

DER AMS-ALGORITHMUS

Eine Soziotechnische Analyse des
Arbeitsmarktchancen-Assistenz-Systems (AMAS)

DER AMS-ALGORITHMUS

Eine Soziotechnische Analyse des
Arbeitsmarktchancen-Assistenz-Systems
(AMAS)

Endbericht

Institut für Technikfolgen-Abschätzung
der Österreichischen Akademie der Wissenschaften
Projektleitung: Doris Allhutter & Astrid Mager
Autor*innen: Doris Allhutter & Astrid Mager (ITA, ÖAW)
Florian Cech & Fabian Fischer (CIS, TU Wien)
Gabriel Grill (Univ. of Michigan at Ann Arbor)

Studie in Kooperation mit der Arbeiterkammer Oberösterreich

Wien, November 2020

IMPRESSUM

Medieninhaber:
Österreichische Akademie der Wissenschaften
Juristische Person öffentlichen Rechts (BGBl 569/1921 idF BGBl I 31/2018)
Dr. Ignaz Seipel-Platz 2, A-1010 Wien

Herausgeber:
Institut für Technikfolgen-Abschätzung (ITA)
Apostelgasse 23, A-1030 Wien
www.oeaw.ac.at/ita

Die ITA-Projektberichte erscheinen unregelmäßig und dienen der Veröffentlichung der Forschungsergebnisse des Instituts für Technikfolgen-Abschätzung. Die Berichte erscheinen in geringer Auflage im Druck und werden über das Internetportal „[epub.oeaw](http://epub.oeaw.ac.at/ita/ita-projektberichte)“ der Öffentlichkeit zur Verfügung gestellt:
epub.oeaw.ac.at/ita/ita-projektberichte

ITA-Projektbericht Nr.: 2020-02
ISSN: 1819-1320
ISSN-online: 1818-6556
epub.oeaw.ac.at/ita/ita-projektberichte/2020-02.pdf



Dieser Bericht unterliegt der Creative Commons Attribution 4.0 International License:
creativecommons.org/licenses/by/4.0/

INHALT

Danksagung.....	5
Zusammenfassung.....	7
1 AMAS aus soziotechnischer Perspektive.....	11
1.1 Studiendesign.....	13
1.2 Arbeitsmarktchancen als Optimierungsproblem	15
2 Entstehungsgeschichte von AMAS	19
3 Soziotechnische Analyse von AMAS.....	23
3.1 Technische Rekonstruktion des Systems	23
3.1.1 Das Stratifikations-Verfahren, logistische Regression und deren Implikationen.....	24
3.1.2 Variablenausprägungen und Konstellationen	28
3.1.3 Populationen und Kategorisierung im Assistenzsystem	30
3.1.4 Modellanpassungen.....	32
3.2 Identifikation kritischer Aspekte	33
3.2.1 Bias und Diskriminierung.....	33
3.2.2 Nachvollziehbarkeit und Transparenz.....	60
4 AMAS in der Beratungspraxis	63
4.1 AMAS im Kernprozess „Arbeitskräfte unterstützen“	63
4.1.1 Einführung von AMAS in die AMS Praxis.....	66
4.1.2 Implementierung der Geschäftsfunktion „Arbeitsmarktchance“	67
4.1.3 Systeminformationen und Vorgaben für AMS-Mitarbeiter*innen.....	72
4.2 Spannungen und Ambivalenzen in der AMS Praxis	75
5 Szenarien zur Illustration der Ergebnisse.....	81
6 Vergleichbare Systeme in anderen Ländern	87
6.1 Daten & Methoden	87
6.2 Profiling und Ressourcenzugang	88
6.3 Ausgewählte Ländervignetten.....	89
6.3.1 Niederlande	89
6.3.2 Polen	89
6.3.3 Schweden	90
6.3.4 Schweiz.....	91
6.3.5 Deutschland	92
6.3.6 Dänemark.....	94
6.4 Trends und Probleme.....	95
7 Schlussfolgerungen und Empfehlungen.....	97
7.1 Zielheterogenität und Ungleichheitsverhältnisse	97
7.2 Empfehlungen.....	99
7.2.1 Empfehlungen zu AMAS.....	100
7.2.2 Empfehlungen zum Einsatz algorithmischer Systeme im öffentlichen Sektor	102
7.2.3 Institutionelle Empfehlungen.....	103
Literatur.....	107
Dokumente-Verzeichnis	111

