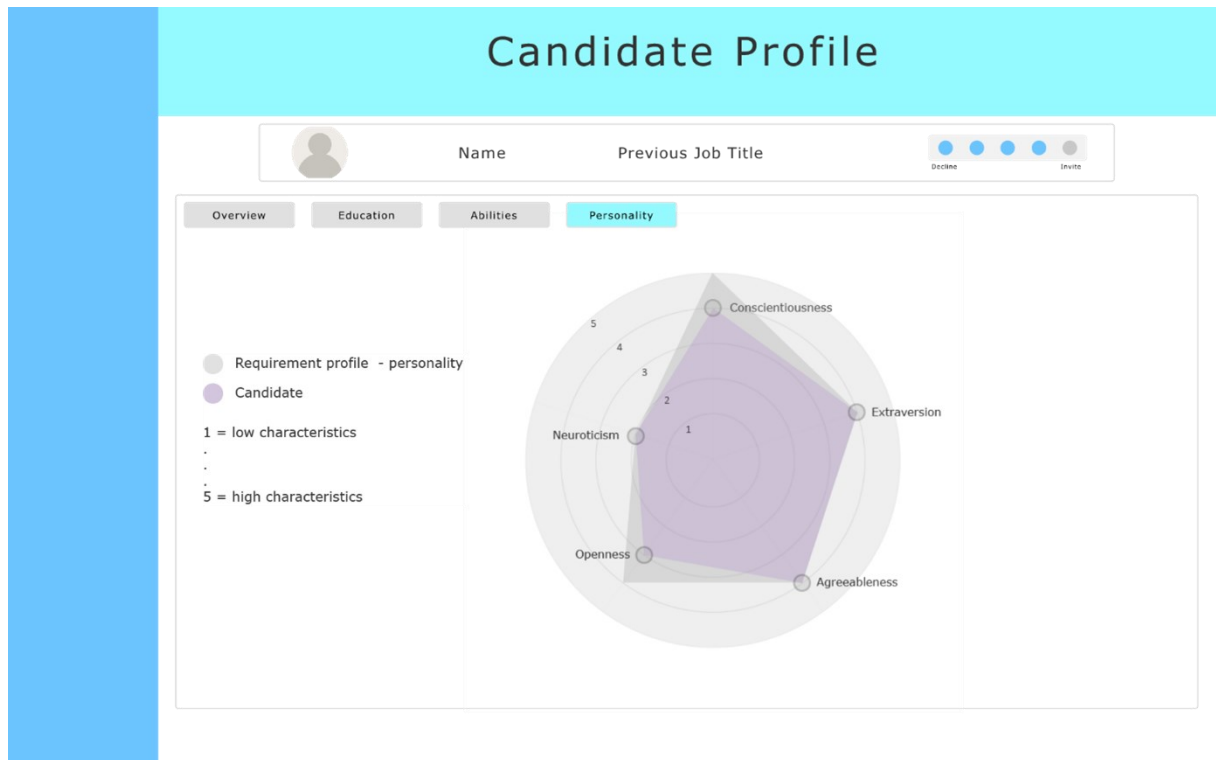


Abbildung 5.3: Gesamtübersicht – Ebene 1

Alle Angaben wurden durch eine Künstliche Intelligenz, mit Hilfe einer intelligenten Sprachanalyse, verarbeitet. Anschließend wurden allen Angaben, auf Basis eines Algorithmus, Zahlenwerten zugeordnet und diese zu den ersichtlichen Diagrammen und Summenwerten aggregiert. **Da sich in bisherigen Studien gezeigt hat, dass bei dem Einsatz von intelligenten Sprachanalysen Fehler auftreten können, ist es unerlässlich die Bewertung des Dashboards zu überprüfen.** Es sollten daher alle Informationen zu den Bewerber/innen genau betrachtet werden, bevor eine Entscheidung getroffen wird.

Bitte beachten Sie, dass das Dashboard lediglich eine Entscheidungshilfe ist und nicht die endgültige Auswahl trifft. Nach dem §22 der Datenschutz-Grundverordnung haben betroffene Personen (hier die Bewerber/innen) das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung unterworfen zu werden. **Dies bedeutet, dass Sie dazu verpflichtet sind, die Bewertung des Systems zu kontrollieren. Die Verantwortung darüber, welche Bewerber/innen zu einem persönlichen Gespräch eingeladen werden und welche nicht, liegt allein bei Ihnen.**

Abbildung 5.4: Einführungstexte der Experimentalgruppen - oben Informationsgruppe, unten Verantwortungsgruppe

5.2.2 Ergebnisse

Die Ergebnisse zeigen, dass sich Personen der Informationsgruppe besser über das Dashboard informiert fühlten und eher über Systemfehler Bescheid wussten ($M = 4.0, SD = 1.1$) als Personen der Kontrollgruppe ($M = 2.6, SD = 1.5, t(38) = -3.25, p = .002$). Erstaunlicherweise war der Unterschied zwischen der Informationsgruppe und der Verantwortungsgruppe nicht signifikant ($M = 3.6, SD = 1.3, t(38) = 0.93, p = .36$). Auch fühlten sich Personen der Verantwortungsgruppe nicht stärker für die Entscheidung verantwortlich als Personen der anderen Einführungsbedingungen ($F(2,59) = 2.35, p = .11$). Allerdings zeigt sich die Tendenz, dass Personen der Verantwortungsgruppe ihr Verantwortungsgefühl für die Entscheidung höher einschätzten ($M = 3.4, SD = 1.3$) als Personen der Informationsgruppe ($M = 2.6, SD = 1.4$) und der Kontrollgruppe ($M = 2.5, SD = 1.5$).

In beiden Personalvorauswahlaufgaben wurden die eingebauten Fehler größtenteils entdeckt. Bei der Auswahl Aufgabe für die Führungsposition in einer Marketingabteilung (erste Auswahl Aufgabe) entdeckten nur sieben Personen keinen der drei Fehler. Mehr als die Hälfte (31 Personen) entdeckten alle drei Fehler. Bei der Auswahl Aufgabe für die Führungsposition in einer psychosozialen Einrichtung (zweite Auswahl Aufgabe) verhielt es sich ähnlich. Drei Personen entdeckten keinen Fehler, während 33 Personen alle drei Fehler entdeckten (siehe Tabelle 5.2).

Tabelle 5.2: Anzahl der entdeckten Fehler in den zwei Auswahl Aufgaben

Anzahl entdeckter Fehler	Auswahl Aufgabe Marketing	Auswahl Aufgabe Psychosozial
0	7	3
1	7	8
2	15	16
3	31	33

Vertrauen und Vertrauenswürdigkeit

Vertrauen ist ein Faktor, der beeinflusst, ob und in welchem Ausmaß die Empfehlung eines Systems angenommen wird (Philipsen u. a., 2018). Daher ist eine wichtige Voraussetzung für die Etablierung eines

Dashboards in der Personalvorauswahl, dass Personalist*innen dieses als vertrauenswürdig einstufen und dem Dashboard vertrauen. Vertrauen und Vertrauenswürdigkeit sind zwei verwandte, aber dennoch voneinander abzugrenzende Konstrukte. Vertrauen beschreibt die Bereitschaft einer Person verletzlich zu sein, während die Vertrauenswürdigkeit die Bewertung einer Person über das zu vertrauende Objekt, in diesem Fall das Dashboard, beschreibt (Alarcon u. a., 2017). In Bezug auf die Verwendung eines Dashboards in der Personalauswahl bedeutet Vertrauen, die Bereitschaft der Personalist*innen bei der Entscheidungsfindung auf die Empfehlung des Dashboards einzugehen. Vertrauenswürdigkeit beschreibt dahingegen die Bewertung der Personalist*innen über die Fähigkeiten des Dashboards. Damit Personalist*innen gut mit einem Dashboard arbeiten können, ist es essenziell, dass sie dem Dashboard in angemessener Weise vertrauen und die Bewertung der Bewerber*innen durch das Dashboard als vertrauenswürdig einstufen. Daher wird untersucht, ob die Art der Einführung einen Einfluss auf die wahrgenommene Vertrauenswürdigkeit und das empfundene Vertrauen hat.

Die durchschnittliche Einschätzung der Vertrauenswürdigkeit des Dashboards lag bei 24.4 von 35 maximal erreichbaren Punkten ($SD = 4.2$). Die geringste Vertrauenswürdigkeitseinschätzung lag bei 15, die höchste bei 33 Punkten. Das durchschnittliche Vertrauen in das Dashboard lag bei 9.1 von 15 maximal möglichen Punkten ($SD = 2.2, Min = 4$ Punkte, $Max = 14$ Punkte).

Mehrere Faktoren können das Vertrauen und die Vertrauenswürdigkeit beeinflussen. Einerseits können die Eigenschaften des Systems Einfluss darauf haben, wie vertrauenswürdig eine Person ein System einstuft und inwieweit sie diesem vertraut (Balfe, Sharples und Wilson, 2018). So steht zum Beispiel die Zuverlässigkeit des Systems mit der Vertrauenswürdigkeit in Zusammenhang. Die Erfahrung mit einem verlässlichen System fördert das Vertrauen (Lee und See, 2004), während die Erfahrung mit Systemfehlern das Vertrauen verringern (Philipsen u. a., 2018). Daher kann angenommen werden, dass Personen, die Fehler des Dashboards entdecken, dieses als weniger vertrauenswürdig einstufen und dem Dashboard weniger vertrauen.