Abbildungsverzeichnis

Abbildung 1: Der AMS-Algorithmus – Bestandteile.....	26
Abbildung 2: Personendaten und Konstellationen – Datenverfügbarkeit	30
Abbildung 3: Merkmalsprofil der gesamten Population zu Beginn des Geschäftsfalls: Anzahl, Median, Anteil im unteren/oberen Bereich, Trefferquote nach Geschlecht.....	42
Abbildung 4: RGS-Typen Visualisierung in Wien im Jahr 2016	43
Abbildung 5: Anteil Personen mit max. Pflichtschule in Wien im Jahr 2016	43
Abbildung 6: Population mit Migrationshintergrund und ihr Merkmalsprofil zu Beginn des Geschäftsfall: Anzahl, Median, Anteil im unteren/oberen Bereich, Trefferquote nach Geschlecht und Alter	45
Abbildung 7: Voll valide beobachtbare Population im Verlauf des Geschäftsfall im Überblick: Zahl der Geschäftsfälle, Median, Anteil im unteren/oberen Bereich, Trefferquoten	50
Abbildung 8: Begrenzte Aussagekraft einzelner Kennzahlen.....	56
Abbildung 9: Evaluation der Logistischen Regression	58
Abbildung 10: Vier Phasen des Kernprozesses „Arbeitskräfte unterstützen“]	64
Abbildung 11: Ausschnitt aus Detail „Phase 2 Auftragsabklärung, Vereinbarung & Matching“	65
Abbildung 12: Grafische Benutzungsoberfläche des EDV-Systems „Personenstammdaten“	67
Abbildung 13: Geschäftsfunktion „Arbeitsmarktchance“ im PST	68
Abbildung 14: Kriterien für die Eintragung eines BAM	69
Abbildung 15: Aufstufung von Computer-Arbeitsmarktchance „niedrig“ (CAMN) auf Berater*innen-Arbeitsmarktchance „mittel“ (BAMM).....	70
Abbildung 16: Segmentzusatzinformation zum CAM.....	71
Abbildung 17: Faktoren für Segmentzusatzinformationen	72
Abbildung 18: Kontaktintervalle nach Segmenten	73

Tabellenverzeichnis

Tabelle 1: Variablen und deren Ausprägungen	29
Tabelle 2: Fiktive Beispiele für korrekte und inkorrekte Segmentzuordnungen und Prognosen	53

DANKSAGUNG

Wir möchten uns bei einer Reihe an Institutionen und Personen bedanken, deren Expertise und Kooperation wesentlich für dieses Projekt waren.

Der Arbeiterkammer Oberösterreich, und insbesondere Dr. Dennis Tamesberger, wollen wir für die Finanzierung der Studie, sowie für anregende Diskussionen und Gedanken im Rahmen der Zwischen- und Endpräsentation unserer Studie danken. Das Österreichische Arbeitsmarktservice (AMS) hat uns dankenswerterweise im Rahmen unserer Datenanfrage, sowie eines persönlichen Treffens mit Vertreter*innen von AMS und der Synthesis Forschung GmbH, Einblicke in die Entstehungsgeschichte, technische Funktionsweisen sowie soziale Einbettung von AMAS in die Beratungspraxis gegeben.

Das Projekt wurde durch einen wissenschaftlichen Beirat begleitet, der aus Univ.-Prof. Dr. Bettina Berendt (Professorin für Internet und Gesellschaft an der TU Berlin und Direktorin des Weizenbaum-Instituts), Univ.-Prof. Dr. Iris Eisenberger (Professorin für Öffentliches Recht und Europäisches Wirtschaftsrecht an der Universität Graz), sowie Univ.-Prof. Dr. Johanna Hofbauer (Professorin am Institut für Soziologie und empirische Sozialforschung an der WU Wien) bestand. Wir bedanken uns herzlich für die wertvollen Einsichten und Einschätzungen, die aus dieser Kooperation hervorgingen und es uns ermöglichten unsere Ergebnisse in den internationalen Stand der Forschung an der Schnittstelle von Informatik, Soziologie und Rechtswissenschaften einzubetten.

Darüber hinaus wollen wir uns bei unserem internen Beirat bedanken, der sich aus Kolleg*innen des Instituts für Technikfolgen-Abschätzung (ITA) zusammensetzte. Dr. Karen Kastenhofer, Dr. Titus Udrea sowie Dr. Walter Peissl haben uns insbesondere gegen Ende des Projekts mit hilfreichem Feedback unterstützt. Barbara Saringer-Bory und Manuela Kaitna möchten wir für Lektorat und Layout des Berichts danken.

ZUSAMMENFASSUNG

Das Arbeitsmarktchancen-Assistenz-System (AMAS) teilt Arbeitssuchende in drei Kategorien ein: jene mit „hohen“ (H), „mittleren“ (M) und „niedrigen“ (N) prognostizierten Chancen innerhalb eines bestimmten Zeitraums wieder Arbeit zu finden. Das arbeitsmarktpolitische Ziel, das an die Kategorisierung geknüpft ist, sieht eine Fokussierung der Fördermittel im „mittleren“ Segment vor.

AMAS verfolgt drei konkrete Ziele:

- Erhöhung der Effizienz und Effektivität in der Beratung,
- Verbesserung der Effektivität von Fördermaßnahmen, sowie
- Standardisierung der Vergabe von Fördermitteln und Vermeidung von Willkür

Hier zeigt sich eine Zielheterogenität, die eine Reihe an Herausforderungen mit sich bringt. Zudem bildet AMAS eine historische Sicht auf den Wiedereinstieg ehemals Arbeitsloser ab und folgt der Grundannahme, dass sich zukünftige Ereignisse und Zusammenhänge ähnlich verhalten wie vergangene Beobachtungen. Wie irreführend und potenziell problematisch diese Annahme sein kann, zeigt sich derzeit etwa durch den radikalen Einfluss der COVID-19-Pandemie, deren Auswirkungen auf den Arbeitsmarkt signifikant und aus den Daten unmöglich vorherzusagen waren.

Die vorliegende Studie untersucht soziotechnische Wirkweisen des Systems AMAS und die Konsequenzen seines Einsatzes in der Beratungspraxis. Sie fokussiert auf Einschreibungen gesellschaftlicher Ungleichheitslagen und Verzerrungen (Bias) sowie auf Spannungsfelder, die bei der Verwendung des Systems in derzeitiger Form auftreten und sich in einem Widerspruch zwischen Effizienzsteigerung und Kund*innen-Orientierung des AMS manifestieren. Eine vollständige Evaluierung von AMAS im technischen Sinn war aufgrund von fehlendem Zugang zum System, sowie den zugrundeliegenden Daten nicht möglich und auch nicht Ziel dieser Studie. Rückschlüsse auf grundsätzliche Funktionsweisen und Auswirkungen von AMAS sind auf Basis der analysierten Materialien, sowie der einschlägigen wissenschaftlichen Forschung dennoch möglich und fundiert.

Die *technische Rekonstruktion des Systems* zeigt eine konzeptuelle Trennung des Systems in eine „Informationsbasis“, die gesammelte Daten zu Arbeitssuchenden enthält, und ein „Assistenzsystem“, das die Klassifikation der Arbeitssuchenden anhand der errechneten „Integrationschance“ in den Arbeitsmarkt (IC-Wert) vornimmt. Die Verarbeitung der Informationsbasis erfolgt durch zwei statistische Verfahren: Stratifikation und logistische Regression. Das Stratifikationsverfahren teilt Arbeitssuchende anhand von Personeneigenschaften und anderer, personenbezogener Variablen in sogenannte „Konstellationen“ ein. Die logistische Regression verdeutlicht den Einfluss einzelner Variablenausprägungen auf die „Integrationschance“ und wird dafür verwendet, sog. „Segmentzusatzinformationen“ zu gene-

soziotechnische
Analyse

Informationsbasis
und Assistenzsystem

rieren, die in der Beratungssituation das Zustandekommen der Einstufung erklären sollen. In der Beratungssituation wird den AMS-Berater*innen über eine EDV Maske die vom AMAS generierte Einstufung (IC-Wert) der*s Arbeitssuchenden als Prozentwert und als Klassifikation in eine der drei Gruppen (N, M, H) angezeigt. Der Output des „Assistenzsystems“ ist somit die „Berechnung“ der „Integrationschance“: Im Kern fasst der Algorithmus Personen gleicher oder ähnlicher Personeneigenschaften, in einer Konstellation zusammen und setzt diese in Beziehung zu einem kurz- und einem langfristigen „Integrationskriterium“ (i. e. innerhalb einer bestimmten Zeit für eine festgesetzte Dauer Beschäftigung zu finden). Haben von 50 Personen einer Konstellation 25 Personen das Integrationskriterium erfüllt, ergibt dies eine Prognose der „Integrationschance“ von 50 % für alle zukünftigen Arbeitssuchenden, die dieser Konstellation zugerechnet werden. Es wird somit angenommen, dass eine Chancenhomogenität innerhalb von „Konstellationen“ besteht.

Verzerrungen und Bias

Diese Chancenhomogenität steht jedoch im Widerspruch zu aufgezeigten Formen von Bias (Verzerrungen) des Systems.

- Sogenannter *technischer Bias* entsteht beispielsweise durch eine kleine Anzahl an unscharfen bzw. mehrdeutigen Variablen, die eine komplexe Realität vereinfachen sollen. So werden durch die Simplifizierung der Altersklassen etwa zwei Personen im Alter von 30 und 49 Jahren gleich behandelt, zwei Personen im Alter von jeweils 49 und 51 Jahren unterschiedlich.
- *Emergenter Bias* entsteht über die Zeit hinweg und ist darauf zurückzuführen, dass ein auf historischen Daten basierendes System wie AMAS unzureichend auf gesellschaftliche oder ökonomische Veränderungen, wie z. B. COVID 19 oder sich verändernde Geschlechterverhältnisse, reagieren kann.
- *Gegebener Bias* wird durch bestehende Ungleichheiten am Arbeitsmarkt in AMAS eingeschrieben. So fließen in den IC-Wert historische und kumulative Ungleichbehandlung und Diskriminierung am Arbeitsmarkt ein.