Um diese Annahme zu überprüfen wurden für die beiden Personalvorauswahlaufgaben jeweils zwei Gruppen gebildet. Die erste Gruppe umfasst alle Personen, die keinen oder nur einen Fehler entdeckten. Die zweite Gruppe umfasst alle Personen, die zwei oder drei Fehler entdeckten. Es zeigte sich, dass sich weder die Vertrauenswürdigkeit noch das Vertrauen in das Dashboard zwischen den Gruppen signifikant unterschied. Personen, die keinen oder nur einen Fehler in der ersten Auswahlaufgabe entdeckten, schätzen das Dashboard nicht signifikant vertrauenswürdiger ein ($M = 23.6, SD = 4.8$) und vertrauen diesem auch nicht signifikant mehr ($M = 8.6, SD = 1.9$) als Personen, die mehr als zwei Fehler entdeckten (Vertrauenswürdigkeit: $M = 24.6, SD = 4.0, t(58) = 0.81, p = .42$; Vertrauen: $M = 9.2, SD = 2.2, t(58) = 0.94, p = .35$). Dasselbe ergab sich für die zweite Auswahlaufgabe. Personen, die keinen oder nur einen Fehler entdeckten, schätzen das Dashboard nicht vertrauenswürdiger ein ($M = 24.7, SD = 4.6$) und vertrauen diesem auch nicht mehr ($M = 9.5, SD = 2.3$) als Personen, die zwei oder drei Fehler entdeckten (Vertrauenswürdigkeit: $M = 24.3, SD = 4.1, t(58) = -0.31, p = .76$; Vertrauen: $M = 9.0, SD = 2.2, t(58) = -0.68, p = .50$). Somit konnte die Annahme, dass Personen, die Fehler des Dashboards entdeckten, dieses als weniger vertrauenswürdig einstufen und diesem weniger vertrauen, nicht bestätigt werden.

Weitere Faktoren, die die Vertrauenswürdigkeit und das Vertrauen beeinflussen, sind Transparenz und Verstehbarkeit. Transparenz und Verstehbarkeit bedeuten, dass die Funktionsweise des Dashboards und die Berechnung der Bewertung für die Nutzer*innen nachvollziehbar und zugänglich sind. Transparenz eines Systems kann entweder durch Informationen über die Datenverarbeitung des Systems im Vorhinein oder über eine Erklärung der errechneten Entscheidung im Nachhinein erreicht werden (Felzmann u. a., 2019). Mehr Informationen über das System können somit helfen, die Nachvollziehbarkeit über die errechnete Bewertung zu erhöhen. Ein besseres Verständnis für das System beeinflusst das Vertrauen positiv (Balfe, Sharples und Wilson, 2018; Felzmann u. a., 2019; Ochmann und Laumer, 2019). Deshalb kann angenommen werden, dass Personen, welche vor der Entscheidungsfindung mehr Informationen über die Funkti-

onsweise des Dashboards erhalten, dieses als vertrauenswürdiger einstufen und dem Dashboard mehr vertrauen.

Das Vertrauen in ein Entscheidungshilfesystem hängt darüber hinaus mit der wahrgenommenen Wichtigkeit der Entscheidung zusammen. Personen bewerten ein Entscheidungshilfesystem als weniger vertrauenswürdig, wenn dieses für eine wichtige Entscheidung verwendet wird, als wenn dieses eine eher unwichtige Entscheidung trifft (Ashoori und Weisz, 2019). Es wird angenommen, dass die übertragene Verantwortung einer Entscheidung die wahrgenommene Wichtigkeit dieser Entscheidung erhöht. Daher sollten Personen, denen vor der Entscheidungsfindung die Verantwortung für die Entscheidung übertragen wurde, das Dashboard als weniger vertrauenswürdig einstufen und diesem weniger vertrauen.

Um zu überprüfen, ob die Einführung in das Dashboard die eingeschätzte Vertrauenswürdigkeit und das Vertrauen beeinflussen, wurden die Mittelwerte der Vertrauenswürdigkeit und des Vertrauens der einzelnen Einführungsgruppen miteinander verglichen. Es zeigte sich, dass sich die Einschätzung der Vertrauenswürdigkeit und des Vertrauens nicht signifikant zwischen den verschiedenen Einführungsgruppen unterschieden. Personen, welche zu Beginn mehr Informationen über das Dashboard erhielten, schätzten dieses nicht als vertrauenswürdiger ein ($M = 25.4, SD = 4.7$) und vertrauen diesem nicht mehr ($M = 9.0, SD = 2.4$) als Personen, die eine andere Einführung erhielten (Vertrauenswürdigkeit: $F(2,59) = 1.03, p = .36$; Vertrauen: $F(2,59) = 0.01, p = .99$). Personen, denen zu Beginn ihre Verantwortung über die Entscheidung bewusst gemacht wurde, unterschieden sich ebenfalls nicht signifikant in der eingeschätzten Vertrauenswürdigkeit ($M = 24.3, SD = 4.4$) und ihrem Vertrauen in das Dashboard ($M = 9.1, SD = 1.8$) von der Informationsgruppe und der Kontrollgruppe (Vertrauenswürdigkeit: $M = 23.5, SD = 3.3, F(2,59) = 1.03, p = .36$; Vertrauen: $M = 9.1, SD = 2.3, F(2,59) = 0.01, p = .99$).

Somit konnte die Annahme, dass mehr Informationen über das Dashboard die Vertrauenswürdigkeit des Dashboards und das Vertrauen erhöhen, nicht bestätigt werden. Auch konnte nicht bestätigt werden, dass

mehr Verantwortung für die Entscheidung die Vertrauenswürdigkeit und das Vertrauen in das Dashboard senken.

Kontrollverhalten

Während ein zu geringes Vertrauen mit der Nicht-Nutzung eines Systems einhergehen kann, kann durch ein zu hohes Vertrauen eine falsche Nutzung des Systems auftreten (Lee und See, 2004). Eine Folge der falschen Nutzung ist, dass Personen automatische Empfehlungen annehmen, ohne diese zu kontrollieren. Diese mangelnde Kontrolle kann dazu führen, dass mögliche Systemfehler nicht entdeckt werden (Mosier u. a., 1996). In Bezug auf den Einsatz von Dashboards in der Personalauswahl würde dies bedeuten, dass Personen eventuell falsche Bewertungen des Dashboards annehmen, ohne die dahinterliegenden Daten zu überprüfen.

Das Kontrollverhalten wurde über die Bearbeitungsdauer, die Anzahl der besuchten Seiten und die Anzahl der angesehenen Seiten erfasst. Für die Auswahlaufgaben hatten die Personen maximal 10 Minuten Zeit. Die durchschnittliche Bearbeitungsdauer für die Auswahlaufgabe der Führungsposition einer Marketingabteilung betrug 7 Minuten und 22 Sekunden (SD = 3.3, Min = 59 Sekunden, Max = 10 Minuten). Für die Position der Führungskraft einer psychosozialen Einrichtung betrug sie 6 Minuten und 53 Sekunden (SD = 2.5, Min = 31 Sekunden, Max = 10 Minuten). Im Durchschnitt wurden ungefähr 23 der 52 Seiten des Dashboards besucht (Auswahlaufgabe 1: SD = 14.6, Min = 1 Seite, Max = 50 Seiten; Auswahlaufgabe 2: SD = 12.7, Min = 1 Seite, Max = 48 Seiten). Die Anzahl der angesehenen Seiten betrug bei der Auswahlaufgabe für die Position der Führungsposition einer Marketingabteilung im Durchschnitt ungefähr 61 Seiten (SD = 43.9, Min = 1 Seite, Max = 164 Seiten). Bei der Auswahlaufgabe für die Position einer Führungskraft einer psychosozialen Einrichtung wurden im Durchschnitt ungefähr 57 Seiten (SD = 37.3, Min = 1 Seite, Max = 154 Seiten) angesehen. Dies bedeutet, dass manche Personen nur sehr wenig Zeit mit der Aufgabe verbrachten und nur wenige Informationen betrachteten, bevor sie eine Entscheidung trafen. Allerdings nahmen sich die

meisten Personen mehr Zeit und kontrollierten die Angaben, bevor sie eine Entscheidung trafen.

Mangelnde Überprüfung steht im Widerspruch zu den rechtlichen Anforderungen der DSGVO, die vorsieht, dass automatisierte Entscheidungen durch menschliche Entscheidungsträger*innen überprüft werden müssen (§22 Absatz 1). Um dieser Anforderung gerecht zu werden, muss das Kontrollverhalten der Benutzer*innen durch ein Dashboard gefördert werden. Dies bedeutet, dass Personen dazu angeregt werden sollen, sich mit der Bewertung des Dashboards stärker auseinanderzusetzen und die dargestellten Werte aktiv zu überprüfen. Faktoren, die das Kontrollverhalten erhöhen, sind unter anderem die Erfahrung mit Systemfehlern (Sauer, Chavaillaz und Wastell, 2016), Wissen über das System und die Zuverlässigkeit des Systems (Parasuraman und Manzey, 2010) sowie die empfundene Verantwortung für das Ergebnis (Skitka, Mosier und Burdick, 2000). Auf Basis dieser Befunde, ist anzunehmen, dass das Kontrollverhalten durch inhaltliche Aspekte eines Dashboards beeinflusst werden kann. Die Bereitstellung von Informationen über die Zuverlässigkeit des Dashboards sowie das Verdeutlichen möglicher Systemfehler sollte dazu führen, dass Personen die Bewertung stärker kontrollieren. Daher wird angenommen, dass Personen, welche vor der Entscheidungsfindung mehr Informationen über das Dashboard und dessen Zuverlässigkeit erhalten, mehr Kontrollverhalten aufweisen.

Des Weiteren sollte die empfundene Verantwortung über die Entscheidung das Kontrollverhalten stärken. Die empfundene Verantwortung wird dadurch gefördert, dass die Verpflichtung, die dargestellte Bewertung zu überprüfen, betont wird. Deshalb wird angenommen, dass Personen, welche vor der Entscheidungsfindung darüber informiert werden, dass die Verantwortung der Entscheidung bei ihnen liegt, die dargestellten Informationen genauer betrachten und somit mehr Kontrollverhalten aufweisen.