Service, Beratung, Betreuung

Von der Einstufung zu trennen ist die Entscheidung, die an die Zuordnung zu einem Segment gebunden ist, d. h. der Zugang zu Service-, Beratungs- oder Betreuungsleistungen. Das segmentierte Angebot kann unterschiedliche Bedarfe abdecken. Wenn allerdings bestimmte Personengruppen gehäuft einem Segment zugeordnet werden, weil die Zugehörigkeit zu einer Konstellation stark auf Personencharakteristiken (wie z. B. Geschlecht, Alter, Staatsangehörigkeit) abstellt und andere Einflussfaktoren in geringerem Ausmaß in die Integrationschance einfließen (wie z. B. der „Konjunkturfaktor“), handelt es sich um eine systematische Ungleichbehandlung.

AMAS in der Beratungspraxis

Die Analyse der *Einbettung von AMAS in die Beratungspraxis* zeigt Spannungen und Ambivalenzen für Arbeitsvermittler*innen, für Arbeitssuchende und für die Institution AMS auf. Der Fokus des Kund*innenkontakts verschiebt sich vom persönlichen Förderbedarf einer Einzelperson hin zu Arbeitsmarktchancen auf Basis einer Populationsberechnung. Damit ändert sich die Arbeitspraxis von AMS-Berater*innen gravierend hin zur

Abarbeitung einer Reihe an „Tasks“, die das Assistenz-System erfordert. Obwohl AMAS vom AMS als „zweite Meinung“ für Berater*innen beschrieben wird, besteht die Gefahr, dass der IC-Wert in der Praxis als „Erstmeinung“ gehandhabt wird und den Beratungsprozess stark prägt. Eine Erklärung oder Verhandlung der Einstufung eines*r Kunden*in und eine allfällige Umstufung benötigen mehr Zeit und Engagement, als den berechneten Chancenwert einfach zu übernehmen. Wird eine Umstufung vorgenommen, muss zudem eine Begründung ins System eingetragen werden. Angesichts der knappen Beratungszeit pro Arbeitssuchender*m, liegt eine Priorisierung des IC-Werts nahe. Das heißt, dass zwar die Letztentscheidung und Verantwortung bei den AMS-Berater*innen liegt, dass das System und seine Einbettung in die Praxis jedoch Anreize setzt, eher dem errechneten Chancenwert zu folgen, als diesen zu „korrigieren“. Die Erwartungshaltung des AMS mittels AMAS eine effizientere und effektivere Vergabe von Fördermitteln zu erzielen, steht damit in einem Spannungsverhältnis zu Anforderung an das System als „Zweitmeinung“ zu fungieren. Die Verschiebung vom persönlichen Förderbedarf hin zu Berechnungen auf Populationsebene zieht auch viele Konsequenzen für Arbeitssuchende nach sich. Für sie bedeutet der Fokus auf den IC-Wert und die Einstufung, dass ihre Biographie und Fähigkeiten auf einen scheinbar „objektiven“ Wert reduziert werden, der Auskunft über ihre Chancen am Arbeitsmarkt geben soll. Der Anspruch der kund*innen-orientierten Betreuung steht damit der potentiellen Reduktion der Biographie auf einen Score gegenüber.

Entgegen der Vorstellung, dass die statistisch berechnete Integrationschance (IC-Wert) lediglich eine Zusatzfunktion zur „evidenzbasierten“ Entscheidungsfindung darstellen soll, zeigt die vorliegende Studie weitreichende Konsequenzen auf unterschiedlichen Ebenen auf. Gemessen an den Zielvorgaben Effizienz, Schulungseffektivität und Standardisierung der Beratungspraxis lassen sich daher folgende Schlüsse ziehen:

Schlussfolgerungen

1. Eine Steigerung der Effizienz des Beratungsprozesses geht nur mit einer vorwiegend routinemäßigen Übernahme der computergenerierten Einstufung einher. Dies mag im Sinne einer Dienstleistungsrationalisierung nachvollziehbar sein, im Sinne der Serviceorientierung einer semi-staatlichen Institution, die ihrem öffentlichen Auftrag zur Erhöhung der Chancen von Arbeitssuchenden nachkommt, aber am Ziel vorbei gehen.
2. Die Erhöhung der „Schulungseffektivität“ durch Konzentration der Fördermittel im mittleren Segment zielt nicht auf die Zielgenauigkeit und Qualität der Maßnahme für eine*n AMS-Kunden*in ab, sondern verbindet das Ziel des „effektiven“ Mitteleinsatzes mit einer groben Einteilung in drei Kund*innengruppen.
3. In der Entwicklung des AMAS wurden kaum Verfahren eingesetzt, um Bias (Verzerrungen) im System zu vermeiden und auch in der Anwendung bietet das System keine Anhaltspunkte, um möglichen strukturellen Ungleichbehandlungen vorzubeugen. Gleichstellungsaspekte spielen im Betreuungsprozess eine wichtige Rolle, der eine objektivier-

te, numerische (und nicht hinsichtlich Bias optimierte) Einstufung der Integrationschancen zuwiderlaufen kann.

Empfehlungen

Aus den Ergebnissen der Studie leiten sich Empfehlungen ab, die sich konkret auf AMAS, auf algorithmische Systeme in (semi-)staatlichen Einrichtungen, sowie auf weitreichendere gesetzliche und institutionelle Veränderungen beziehen. Die Empfehlungen zu AMAS reichen von System- und Datentransparenz, sowie einer Bedarfserhebung von Bedürfnissen in der Systementwicklung über Einsichts- und Einspruchsrechte von Betroffenen und neuen Kompetenzen für AMS-Berater*innen bis hin zu mehr Ressourcen für einen besseren Betreuungsschlüssel in der Beratungspraxis. Die Empfehlungen für (semi-)staatliche Institutionen fokussieren primär auf Anti-Diskriminierungsstrategien, öffentliche Konsultationen und eine Informationspflicht für Betroffene, sowie für ausgewählte Personengruppen wie Forscher*innen oder Journalist*innen und (neu geschaffene) Aufsichtsorgane. Die institutionellen Empfehlungen beziehen sich auf größere Veränderungen, die (semi-)automatisierte Entscheidungssysteme nach sich ziehen, wie z. B. die Notwendigkeit gesetzlicher Anpassungen, sowie der Bedarf an neuen Gremien, Aufsichtsbehörden und Auditverfahren, welche zu fundierten gesellschaftlichen Debatten über die weitreichenden Konsequenzen solcher Systeme für Gesellschaft und Politik beitragen können.

1 AMAS AUS SOZIOTECHNISCHER PERSPEKTIVE

Mit Jänner 2021 plante das Arbeitsmarktservice der Republik Österreich (AMS) den flächendeckenden Einsatz ihres Assistenzsystems AMAS (Arbeitsmarktchancen-Assistenz-System), das 2015 bei der Firma Synthesis Forschung GmbH in Auftrag gegeben wurde. Basierend auf einem statistischen Modell der Arbeitsmarktchancen von Arbeitssuchenden teilt das System – landläufig auch als „AMS-Algorithmus“ bezeichnet – Kund*innen in drei Kategorien ein: jene mit „hohen Chancen“, innerhalb des folgenden halben Jahres wieder einen Job zu finden, jene mit mittelmäßigen Chancen am Arbeitsmarkt, und jene, denen das System schlechte Chancen bescheinigt, innerhalb von zwei Jahren ab Zuweisung zum AMS wieder eine Arbeit aufnehmen zu können. Basierend auf dieser Kategorisierung sollen den Kund*innen unterschiedliche Maßnahmen der Unterstützung seitens des AMS zur Verfügung stehen.

Die Einführung des algorithmischen Profiling-Systems verfolgt vorrangig zwei Ziele: Zum einen soll die Effizienz und Effektivität der Beratung erhöht werden, d. h. der Zeit- und Ressourceneinsatz der AMS-Berater*innen soll durch Differenzierung der Kund*innen in „Servicekund*innen“ (mit „hohen“ Arbeitsmarktchancen), „Beratungskund*innen“ (mit „mittleren Arbeitsmarktchancen“) und „Betreuungskund*innen“ (mit „niedrigen Arbeitsmarktchancen“) zielgenauer werden. Zum zweiten soll die Effektivität von Arbeitsmarktprogrammen und -maßnahmen verbessert und im Hinblick auf arbeitsmarktpolitische Ziele der Bundesregierung Ressourcen für qualifizierende und beschäftigungsfördernde Maßnahmen überwiegend im „mittleren“ Segment eingesetzt werden [BER_13 2019].¹ Kund*innen, denen AMAS „hohe Arbeitsmarktchancen“ vorhersagt, würden auch ohne qualifizierende und beschäftigungsfördernde Maßnahmen schnell einen Arbeitsplatz finden. Kund*innen, denen das System „niedrige Arbeitsmarktchancen“ prognostiziert, würden dagegen trotz dieser Maßnahmen keinen schnellen Wiedereinstieg erreichen und könnten mehr von „unterstützenden Angeboten“ profitieren. Als drittes, weniger zentrales Ziel, findet sich in einigen Dokumenten des AMS die Anforderung, dem

Ziele des
Arbeitsmarktchancen-
Assistenz-Systems

¹ Aufgrund der in Österreich Mitte März 2020 einsetzenden Corona-Krise wurden sowohl hinsichtlich des Einsatzes von AMAS, sowie hinsichtlich der Umsetzung dieses arbeitsmarktpolitischen Ziels Änderungsmaßnahmen nötig und die flächendeckende Einführung des Systems wurde auf Jänner 2021 verschoben. Aufgrund eines Bescheids der Österreichischen Datenschutzbehörde vom August 2020 [DSB 2020] wurde der Einsatz des Systems ab 1.1.2021 untersagt, sofern bis dahin keine gesetzliche Grundlage dafür geschaffen wird (siehe Kapitel 2). Zum derzeitigen Zeitpunkt (November 2020) ist der Betrieb von AMAS ausgesetzt.

Vorwurf der willkürlichen Vergaben von Fördermaßnahmen durch eine evidenzbasierte Vergabe entgegen zu wirken.