Um zu untersuchen, ob die Einführung eine Auswirkung auf das Kontrollverhalten hat, wurden ein Wert für das Kontrollverhalten auf Basis der oben genannten Parameter errechnet. Je mehr Zeit für die Auswahlaufgabe verwendet und je mehr Seiten vor der Entscheidung besucht und angesehen wurden, desto hö-

her ist der Wert (Kontrollverhalten = (Bearbeitungszeit beider Aufgaben + Anzahl der besuchten Seiten beider Auswahlaufgaben + Anzahl angesehenen Seiten beider Aufgaben)/100). Das Minimum des Kontrollverhalten-Wertes lag bei ungefähr 2 Punkten, das Maximum bei ungefähr 16 Punkten (M = 10.2 Punkte, SD = 3.8). Es zeigte sich ein signifikanter Unterschied im Kontrollverhalten zwischen den Einführungsgruppen ($F(2,37) = 4.36, p = .02$). Personen, die zu Beginn mehr Informationen über das Dashboard und eventuelle Verarbeitungsfehler erhielten, wiesen ein signifikant höheres Kontrollverhalten auf (M = 11.6 Punkte, SD = 2.8) als Personen der Kontrollgruppe (M = 8.0 Punkte, SD = 4.7, $t(31) = -2.95, p = .006$). Des Weiteren wiesen Personen, denen zu Beginn die Verantwortung für die Entscheidung bewusst gemacht wurde, ein höheres Kontrollverhalten auf (M = 11.0 Punkte, SD = 2.8), als Personen der Kontrollgruppe ($t(31) = -2.46, p = .02$).

Ein weiterer Faktor, der das Kontrollverhalten beeinflusst, ist der Detaillierungsgrad der Ergebnisdarstellung. Ausschlaggebend ist hierbei, wie stark die Daten zusammengefasst werden, bevor die Benutzer*innen eine Rückmeldung erhalten. Wie bei der Konzeption des Dashboards angesprochen wurde, soll das Design des Dashboards genügend Informationen enthalten, um die Fähigkeiten des/r Personalist*in auszuschöpfen. Das alleinige Darstellen eines allgemeinen Übereinstimmungswertes wird dieser Anforderung nur bedingt gerecht, da Personen ihre Bewertungsfähigkeiten bei einem Übereinstimmungswert nur begrenzt einsetzen können. Die Darstellung eines einzelnen Übereinstimmungswertes kann dazu führen, dass Personen sich mit den dahinterliegenden Daten weniger stark befassen und die Informationen oberflächlicher betrachten (Parasuraman und Manzey, 2010). Daher kann angenommen werden, dass Personen, die einen allgemeinen Übereinstimmungswert für die Bewerber*innen erhalten, ein geringeres Kontrollverhalten aufweisen.

Um diese Annahme zu überprüfen, wurden die Mittelwerte des Kontrollverhaltens der Gruppen mit und ohne allgemeinen Übereinstimmungswert miteinander verglichen. Der Unterschied zwischen den Gruppen war nicht signifikant ($t(58) = 1.02, p = .31$). Allerdings zeigte der Vergleich aller sechs Experimental-

gruppen, dass die Gruppe, welche keine spezielle Einführung in das Dashboard (Kontrollgruppe) und einen allgemeinen Übereinstimmungswert erhielt, das geringste Kontrollverhalten (M = 7.2 Punkte, SD = 5.5) und die Gruppe, die zu Beginn mehr Informationen über das Dashboard sowie eventuelle Systemfehler und keinen allgemeinen Übereinstimmungswert erhielt, das höchste Kontrollverhalten aufwies (M = 12.5 Punkte, SD = 2.7; Abbildung 5.5).

Leistung

Eine der wichtigsten Anforderungen an Entscheidungshilfesysteme ist, dass sie dabei helfen, die bestmögliche Entscheidung zu treffen. Ziel ist es durch den Einsatz von künstlicher Intelligenz und Datenvisualisierungen schnelle und gute Personalauswahlentscheidungen zu treffen (Tallgauer, Festing und Fleischmann, 2020). Obwohl bisherige Studien zeigen, dass Algorithmen zur Bewertung von Bewerber*innen reliable Ergebnisse liefern können (Campion u. a., 2016), sollte dennoch deren Auswirkungen auf die Leistung von Personalist*innen geprüft werden. Daher wurde untersucht, ob verschiedene Einführungen und unterschiedliche Darstellungsformen einen Einfluss auf die Güte der Personalvorauswahlentscheidungen und die subjektiv eingeschätzte Qualität der eigenen Entscheidungen haben.

Ein mangelndes Kontrollverhalten kann die Leistung verringern, da die Empfehlungen des Systems nicht geprüft werden. Dies bedeutet, dass die zugrunde liegenden Daten der errechneten Bewertung der Bewerber*innen nicht betrachtet und kontrolliert werden. Dadurch bleiben subtile Fehlfunktionen des Systems unerkannt und fehlerhafte Empfehlungen werden angenommen. Dies hat einen negativen Einfluss auf die Leistung, da nicht die beste Entscheidung getroffen wird (Cummings, 2004). Daher kann angenommen werden, dass das Kontrollverhalten und die Leistung in einem positiven Zusammenhang stehen.

Die objektive Leistung wurde anhand der ausgewählten Bewerber*innen und der Reihung der Bewerber*innen errechnet. Dabei wurde auch berücksichtigt, ob fälschlicherweise gut geeignete Personen nicht,

hingegen schlecht geeignete Personen schon eingeladen wurden. Die durchschnittliche Leistung war bei beiden Auswahlaufgaben 7.2 Punkte (SD = 2.1). Die schlechteste Leistung in der ersten Auswahlaufgabe war 2 Punkte und in der zweiten Auswahlaufgabe 3 Punkte, die beste Leistung war in beiden Auswahlaufgaben 10 Punkte, welches der maximal erreichbaren Punktezahl entsprach. Die selbsteingeschätzte Leistung betrug durchschnittlich 4.2 Punkte von max. 5 möglichen Punkten (SD = 0.7, Min = 2, Max = 5). Dies bedeutet, dass Personen grundsätzlich eine gute Leistung zeigten, oftmals die besten Bewerber*innen einluden und ihre Leistung selbst als gut einschätzten.

Um zu überprüfen, ob das Kontrollverhalten positiv mit der Leistung zusammenhängt, wurde der Kontrollverhalten-Wert mit den Punkten der Leistung korreliert. Die Korrelation mit der objektiven Leistung betrug $r = .39$ ($p < .01$) und mit der selbsteingeschätzten Leistung $r = .44$ ($p < .01$). Entsprechend der Annahme zeigte sich, dass Fehler eher entdeckt und besser passende Bewerber*innen eingeladen wurden je mehr Zeit verwendet, je mehr Seiten betrachtet und besucht wurden. Personen, die die Bewertung des Dashboards stärker kontrollierten, luden also auch besser geeignete Bewerber*innen ein.

Eine stärkere Auseinandersetzung mit den dargestellten Daten, sowie eine stärkere Kontrolle der dargestellten Bewertungen führt dazu, dass Systemfehler eher erkannt werden (Parasuraman und Manzey, 2010). Dadurch wird die Leistung gesteigert, da fehlerhaften Bewertungen nicht übernommen werden. Das Wissen über Systemfehler kann somit dazu führen, dass dahinterliegende Daten stärker kontrolliert werden und somit bessere Entscheidungen getroffen werden. Daher kann angenommen werden, dass Personen, welche mehr Informationen über das Dashboard erhalten und über eventuelle Systemfehler Bescheid wissen, eine bessere Leistung erzielen.

Die übertragende Verantwortung für die Entscheidung erhöht ebenfalls das Kontrollverhalten (Skitka, Mosier und Burdick, 2000). Dies bedeutet, dass Personen, welche sich ihrer Verantwortung über die Entscheidung bewusst sind, die dargestellte Bewertung stärker kontrollieren und Systemfehler eher entdecken. Deshalb kann davon ausgegangen werden, dass

Personen, denen vor der Entscheidungsfindung Verantwortung für die Entscheidung übertragen wurde, eine bessere Leistung erzielen.

Zur Überprüfung dieser Annahmen wurden die Mittelwerte der Leistung der einzelnen Einführungsgruppen miteinander verglichen. Es zeigte sich, dass die Einführungsgruppen sich in ihrer objektiven Leistung ($F(2, 59) = 7.39$, $p = .001$) und der selbsteingeschätzten Leistung signifikant voneinander unterschieden ($F(2, 34) = 4.05$, $p = .03$; Abbildung 5.6). Die Gruppe, welche zu Beginn mehr Informationen über das Dashboard erhielt, erbrachte eine signifikant bessere Leistung ($M = 16.6$, $SD = 3.5$) als Personen, denen die Verantwortung für die Entscheidung bewusst gemacht wurde ($M = 13.6$, $SD = 3.2$, $t(38) = 2.88$, $p = .007$) und Personen, die keine spezielle Einführung erhielten ($M = 13.0$, $SD = 2.9$, $t(38) = -3.60$, $p = .001$). Der Unterschied in der objektiven Leistung zwischen der Verantwortungsgruppe und der Kontrollgruppe ist nicht signifikant ($t(38) = -0.62$, $p = .54$). Jedoch ist der Unterschied in der selbsteingeschätzten Leistung zwischen diesen beiden Gruppen signifikant ($t(23) = -2.55$, $p = .02$). Personen, denen zu Beginn die Verantwortung für die Entscheidung bewusst gemacht wurde, schätzen ihre Leistung signifikant höher ein ($M = 17.2$, $SD = 1.5$) als Personen der Kontrollgruppe ($M = 14.4$, $SD = 4.8$). Personen, die mehr Informationen über das Dashboard erhielten, schätzen ihre Leistung signifikant besser ein ($M = 17.7$, $SD = 2.1$) als Personen der Kontrollgruppe ($t(26) = -2.88$, $p = .008$). Der Unterschied der selbsteingeschätzten Leistung zwischen der Verantwortungsgruppe und der Informationsgruppe ist nicht signifikant ($t(38) = 0.86$, $p = .40$).