New Public Management

Erste Versuche zur Einführung eines Profiling-Systems am AMS gab es bereits im Jahr 2008 und der langjährige Prozess mündete schließlich im Herbst 2018 im Testbetrieb von AMAS. Dieser Prozess knüpft an politische und organisationale Entwicklungen an, die 1994 in einem Umbau der österreichischen „Arbeitsämter“ in teilautonome Serviceeinrichtungen resultierte. Wie Penz et al. (2015) in ihrer ländervergleichenden Studie zur „Transformation vom versorgenden zum aktivierenden Sozialstaat“ in Österreich, Deutschland und der Schweiz dargelegt haben, ging damit eine schrittweise Einführung von Techniken des New Public Management² einher. Die Umgestaltung bürokratischer Verwaltungsarbeit in eine selbstverwaltete Dienstleistungsorganisation verändert „die Anforderungen an das berufliche Handeln von Beschäftigten (semi-)staatlicher Institutionen. Gefordert sind höhere Effizienz und Effektivität öffentlicher Dienstleistungen sowie Bürger*innennähe, also kundenorientierte Interaktionsarbeit“ (Penz et al. 2015, 21).

Serviceorientierung und Effizienz

Auch eine AMS-interne Präsentation über die Einführung des Arbeitsmarktchancen-Assistenz-Systems AMAS vom Februar 2020 spricht sowohl die Kund*innenorientierung – im Sinne von Dienstleistungsarbeit und Problemverwaltung – als auch Effizienzkriterien an [SCHU_7 2020, S. 6]: die Aufgabe eines öffentlichen Arbeitsmarktservices ist es, Arbeitssuchende mit Arbeitgeber*innen zusammenzuführen (zu matchen), die Chancen von Arbeitssuchenden durch Qualifizierung und den Abbau von Vermittlungshindernissen zu erhöhen und die dafür zur Verfügung stehenden Mittel effizient einzusetzen. Grundbedingungen um die Ausgangslage der AMS-Kund*innen gut einzuschätzen seien „Kompetenz, Wissen, Erfahrung und Zeit“ der AMS-Mitarbeiter*innen. Und es stelle sich die Frage: „Können wir dafür auch das ‚Wissen in den Daten‘ nutzen?“ [SCHU_7 2020, S. 6]. Der technologische Fortschritt ermögliche „eine systematische Bearbeitung von großen Datensätzen“ und es sei ein „Anspruch des AMS ‚state of the art‘ zu sein“ [SCHU_7 2020, S. 7]. Aus organisationaler Sicht stellt diese Form der technischen Modernisierung einer post-bürokratischen Dienstleistungsorganisation einen naheliegenden nächsten Schritt dar.

Daten bilden eine spezifische Perspektive ab

Die Einführung eines datenbasierten Assistenzsystems stellt sich als Modernisierungsprojekt dar, das „evidenzbasierte Entscheidungsfindung“ mit Vorstellungen von Effizienz, Objektivität und Genauigkeit verbindet.

² New Public Management bezeichnet einen Reformprozess, der Bürokratien „in wettbewerbsorientierte öffentliche Dienstleistungsunternehmen verwandeln soll. Output-Steuerung (oder management by objectives) anhand von Wirkungsindikatoren (key performance indicators, KPI) beginnt die starren Richtlinien und hierarchischen Vorgaben der alten bürokratischen Ordnung zu ersetzen, balanced scorecards ermöglichen eine interne vergleichende Qualitätskontrolle, über benchmarking wird die Dienstleistungsqualität von organisatorischen Subeinheiten zueinander in Beziehung gesetzt“ (Penz et al. 2015, S. 23).

Die Vorstellung, dass Big Data eine detaillierte, umfassende und „wahrheitsgemäße“ Darstellung der „Realität“ ermöglichen (Daston & Galison 2007; Rieder & Simon 2016; McQuillan 2018), blendet allerdings aus, dass Daten immer in einem bestimmten Kontext und mit einem bestimmten Ziel erstellt werden (Bowker 2005; Boyd & Crawford 2012). Daten geben daher nur eine bestimmte Perspektive auf einen simplifizierten Ausschnitt von „Realität“ wieder (Kitchin 2014, S. 133). Auch das System AMAS spiegelt in diesem Sinne bestimmte sozialpolitische Perspektiven, Wertentscheidungen und historisch gewachsene organisationale Normen wieder. Die administrativen Daten, die die Basis des Assistenzsystems bilden, stellen einen spezifischen Blick auf Arbeitsmarktchancen dar.