Während Entscheidungshilfen mit einer stärkeren Zusammenfassung der Daten bei korrekter Funktionsweise mit einem Anstieg der Leistung einhergehen, hat ein Ausfall des Systems eine stärker negative Auswirkung auf die Leistung als Entscheidungshilfen, die Daten weniger stark aggregieren (Wickens u. a., 2010). In Hinblick auf ein Dashboard in der Personalauswahl bedeutet dies, dass durch die Darstellung eines einzelnen Übereinstimmungswertes Fehler weniger offensichtlich sind, seltener entdeckt werden und somit die Leistung verringert wird. Es wird angenommen, dass Personen, welche einen allgemeinen Über-

einstimmungswert für die Passung der Bewerber*innen erhalten, Fehler weniger häufig entdecken und somit eine geringere Leistung erzielen.

Die Unterschiede in der objektiven und der selbsteingeschätzten Leistung in Abhängigkeit der Darstellungsform der Ergebnisse waren nicht signifikant (objektive Leistung: $t(58) = 1.02, p = .31$; selbsteingeschätzte Leistung: $t(58) = 0.04, p = .97$). Das bedeutet, dass die Leistung nicht davon abhängt, ob Personen einen allgemeinen Übereinstimmungswert für die Passung der Bewerber*innen erhalten ($M = 13.9, SD = 3.1$) oder nicht ($M = 14.8, SD = 3.9$). Personen, denen ein allgemeiner Übereinstimmungswert angezeigt wurde, schätzten ihre Leistung nicht schlechter ein ($M = 16.4, SD = 3.7$) als Personen, denen kein Übereinstimmungswert zur Verfügung stand ($M = 16.4, SD = 3.2$).

Die Annahme, dass mehr Information über das System zu einer besseren Leistung führt, kann somit bestätigt werden, da die Informationsgruppe eine bessere Leistung erzielte und ihre Leistung besser einschätzte als die Kontrollgruppe. Allerdings kann die Annahme, dass ein gesteigertes Verantwortungsgefühl zu einer besseren Leistung führt, nur teilweise bestätigt werden, da sich nur die selbsteingeschätzte Leistung signifikant von der Kontrollgruppe unterschied. Die Annahme, dass ein allgemeiner Übereinstimmungswert einen negativen Einfluss auf die Leistung hat, kann nicht bestätigt werden, da sich weder die objektive noch die selbsteingeschätzte Leistung zwischen den Gruppen unterschied.

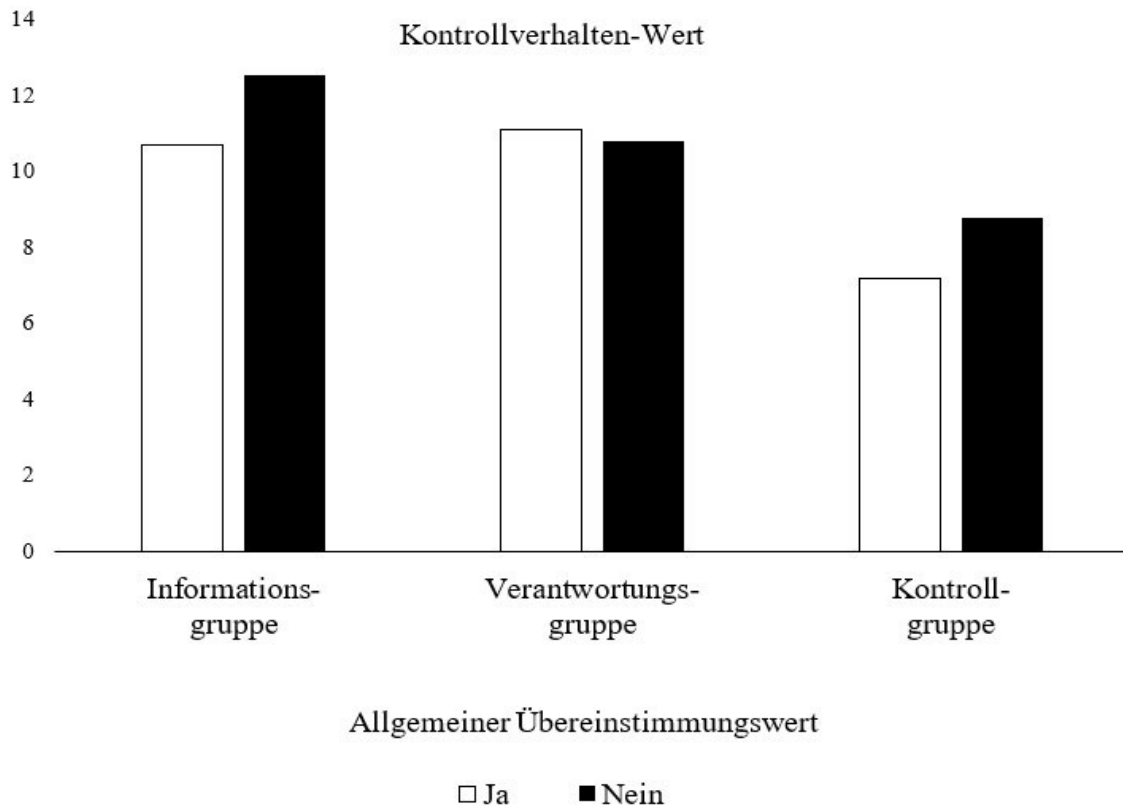


Abbildung 5.5: Kontrollverhalten der Informationsgruppe, Verantwortungsgruppe und Kontrollgruppe mit und ohne allgemeinem Übereinstimmungswert

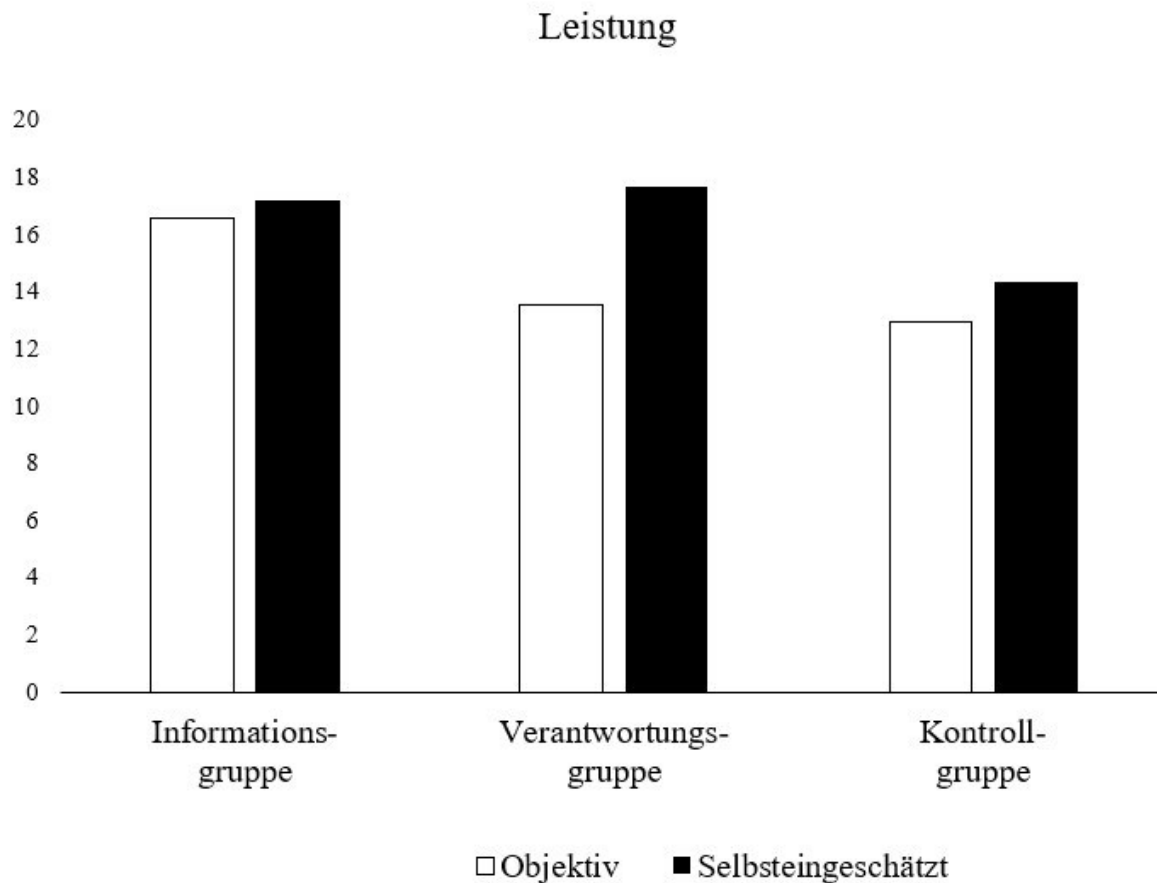


Abbildung 5.6: Objektive und selbsteingeschätzte Leistung für die Informationsgruppe, Verantwortungsgruppe und die Kontrollgruppe

Prozedurale Gerechtigkeit

Eine weitere Anforderung an Entscheidungshilfesysteme ist, dass diese zu gerechten Entscheidungen beitragen. Diese Voraussetzung ist auch in der DSGVO verankert, welche auf eine gerechte Verarbeitung von personenbezogenen Daten abzielt (Butterworth, 2018). Gerechtigkeit wird als die gleiche Behandlung aller Personen definiert. In Hinblick auf die Personalauswahl ist vor allem die prozedurale Gerechtigkeit wichtig. Unter prozeduraler Gerechtigkeit versteht man die empfundene Gerechtigkeit einer Entscheidungsfindung (Bauer u. a., 2001). Die Wahrnehmung prozeduraler Gerechtigkeit wird unter anderem dadurch beeinflusst, ob Regeln einheitlich angewendet

werden, zutreffende Informationen die Entscheidungsgrundlage bilden und die Entscheidung unvoreingenommen ist (Maier u. a., 2007). Da Auswahlverfahren durch eine künstliche Intelligenz konsistent und unvoreingenommen sein sollten, sollten sie die Standards prozeduraler Gerechtigkeit erfüllen (Ochmann und Laumer, 2019). Dennoch sollte überprüft werden, wie gerecht Benutzer*innen die Verwendung des Dashboards einschätzen. Insbesondere da auf künstlicher Intelligenz beruhende Auswahlverfahren von Bewerber*innen als ungerechter empfunden werden als von Menschen durchgeführte Verfahren (Langer, König und Papathanasiou, 2019).

Die durchschnittliche Bewertung der Gerechtigkeit des Einsatzes eines Dashboards für die Personalvor-

auswahl betrug 24.6 von 35 maximal möglichen Punkten ($SD = 4.8$, $Min = 16$, $Max = 33$). Dies bedeutet, dass Personen die Personalvorauswahl mit Hilfe eines Dashboards als ein eher gerechtes Verfahren empfinden.

Faktoren, die die empfundene prozedurale Gerechtigkeit von algorithmischen Entscheidungsprozessen beeinflussen, sind Transparenz und Verständlichkeit. Personalist*innen benötigen Informationen darüber, wie der Algorithmus eine Entscheidung trifft, um beurteilen zu können, ob das Verfahren gerecht ist. Dies beinhaltet Informationen darüber, welche Daten berücksichtigt, wie diese analysiert und gewichtet werden (Ochmann und Laumer, 2019). Daher kann angenommen werden, dass Personen, welche vor der Entscheidungsfindung mehr Informationen über das Dashboard erhalten, die Entscheidungsfindung mit Hilfe des Dashboards als gerechter empfinden.

Entgegen dieser Annahme, hatte die Einführung keinen Einfluss auf die wahrgenommene prozedurale Gerechtigkeit ($F(2,59) = 0.55$, $p = .58$). Die Einschätzung der prozeduralen Gerechtigkeit unterschied sich nicht für die Informationsgruppe ($M = 25.5$, $SD = 4.6$), die Verantwortungsgruppe ($M = 24.5$, $SD = 4.8$) und die Kontrollgruppe ($M = 23.9$, $SD = 5.2$). Dies bedeutet, dass Personen das Dashboard nicht als gerechter empfinden, wenn sie vor der ersten Benutzung mehr Informationen über die Funktionsweise des Dashboards erhalten.

Des Weiteren können verschiedene Darstellungsformen einen Einfluss auf die Gerechtigkeit haben. Die Entscheidungsfindung auf Basis eines Übereinstimmungswertes fordert nur eine geringe Auseinandersetzung mit den dahinterliegenden Daten. Im Falle einer oberflächlichen Auseinandersetzung mit den Daten sind jedoch die rechtlichen Anforderungen der Überprüfung von automatisierten Entscheidungen (DSGVO §22 Absatz 1) nicht hinreichend erfüllt (Malgieri und Comandé, 2017). Darüber hinaus verringert ein allgemeiner Übereinstimmungswert die Verständlichkeit, da weniger Information über das Zustandekommen des Wertes ersichtlich ist. Daher kann angenommen werden, dass Personen, die mit einem Dashboard mit einem allgemeinen Übereinstimmungswert arbeiten, dieses als weniger gerechter wahrnehmen.

In Übereinstimmung mit dieser Annahme unterschied sich die Gruppe, die einen allgemeinen Übereinstimmungswert erhielt, signifikant in ihrer Einschätzung der prozeduralen Gerechtigkeit von der Gruppe, welche keinen Übereinstimmungswert erhielt ($t(58) = 2.40$, $p = .02$). Personen, denen ein allgemeiner Übereinstimmungswert angezeigt wurde, bewerteten die Auswahl mit Hilfe des Dashboards als weniger gerecht ($M = 23.2$, $SD = 3.9$) als Personen, die keinen allgemeinen Übereinstimmungswert sahen ($M = 26.0$, $SD = 5.3$). Dadurch wird die Annahme bestätigt, dass die Darstellung eines allgemeinen Übereinstimmungswertes eine negative Auswirkung auf die Einschätzung der Gerechtigkeit hat.

Benutzer*innenfreundlichkeit

Ein weiterer Faktor zur erfolgreichen Etablierung eines Dashboards in der Personalvorauswahl, ist die Benutzer*innenfreundlichkeit. Unter Benutzer*innenfreundlichkeit versteht man die wahrgenommene Einfachheit der Bedienung der Technologie und die Auffassung, dass der Einsatz der Technologie bei der Arbeit problemlos funktioniert (Harrati u. a., 2016). Die Benutzer*innenfreundlichkeit ist insofern wichtig, als sie ausschlaggebend dafür ist, ob Personen ein System in ihrer Arbeit tatsächlich nutzen. Personen, die ein System als benutzer*innenfreundlich empfinden, neigen eher dazu es tatsächlich zu verwenden (Eißer, Torrini und Böhm, 2020).

Die durchschnittliche Benutzer*innenfreundlichkeit lag bei 42.7 von maximal 50 Punkten ($SD = 4.6$, $Min = 31$, $Max = 50$). Dies bedeutet, dass die meisten Personen das Dashboard als sehr benutzer*innenfreundlich einstufen.

Die bisherige Literatur hat gezeigt, dass das Wissen über ein System einen positiven Einfluss auf die wahrgenommene Benutzer*innenfreundlichkeit hat (Venkatesh und Bala, 2008). Darüber hinaus empfehlen Gestaltungsrichtlinien zur Erhöhung der Benutzer*innenfreundlichkeit, dass ein System Informationen über den Funktionsbereich und die Verlässlichkeit geben soll (Amershi u. a., 2019). Auf Grund dessen wird angenommen, dass Personen, welche vor der Entschei-

dungsfindung mehr Informationen über das Dashboard erhalten, dieses als benutzer*innenfreundlicher einschätzen.

Darüber hinaus könnte die Darstellungsform einen Einfluss auf die Benutzer*innenfreundlichkeit haben. Je leichter zugänglich Daten sind, welche für die Entscheidungsfindung wichtig sind, desto benutzer*innenfreundlicher wird ein System bewertet (Wixom und Todd, 2005). Ein allgemeiner Übereinstimmungswert könnte den Zugang zu relevanten Informationen erschweren und somit einen negativen Einfluss auf die Benutzer*innenfreundlichkeit haben. Daher wird angenommen, dass Personen, die mit einem Dashboard mit einem allgemeinen Übereinstimmungswert arbeiten, das Dashboard als weniger benutzer*innenfreundlich einstufen.

Der Vergleich der Mittelwerte ergab, dass sich weder die Einführung ($F(2,59) = 0.85, p = .43$) noch die Darstellungsform der Ergebnisse signifikant auf die Benutzer*innenfreundlichkeit des Dashboards auswirkten ($t(2) = -0.23, p = .82$). Das bedeutet, dass die eingeschätzte Benutzer*innenfreundlichkeit sowohl in der Informationsgruppe ($M = 43.3, SD = 4.9$), der Verantwortungsgruppe ($M = 43.3, SD = 3.3$) als auch der Kontrollgruppe ($M = 41.7, SD = 5.3$) ungefähr gleich hoch war. Auch war die Benutzer*innenfreundlichkeit in der Gruppe mit einem allgemeinen Übereinstimmungswert ($M = 42.9, SD = 3.9$) ungefähr gleich hoch wie in der Gruppe ohne einem allgemeinen Übereinstimmungswert ($M = 42.6, SD = 5.2$).

Dies bedeutet, dass Personen das Dashboard unabhängig von der Einführung und der Darstellungsform der Ergebnisse als sehr benutzer*innenfreundlich einstufen. Somit konnte die Annahme, dass mehr Informationen über das Dashboard und dessen Funktionsweise, die Benutzer*innenfreundlichkeit erhöhen, nicht bestätigt werden. Die Annahme, dass eine stärker zusammenfassende Darstellung in Form eines allgemeinen Übereinstimmungswertes die Benutzer*innenfreundlichkeit senkt, wurde ebenfalls nicht bestätigt.

5.2.3 Explorative Analysen

Um zu überprüfen, ob die Technikaffinität, die Gewissenhaftigkeit und die bisherige Erfahrung mit der Personalauswahl Einfluss auf die gefundenen Effekte haben, wurden diese in die Berechnungen miteinbezogen. Es zeigte sich, dass alle dargestellten Effekte bestehen bleiben, wenn die Technikaffinität, die Gewissenhaftigkeit und die bisherige Erfahrung berücksichtigt werden. Dies bedeutet, dass der Einfluss der Einführung und der Darstellungsform auf das Kontrollverhalten und die Leistung weiterhin bestehen, wenn Technikaffinität und Gewissenhaftigkeit berücksichtigt werden. Die Vorerfahrung hat keinerlei Einfluss auf Leistung, Kontrollverhalten, Vertrauenswürdigkeit, Vertrauen, Gerechtigkeit oder Benutzer*innenfreundlichkeit.

6 Anforderungskatalog und Empfehlungen

Im Folgenden werden die vorgestellten Ergebnisse in 5 Anforderungen und Empfehlungen zusammengefasst. Dabei werden die Ergebnisse aus allen Teilstudien berücksichtigt.