1.1 STUDIENDESIGN

Wie zahlreiche Studien gezeigt haben, kann der Einsatz algorithmischer Systeme zur Bewertung oder Klassifizierung von Menschen insbesondere in der öffentlichen Verwaltung weitreichende Konsequenzen haben (AI Now Institute 2018; Eubanks 2018; Veale et al. 2018; Gangadharan & Niklas 2019). Die vorliegende Studie untersucht daher den AMS-Algorithmus aus einer soziotechnischen Perspektive und analysiert, welche organisationalen und technischen Entscheidungen bei der Entwicklung von AMAS getroffen wurden, und welche Konsequenzen der Einsatz des Systems in der Beratungspraxis nahelegt. Zu diesem Zweck wurde die Studie unter der Projektleitung von Doris Allhutter und Astrid Mager (Institut für Technikfolgen-Abschätzung (ITA) der Österreichischen Akademie der Wissenschaften) und in Kooperation mit Florian Cech und Fabian Fischer (Centre for Informatics & Society der TU Wien) sowie Gabriel Grill (University of Michigan at Ann Arbor) durchgeführt. Die Analyse beleuchtet die Spannungsfelder, Herausforderungen und potentiellen Bias, die das System in den Beratungsprozess einbringt, und diskutiert diese Dynamiken im Rahmen der Zielsetzungen des AMS, sowie darüber hinausgehend hinsichtlich des Einsatzes evidenzbasierter algorithmischer Entscheidungssysteme in (semi-)staatlichen Institutionen.

Die vorliegende soziotechnische Analyse von AMAS rekonstruiert und hinterfragt die technischen Grundlagen des Systems, sowie dessen Einbettung in die Betreuungspraxis entlang folgender Forschungsfragen:

Forschungsfragen

- Welche Zielsetzungen verfolgt AMAS auf organisationaler und operativer Ebene? Welche Leitprinzipien informieren die technische Ausgestaltung des Systems?
- Wie wurde AMAS technisch umgesetzt? Welche Daten über vergangene Berufsverläufe werden herangezogen und aus welchen Bestandteilen besteht der AMS-Algorithmus?
- Welche Formen von Bias, Diskriminierung und Fehlerquoten spielen dabei eine Rolle und wie wurden diese konkret berücksichtigt?

- Welche Auswirkungen kann der Einsatz des Systems für die gängige AMS-Praxis und die Arbeitssuchenden haben?
- Welche Empfehlungen hinsichtlich Bias und Diskriminierung, Transparenz und politischer Steuerungs-Prozesse und Governance-Institutionen leiten sich daraus ab?

Literatur, Material und Methode

Die Technikfolgenabschätzung betrachtet technologische Entwicklungen in ihrer soziotechnischen Dimension, d. h. als etwas, das nicht ausschließlich technikzentrierten Logiken folgt, sondern in sozialen Kontexten entsteht und diese auch mitgestaltet. In diesem Sinne analysiert dieser Bericht die soziotechnischen Dimensionen von AMAS auf Basis aktueller Forschung, konkret aus den *Critical Data Studies* und dem Bereich *Fairness, Accountability and Transparency in Sociotechnical Systems*³. Mittels einer tiefgehenden Dokumentenanalyse (Silverman 2015; Bowen 2009) interner und öffentlich publizierter Dokumente des AMS untersuchen wir die konzeptionellen, technischen und sozialen Implikationen dieses Systems. Die der Analyse zugrunde liegenden Dokumente sind in einem Verzeichnis im Anhang dieses Berichts aufgelistet und mit Abkürzungen versehen, die im Text in eckigen Klammern und Großbuchstaben referenziert sind.⁴

Seitens des AMS und der Firma Synthesis Forschung GmbH waren Informationen zu AMAS bis zu Beginn dieser Studie im August 2019 spärlich [DOK_1 2018; NOTES_6 2019] und wurden vor allem durch Anfragen von Institutionen und Aufsichtsorganen, wie dem Rechnungshof Österreich [BER_15 2017], der Anwaltschaft für Gleichbehandlungsfragen [NOTES_10 2019], der Volksanwaltschaft [BER_14 2019], sowie der zivilgesellschaftlichen Organisation Epicenter.Works [NOTES_7 2019]⁵ öffentlich zugänglich. Auf Basis einer fokussierten Analyse dieser Dokumente (Allhutter et al. 2020) stellten wir im Oktober 2019 eine detaillierte Anfrage an das AMS [NOTES_8 2019], die im Rahmen eines Treffens mit Vertreter*innen des AMS und von Synthesis Forschung am 23. Oktober 2019 besprochen und konkretisiert wurde. Im Februar 2020 wurde diese Anfrage vom AMS in Kooperation mit Synthesis Forschung schriftlich beantwortet [BRIEF_3 2020]. Die Beantwortung referenzierte eine Liste an AMS-internen Dokumenten, wie technische Dokumentationen, Richtlinien, Protokolle, Präsentationen, Schulungsunterlagen etc. sowie rechtliche Dokumente, die uns das AMS zur Verfügung stellte (siehe Dokumente-Verzeichnis). Zwei weitere Rückfragen unsererseits zur Klärung offener Punkte wurden von Synthesis Forschung schriftlich beantwortet (NOTES_5 2020; NOTES_9 2020). Dieser Bericht speist sich aus der Analyse all dieser Dokumente entlang unserer Forschungsfragen.

Ziel der Studie

Im Rahmen dieser Studie hatten wir (u. a. aus datenschutzrechtlichen Gründen) keinen Zugriff auf die verarbeiteten Daten des Data Warehouse des AMS oder auf das System AMAS selbst. Eine tatsächliche Evaluierung

³ <https://facctconference.org/>

⁴ Siehe Dokumente-Verzeichnis am Ende des Berichts.