1. KI sollte Personalist*innen in ihren Aufgaben unterstützen und nicht automatisieren.

Die durchgeführte Literaturanalyse als auch die durchgeführten Interviews zeigen, dass sich KI nicht für alle Schritte der Personalauswahl eignet. So werden in der Literatur insbesondere die Zielgruppenansprache, das Screening sowie die (Vor-)Auswahl als vielversprechende Einsatzszenarien genannt. Auch die Interviewpartner*innen verweisen darauf, dass insbesondere der Einsatz von KI in der Vorauswahl und der Informationsbereitstellung sinnvoll erscheint. Dort kann KI die Datensammlung, Informationsbereitstellung und Entscheidungsvorbereitung unterstützen. Es zeigte sich klar, dass KI nur einen Teil des Personalauswahlprozesses unterstützen kann und das allen Beteiligten ein persönlicher Kontakt sehr wichtig ist. Wobei klar ist, dass in solchen Phasen des Personalauswahlprozesses, in denen die Informationen zahlreicher Bewerber*innen bearbeitet werden müssen, KI-Unterstützung erwünscht ist und bei den finalen Auswahlritten mit wenigen Bewerber*innen der persönliche Kontakt aufrechterhalten bleiben soll. Dabei sollte das Ziel des Einsatzes von KI in der Personalauswahl nicht eine Automatisierung und Kostenreduktion des gesamten Prozesses sein, sondern eine Entscheidungsunterstützung (Thalmann, 2018). Empfohlen wird ein gezielter Einsatz von KI um die Arbeitsbelastung von Personalist*innen in den ersten Schritten der Personalauswahl zu reduzieren und ihnen somit mehr Zeit für eine qualitativ hochwertige Personalauswahl zu geben.

2. KI sollte Bewerber*innen einen Mehrwert bieten und nicht abschrecken.

In der Literatur wird das „Chatbot Disclosure Dilemma“ dahingehend beschrieben, dass Nutzer*innen

Chatbots bzw. ihre Entscheidungen so lange sehr gut akzeptieren, wie sie nicht wissen, dass es sich um Software und nicht um einen Menschen handelt (Mozafari, Weiger und Hammerschmidt, 2020). Sollte man also nun den Chatbot enttarnen und somit möglicherweise mangelnde Bewerber*innenakzeptanz in Kauf nehmen? JA lautet hier unsere klare Antwort, denn dies stärkt das Vertrauen in den zukünftigen Arbeitgeber maßgeblich. Nicht nur die klare Kennzeichnung des Chatbots selbst wird empfohlen, sondern auch eine klare Information, welche Aufgaben der Chatbot im Bewerbungsprozess übernimmt. Darüber hinaus wird empfohlen, den Bewerber*innen die Möglichkeit einzuräumen, sich gegen die KI zu entscheiden und an einem traditionellen Prozess teilzunehmen. Mit diesen beiden Anforderungen kann der Sorge der Personalist*innen entgegengewirkt werden, dass der Einsatz von KI potentielle (und insbesondere besonders gut geeignete) Bewerber*innen abschreckt. Unsere Untersuchungen haben gezeigt, dass Bewerber*innen den Mehrwert von KI-basierten Services erkennen und diesen auch annehmen. Insbesondere wird die 24/7 Verfügbarkeit geschätzt, die Möglichkeit verschiedene Plattformen zu nutzen und hier insbesondere mobile Services und die Mehrsprachigkeit. Wie in der Literatur beschrieben, stellen auch Bewerber*innen höhere Ansprüche in Bezug auf Transparenz und Erklärung an Chatbots als an vergleichbare menschliche Akteure (Jarrah, 2018). Unternehmen, die KI einsetzen, sollten daher auf diese Anforderungen der Bewerber*innen eingehen und diese Möglichkeit anbieten.

3. KI sollte nur mit einer Erklärungskomponente eingesetzt werden.

Eine der großen Herausforderungen von KI ist deren mangelnde Transparenz in Bezug auf Entscheidungen, die auch häufig als Black-Box Charakter bezeichnet wird. In der Literatur wird Transparenz in vielen ethischen Anforderungen, wie beispielsweise denen der IEEE, ACM oder selbst dem vom Vatikan aufgestellten Katalog, gefordert. Besonders bekannt sind in diesem Zusammenhang die FAT Anforderungen, die eine *faire* also von Nutzer*innen als fair bewertete, *accountable*, also zurechnungsfähige und damit auch rechtlich gegenüber einer rechtlichen Person durchsetzbare, und

transparente, also eine in ihren Entscheidungen nachvollziehbare, KI fordern. Aus technischer Sicht wird die sich erklärende KI (oder eXplainable AI = XAI) hier als Lösungsvorschlag diskutiert und es findet derzeit sehr viel Forschung und Entwicklung in diesem Bereich statt (Thalmann u. a., 2021). Bei unseren empirischen Untersuchungen hat sich gezeigt, dass die klassische Black Box KI sowohl bei Personalist*innen als auch bei Bewerber*innen ein Akzeptanzproblem hat. Wird die KI jedoch mit einer Erklärungskomponente ausgestattet, beispielsweise auf Basis von XAI, verbessert sich die Akzeptanz deutlich. Wichtig ist hier hervorzuheben, dass dies für Bewerber*innen als auch für Personalist*innen gilt. Wir empfehlen daher für den Einsatz von KI in der Personalauswahl, dass KI, die Entscheidungen trifft (Einstufungen, Bewertungen etc.), mit einer Erklärungskomponente ausgestattet werden sollte und somit Nutzer*innen in jedem Fall Erklärungen angeboten werden.

4. KI sollte gezielt für mehr Vielfalt und gegen Diskriminierung eingesetzt werden.

In unseren Interviews zeigte sich, dass es häufig das Vorurteil gibt „KI sei diskriminierend“. Dies deckt sich auch mit einer Reihe von aktuellen Meldungen in den Medien. Technisch muss allerdings festgehalten werden, dass KI nur so gut bzw. so schlecht ist wie die Daten, mit denen sie trainiert wurde. Sind also Vorurteile in den historischen und von Menschen produzierten Daten vorhanden, dann lernt die KI diese. Allerdings gibt es eine ganz Reihe von technischen Ansätzen, mit denen Vorurteile erkannt und entfernt werden können (Corbett-Davies u. a., 2017). Darüber hinaus haben sich viele Anbieter darauf spezialisiert Vorurteile zu entfernen und bieten gezielt KI-Systeme mit dem Versprechen „Vorurteile zu reduzieren und mehr Vielfalt zuzulassen“ an. Erste Feldversuche mit solchen Systemen zeigen, dass insbesondere diskriminierte Gruppen von solchen Systemen profitieren. Es lässt sich also festhalten, dass KI maximal so diskriminierend sein kann, wie dies in den von Menschen durchgeführten Fällen, die als Trainingsdaten verwendet wurden, passiert ist. Im besten Fall können diese Vorurteile jedoch mit Hilfe von technischen Ansätzen

entfernt bzw. kompensiert werden. Unsere empirischen Untersuchungen haben darüber hinaus gezeigt, dass Personen die sich selbst diskriminiert fühlen eine positivere Einstellung als nicht-diskriminierte Personen gegenüber KI-Systemen (selbst ohne Erklärungskomponente) haben. Sollen also insbesondere Personen aus diskriminierten Gruppen angesprochen und deren Anteil in der Belegschaft erhöht werden, kann der Einsatz von KI-Systemen (die speziell für den nichtdiskriminierenden Einsatz vorbereitet wurden) durchaus sinnvoll sein. Unser Fazit ist daher, KI ist nicht per se diskriminierend und insbesondere nicht im Vergleich mit menschlichen Akteuren, sondern kann durchaus zum Abbau von Diskriminierung eingesetzt werden.

5. KI braucht Rechtssicherheit für den Einsatz

Die mangelnde Rechtssicherheit im Hinblick auf den Einsatz von KI wird sowohl in der Literatur (Königstorfer und Thalmann, 2020), als auch von den interviewten Personalist*innen als eine wichtige Barriere genannt. Auf Grund des Black-Box Charakters ist aktuell eine Auditierung oder Softwarevalidierung für KI-Systeme sehr schwierig. Besonders für den Einsatz in sensiblen Bereichen, wie dem Personalwesen, ist dies jedoch erforderlich. In unserem Designexperiment haben wir die Herausforderungen an eine rechtssichere Umsetzung gezeigt. Insbesondere die Beschränkung der automatisierten Datenverarbeitung durch §22 DSGVO ist bei der Entscheidungsunterstützung eine Herausforderung. Bei welchem Grad von Datenaufbereitung kann man von einer automatisierten Datenverarbeitung reden und wie stark muss die Prüfung durch den Menschen sein? Diese Fragen sind nur sehr schwer zu beantworten und daher ist eine Evaluation notwendig. Einzelne Hersteller bemühen sich um solche Evaluationen für Standardsoftware. Aber insbesondere für Individualsoftware oder stark angepasste Standardsoftware wird dies ein kritischer Punkt bleiben. Unsere Anforderung ist daher, dass KI auditiert bzw. evaluiert werden muss, damit der rechtliche Rahmen für Unternehmen und auch für die Personalist*innen, die diese Software einsetzen, geklärt wird.

7 Ausblick

Künstliche Intelligenz wird aktuell nur sehr vereinzelt in österreichischen Unternehmen für die Personalauswahl eingesetzt. Sowohl in der Literatur als auch von den interviewten Personalverantwortlichen werden viele Einsatzpotentiale genannt und vielfältige potenzielle Vorteile beschrieben. Es bestehen jedoch eine Reihe von Vorbehalten und Unklarheiten, die eine breite Adoption von KI in der Personalauswahl bremsen.

Eine wichtige Frage ist die Reife der Technologie und die Anwendbarkeit in der Praxis. Bisher gibt es noch nicht viele positive Beispiele und viele Unternehmen warten daher auf klare Signale. Weiters ist nicht klar, wie KI für häufig wechselnde und kleinteilige Personalauswahlprozesse eingesetzt werden kann und wie hoch der Aufwand hier ist. Aus Sicht der Studienautor*innen werden diese Best Practice Fälle jedoch in naher Zukunft zu erwarten sein und sich damit auch die Verbreitung erhöhen. Die Erwartung speist sich insbesondere aus dem zunehmenden Angebot von entsprechenden Lösungen und der Aktivität der Anbieter.

Der Black Box Charakter und die mangelnde Auditierung oder Softwarevalidierung für KI ist eine weitere Barriere für viele Unternehmen. Da es sich bei

Personalauswahlprozessen um sensible Bereiche handelt, sehen viele Unternehmen eine solche Auditierung jedoch als erforderlich an. Dieses Audit schließt auch explizit die Fragen nach Fairness und Verantwortung mit ein. Aus Sicht der Studienautor*innen würde eine solche erfolgreiche Auditierung und Validierung die Verbreitung und die Adoption von KI stark unterstützen.

Ein weiterer häufig genannter Bedenkenpunkt, der war, ob der Einsatz von KI nicht potenzielle Bewerber*innen abschreckt und insbesondere besonders geeignete Kandidat*innen. Aus unseren Untersuchungen zur Nutzer*innenperspektive hat sich gezeigt, dass Kandidat*innen den Zusatznutzen von KI-basierten Services und insbesondere von Chatbots erkennen. Allerdings sollten dabei auch Grundanforderungen der Transparenz und insbesondere die Erklärung von Entscheidungen erfüllt sein. Weiters erscheint es wichtig eine Wahlfreiheit anzubieten und den Nutzen für die Bewerber*innen klar zu kommunizieren. Aus Sicht der Studienautor*innen wird auch hier eine steigende Akzeptanz zu erwarten sein und insbesondere, wenn die Grundanforderungen an faire KI umgesetzt werden. Diesbezüglich ist zu erwarten, dass weitere Qualitätsauszeichnungen und Zertifizierungen für faire KI-Systeme auf den Markt kommen.

Literatur

- Acikgoz, Yalcin u. a. (2020). „Justice perceptions of artificial intelligence in selection“. In: *International Journal of Selection and Assessment* 28.4, S. 399–416.
- Aggarwal, Charu C und Cheng Xiang Zhai (2012). *Mining text data*. Springer Science & Business Media.
- AI for Recruiting Software / SmartAssistant. <https://www.smartrecruiters.com/recruitingsoftware/artificial-intelligence-recruitment-tools>. Accessed: 2020-08-09.
- Alarcon, Gene M u. a. (2017). „A descriptive model of computer code trustworthiness“. In: *Journal of Cognitive Engineering and Decision Making* 11.2, S. 107–121.
- Alfaro, Elena u. a. (2019). „BBVA’s Data Monetization Journey.“ In: *MIS Quarterly Executive* 18.2.
- Amershi, Saleema u. a. (2019). „Guidelines for human-AI interaction“. In: *Proceedings of the 2019 chi conference on human factors in computing systems*, S. 1–13.
- Armstrong, Michael und Stephen Taylor (2014). *Armstrong’s Handbook of Human Resource Management Practice: Edition 13*. Kogan Page.
- Ashoori, Maryam und Justin D Weisz (2019). „In AI we trust? Factors that influence trustworthiness of AI-infused decision-making processes“. In: *arXiv preprint arXiv:1912.02675*.
- Balfe, Nora, Sarah Sharples und John R Wilson (2018). „Understanding is key: An analysis of factors pertaining to trust in a real-world automation system“. In: *Human factors* 60.4, S. 477–495.
- Bauer, Talya N u. a. (2001). „Applicant reactions to selection: Development of the selection procedural justice scale (SPJS)“. In: *Personnel psychology* 54.2, S. 387–419.
- Binns, Reuben u. a. (2018). „It’s Reducing a Human Being to a Percentage“. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*. Hrsg. von Regan Mandryk u. a. New York, USA: ACM Press, S. 1–14. isbn: 9781450356206. doi: 10.1145/3173574.3173951.
- Black, J Stewart und Patrick van Esch (2020). „AI-enabled recruiting: What is it and how should a manager use it?“ In: *Business Horizons* 63.2, S. 215–226.
- (2021). „AI-enabled Recruiting in The War for Talent“. In: *Business Horizons*.
- Brkan, Maja (2019). „Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond“. In: *International journal of law and information technology* 27.2, S. 91–121.
- Brynjolfsson, Erik und ANDREW McAfee (2017). „The business of artificial intelligence“. In: *Harvard Business Review* 7, S. 3–11.
- Butterworth, Michael (2018). „The ICO and artificial intelligence: The role of fairness in the GDPR framework“. In: *Computer Law & Security Review* 34.2, S. 257–268.
- Campion, Michael C u. a. (2016). „Initial investigation into computer scoring of candidate essays for personnel selection.“ In: *Journal of Applied Psychology* 101.7, S. 958.
- Cherkassky, Vladimir und Filip M Mulier (2007). *Learning from data: concepts, theory, and methods*. John Wiley & Sons.
- Corbett-Davies, Sam u. a. (2017). „Algorithmic decision making and the cost of fairness“. In: *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, S. 797–806.
- crystalknows.com (2020). *Hiring*. url: <https://www.crystalknows.com/hiring> Zugriff am 26.04.2021.
- Cummings, Mary (2004). „Automation bias in intelligent time critical decision support systems“. In: *AIAA 1st Intelligent Systems Technical Conference*, S. 6313.
- Dashboard Design Fundamentals - The Definitive Guide to Dashboard Design*. <https://www.logianalytics.com/dashboarddesignguide/dashboard-designfundamentals>. Accessed: 2020-09-09.
- Dictionary, Merriam-Webster (2021). *Artificial Intelligence*. url: <https://www.merriam-webster.com/dictionary/artificial%20intelligence/> Zugriff am 26.04.2021.
- Dineen, Brian R, Raymond A Noe und Chongwei Wang (2004). „Perceived fairness of web-based applicant screening procedures: Weighing the rules of justice and the role of individual differences“. In: *Human Resource Management: Published in Cooperation with the School of Business Administration, The University*

- of Michigan and in alliance with the Society of Human Resources Management 43.2-3, S. 127–145.
- Di Romualdo, Anthony, Dorothee El-Khoury und Franco Girimonte (2018). „HR in the digital age: How digital technology will change HR’s organization structure, processes and roles“. In: *Strategic HR Review*.
- Eißer, Judith, Mario Torrini und Stephan Böhm (2020). „Automation Anxiety as a Barrier to Workplace Automation: An Empirical Analysis of the Example of Recruiting Chatbots in Germany“. In: *Proceedings of the 2020 on Computers and People Research Conference*, S. 47–51.
- El Ayadi, Moataz, Mohamed S Kamel und Fakhri Karray (2011). „Survey on speech emotion recognition: Features, classification schemes, and databases“. In: *Pattern recognition* 44.3, S. 572–587.
- Esch, Patrick van, J Stewart Black und Denni Arli (2020). „Job candidates’ reactions to AI-Enabled job application processes“. In: *AI and Ethics*, S. 1–12.
- Eubanks, Ben (2018). *Artificial intelligence for HR: use AI to support and develop a successful workforce*. Kogan Page Publishers.
- Feloni, Richard (2017). „Consumer goods giant Unilever has been hiring employees using brain games and artificial intelligence-and it’s a huge success“. In: *Business Insider*.
- Felzmann, Heike u. a. (2019). „Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns“. In: *Big Data & Society* 6.1, S. 2053951719860542.
- Gonzalez, Manuel F u. a. (2019). „Where’s the IO?“ Artificial intelligence and machine learning in talent management systems“. In: *Personnel Assessment and Decisions* 5.3, S. 5.
- Harrati, Nouzha u. a. (2016). „Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis“. In: *Computers in Human Behavior* 61, S. 463–471.
- Herrmann, Brigitte (2016). *Die Auswahl: wie eine neue starke Recruiting-Kultur den Unternehmenserfolg bestimmt*. John Wiley & Sons.
- Hunter, Samuel T, Liliya Cushenbery und Tamara Friedrich (2012). „Hiring an innovative workforce: A necessary yet uniquely challenging endeavor“. In: *Human resource management review* 22.4, S. 303–322.
322. *Ideal | AI Powered Screening and Matching*. [https:// ideal.com](https://ideal.com). Accessed: 2021-03-22.
- Jarrahi, Mohammad Hossein (2018). „Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making“. In: *Business Horizons* 61.4, S. 577–586.
- Königstorfer, Florian und Stefan Thalmann (2020). „Applications of Artificial Intelligence in commercial banks–A research agenda for behavioral finance“. In: *Journal of Behavioral and Experimental Finance* 27, S. 100352.
- Langer, Markus, Cornelius J König und Vivien Busch (2020). „Changing the means of managerial work: effects of automated decision support systems on personnel selection tasks“. In: *Journal of Business and Psychology*, S. 1–19.
- Langer, Markus, Cornelius J König und Maria Papathanasiou (2019). „Highly automated job interviews: Acceptance under the influence of stakes“. In: *International Journal of Selection and Assessment* 27.3, S. 217–234.
- Langer, Markus u. a. (2019). „Highly automated interviews: applicant reactions and the organizational context“. In: *Journal of Managerial Psychology*.
- Lee, John D und Katrina A See (2004). „Trust in automation: Designing for appropriate reliance“. In: *Human factors* 46.1, S. 50–80.
- Lee, Min Kyung (2018). „Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management“. In: *Big Data & Society* 5.1, S. 205395171875668. issn: 2053-9517. doi: 10.1177/2053951718756684.
- Maier, Günter W u. a. (2007). „Gerechtigkeitseinschätzungen in Organisationen“. In: *Diagnostica* 53.2, S. 97–108.
- Malgieri, Gianclaudio und Giovanni Comandé (2017). „Why a right to legibility of automated decision making exists in the general data protection regulation“. In: *International Data Privacy Law*.
- McCarthy, John (2007). *What is AI?* url: <http://jmc.stanford.edu/articles/whatisai.html>. Zugriff am 26.04.2021.
- Mendoza, Isak und Lee A. Bygrave (2017). „The Right Not to be Subject to Automated Decisions Based on Profiling“. In: *EU Internet Law: Regulation and Enforcement*. Hrsg. von Tatiana-Eleni Synodinou u. a. Cham: Springer International Publishing, S. 77–98.

- isbn: 978-3-319-64955-9. doi: 10.1007/978-3-319-64955-9_4. url: https://doi.org/10.1007/978-3-319-64955-9_4.
- Mirowska, Agata (2020). „AI Evaluation in Selection“. In: *Journal of Personnel Psychology*.
- Moor, James (2006). „The Dartmouth College artificial intelligence conference: The next fifty years“. In: *Ai Magazine* 27.4, S. 87–87.
- Mortier, Richard u. a. (2014). „Human-data interaction: The human face of the data-driven society“. In: *Available at SSRN 2508051*.
- Mosier, Kathleen L u. a. (1996). „Automation bias, accountability, and verification behaviors“. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Bd. 40. 4. SAGE Publications Sage CA: Los Angeles, CA, S. 204–208.
- Mozafari, Nika, Welf H Weiger und Maik Hamerschmidt (2020). „The Chatbot Disclosure Dilemma: Desirable and Undesirable Effects of Disclosing the Non-Human Identity of Chatbots“. In:
- Nørskov, Sladjana u. a. (2020). „Applicant Fairness Perceptions of a Robot-Mediated Job Interview: A Video Vignette-Based Experimental Survey“. In: *Frontiers in Robotics and AI* 7, S. 163.
- Ochmann, Jessica und Sven Laumer (2019). „Fairness as a determinant of AI Adoption in Recruiting: An interview-based study.“ In: *Proceedings of the DIGIT*.
- Parasuraman, Raja und Dietrich H Manzey (2010). „Complacency and bias in human use of automation: An attentional integration“. In: *Human factors* 52.3, S. 381–410.
- Philipsen, Ralf u. a. (2018). „Evaluating strategies to restore trust in decision support systems in crosscompany cooperation“. In: *International Conference on Applied Human Factors and Ergonomics*. Springer, S. 115–126.
- Pillai, Rajasshrie und Brijesh Sivathanu (2020). „Adoption of artificial intelligence (AI) for talent acquisition in IT/ITeS organizations“. In: *Benchmarking: An International Journal*.
- Rathi, R (2018). „Artificial intelligence and the future of hr practices“. In: *IJAR* 4.6, S. 113–116.
- Raub, McKenzie (2018). „Bots, bias and big data: artificial intelligence, algorithmic bias and disparate impact liability in hiring practices“. In: *Ark. L. Rev.* 71, S. 529.
- Ripley, Brian D (2007). *Pattern recognition and neural networks*. Cambridge university press.
- Russel, S und P Norvig (2016). „Artificial Intelligence: A Modern Approach Global Edition“. In: *Pearson*.
- Sauer, Juergen, Alain Chavaillaz und David Wastell (2016). „Experience of automation failures in training: effects on trust, automation bias, complacency and performance“. In: *Ergonomics* 59.6, S. 767–780.
- Schikora, Claudius, Sonia Galster und Daniela Högerl (2020). „Digitalisierung im Recruiting: Chatbots“. In: *Führen und Managen in der digitalen Transformation*. Springer, S. 265–283.
- Schuler, Heinz (2013). „Personalauswahl“. In: *Handbuch Strategisches Personalmanagement*. Hrsg. von Ruth Stock-Homburg. Wiesbaden: Springer Fachmedien Wiesbaden, S. 29–58. isbn: 978-3-658-00431-6. doi: 10.1007/978-3-658-00431-6_3. url: https://doi.org/10.1007/978-3-658-00431-6_3.
- Schulz, Ludwig M (2014). *Das Geheimnis erfolgreicher Personalbeschaffung*. Springer.
- Schumann, Candice u. a. (May 2020). „We Need Fairness and Explainability in Algorithmic Hiring“. In: *AAMAS '19: International Conference on Autonomous Agents and Multiagent Systems*, S. 1716–1720. isbn: 978-1-4503-7518-4.
- Sekhri, Alka und Dr Jagvinder Cheema (2019). „The new era of HRM: AI reinventing HRM functions“. In: *International Journal of Scientific Research and Review* 7.3.
- Skitka, Linda J, Kathleen Mosier und Mark D Burdick (2000). „Accountability and automation bias“. In: *International Journal of Human-Computer Studies* 52.4, S. 701–717.
- Sosulski, Kristen (2018). *Data visualization made simple: insights into becoming visual*. Routledge.
- Tallgauer, Maximilian, Marion Festing und Florian Fleischmann (2020). „Big Data im Recruiting“. In: *Digitalisierung im Recruiting*. Springer, S. 25–39.
- Thalman, Stefan (2018). „Data driven decision support“. In: *it-Information Technology* 60.4, S. 179–181. Thalman, Stefan u. a. (2021). „Automatisiertes Entscheiden mit Hilfe von künstlicher Intelligenz braucht Erklärung!“ In: *Smart Regulation: Vertrag, Unternehmung und Markt*. Hrsg. von Patrick Leyens, Iris Eisenberger und Rainer Niemann.

- Tufte, E.R. (1990). *The visual display of quantitative information*.
- Ullah, Robindro und Michael Witt (2018). *Praxishandbuch Recruiting: Grundlagenwissen-Prozess-Know-how Social Recruiting*. Schäffer-Poeschel.
- Upadhyay, Ashwani Kumar und Komal Khandelwal (2018). „Applying artificial intelligence: implications for recruitment“. In: *Strategic HR Review*.
- Van Esch, Patrick und J Stewart Black (2019). „Factors that influence new generation candidates to engage with and complete digital, AI-enabled recruiting“. In: *Business Horizons* 62.6, S. 729–739.
- Van Esch, Patrick, J Stewart Black und Joseph Ferolie (2019). „Marketing AI recruitment: The next phase in job application and selection“. In: *Computers in Human Behavior* 90, S. 215–222.
- Venkatesh, Viswanath und Hillol Bala (2008). „Technology acceptance model 3 and a research agenda on interventions“. In: *Decision sciences* 39.2, S. 273–315.
- Verhoeven, Tim (2020). *Digitalisierung im Recruiting. Wie sich Recruiting durch künstliche Intelligenz, Algorithmen und Bots verändert*. Springer Gabler.
- Verlinden, Neelie (2019). *9 Intriguing Uses of AI in Recruitment in 2019*. url: <https://harver.com/blog/ai-in-recruitment-2019/> Zugriff am 26.04.2021.
- Vial, Gregory (2019). „Understanding digital transformation: A review and a research agenda“. In: *The Journal of Strategic Information Systems* 28.2, S. 118–144.
- Waizenegger, Lena u. a. (2020). „An affordance perspective of team collaboration and enforced working from home during COVID-19“. In: *European Journal of Information Systems* 29.4, S. 429–442.
- Weizenbaum, Joseph (1966). „ELIZA—a computer program for the study of natural language communication between man and machine“. In: *Communications of the ACM* 9.1, S. 36–45.
- Wickens, Christopher D u. a. (2010). „Stages and levels of automation: An integrated meta-analysis“. In: *Proceedings of the human factors and ergonomics society annual meeting*. Bd. 54. 4. Sage Publications Sage CA: Los Angeles, CA, S. 389–393.
- Wilke, Claus O (2019). *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media.
- Wixom, Barbara H und Peter A Todd (2005). „A theoretical integration of user satisfaction and technology acceptance“. In: *Information systems research* 16.1, S. 85–102.
- Yau, Nathan (2013). *Data points: visualization that means something*. John Wiley & Sons.
- Yigitbasioglu, Ogan M und Oana Velcu (2012). „A review of dashboards in performance management: Implications for design and research“. In: *International Journal of Accounting Information Systems* 13.1, S. 41–59.
- Zerilli, John u. a. (2019). „Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?“ In: *Philosophy & Technology* 32.4, S. 661–683. issn: 2210-5433. doi: 10.1007/s13347-0180330-6.

8 Anhang



Künstliche Intelligenz in der Personalauswahl

Interviewleitfaden

- Kurze Vorstellung der Interviewerinnen und Beschreibung des Projekts
- Kurze Vorstellung Ihrer Person (Position im Unternehmen)
- Beschreibung des Personalauswahlprozesses in Ihrem Unternehmen
- Beschreibung bisheriger Erfahrungen von KI in der Personalauswahl
- Diskussion über den Einsatz von Conversational Agents (Chatbots) und Dashboards (graphische Datenvisualisierungen) in der Personalauswahl, sowie (eventuell) über weitere Einsatzgebiete von KI in Ihrem Unternehmen. Es sollen folgende Punkte besprochen werden:
 - Einsatzmöglichkeiten (konkrete Tätigkeiten, Stellenprofile)
 - Vorteile und Herausforderungen von KI in der Personalauswahl, sowie (eventuell) persönliche Erfahrungen
 - Veränderungen durch KI in der Personalauswahl in Bezug auf die Tätigkeit und Arbeitsbedingungen von PersonalerInnen, sowie in Bezug auf den Personalauswahl- und Bewerbungsprozess
 - Anforderungen an KI in der Personalauswahl
- Diskussion über zukünftige Entwicklungen und Trends von KI in der Personalauswahl
- Demographische Merkmale und kurze Beschreibung des Unternehmens

Abbildung 8.1: Interviewleitfaden