⁵ <https://epicenter.works/>

der technischen Funktionsweisen von AMAS ist uns aus diesem Grund nicht möglich und war auch nicht Ziel dieser Studie. Eine Unterscheidung zwischen vorliegender soziotechnischer Analyse und einer technischen Evaluierung ist relevant, da beispielsweise die Berechnung von Fehlerraten oder von konkreten Integrationschancenwerten nur mit Zugriff auf die Daten und das System möglich ist. Auch Bias oder Fairness der algorithmischen Klassifizierung können empirisch nur nachgewiesen werden, wenn anhand von Daten festgestellt werden kann, ob unterschiedliche Personengruppen vor dem Einsatz des Systems besser oder schlechter gefördert wurden als nach der Einführung. Rückschlüsse auf die grundsätzlichen Funktionsweisen und Auswirkungen von AMAS sind allerdings auf Basis der technischen Dokumentation und unsere Anfrage sowie unter Einbeziehung der Faktenlage aus der einschlägigen Forschung wissenschaftlich möglich und fundiert.

Der Bericht ist wie folgt aufgebaut: Kapitel 2 stellt die Entstehungsgeschichte des AMS-Algorithmus dar. Kapitel 3 widmet sich der technischen Rekonstruktion von AMAS und bietet eine kritische Analyse des Systems in Hinblick auf Fragen von inhärenten Bias, Diskriminierung, Fehlerraten sowie Transparenz und Nachvollziehbarkeit („*Accountability*“). Kapitel 4 behandelt die Einführung von AMAS in die Arbeitspraxis des AMS und geht insbesondere auf die Spannungen und Ambivalenzen ein, die das System für Mitarbeiter*innen und Kund*innen des AMS aufwirft. Kapitel 5 präsentiert vier illustrative Szenarien, anhand derer das Zusammenwirken der technischen und sozialen Dynamiken exemplarisch vorgeführt und hinsichtlich Zielerreichung durch den AMS-Algorithmus diskutiert wird. An unterschiedlichen Stellen des Berichts werden vergleichbare Systeme, die in anderen Ländern im Einsatz sind bzw. waren, angesprochen und später gesammelt in Form von Ländervignetten in Kapitel 6 vorgestellt. Kapitel 7 zieht Schlussfolgerungen aus der Analyse und formuliert darauf basierende Empfehlungen. Am Ende des Berichts finden sich eine Literaturliste sowie ein Dokumente-Verzeichnis, die alle Quellen auflistet, die diesem Bericht zugrunde liegen.

Aufbau des Berichts

1.2 ARBEITSMARKTCHANCEN ALS OPTIMIERUNGSPROBLEM

Während sich die Vermittlungstätigkeit des AMS weiterhin auf alle Kund*innen gleichermaßen erstreckt, verfolgt AMAS eine „neue Strategie“ in der Betreuung, indem die Bereiche Beratung, Qualifizierung und Förderung auf die vom System errechneten Arbeitsmarktintegrationschancen abstellen [SCHU_7 2020, S. 12]. Konkret bedeutet das, dass arbeitslos gemeldete mit „hohen Integrationschancen“ als „Servicekund*innen“ in der Servicezone des AMS betreut werden, wo der primäre Fokus auf der Vermittlung der Arbeitssuchenden und nicht auf deren Förderung oder Schu-

Leitprinzipien des Systemdesigns

lung liegt. Arbeitssuchenden mit „mittleren Integrationschancen“ stehen als „Beratungskund*innen“ nach Maßgabe Mittel für qualifizierende und beschäftigungsfördernde Maßnahmen zur Verfügung. Kund*innen mit „niedrigen Integrationschancen“ werden in die sogenannten BBEN (Beratungs- und Betreuungseinrichtungen Neu) und in „zweckmäßigere Angebote“ ausgliedert. BBEN befinden sich als neue Betreuungsvariante derzeit noch im Aufbau und beinhalten primär stabilisierende und unterstützende Maßnahmen [BER_1 2019]. In den arbeitsmarktpolitischen Zielen 2020 wurde (vor der Corona-Krise) als Ziel die Erhöhung der „Schulungseffektivität nur für das mittlere Segment“ genannt und als Indikator wird der Erfolg der Arbeitsmarktförderung und die Reduzierung der Geschäftsfalldauer im „hohen“ und „mittleren“ Segment definiert. Dieser Erfolg wird durch eine *Balanced Score Card* gemessen, die als Leistungsindikator eine AMS-interne vergleichende Qualitätskontrolle ermöglicht. Die Förderquote im „niedrigen“ Segment wurde 2020 „im Hinblick auf den Aufbau des BBEN Angebots“ ausgesetzt [BER_1 2019, S. 17]. In der Ressourcenverteilung wird damit ein Schwerpunkt im „mittleren“ Segment gesetzt, während bei Kund*innen mit „hohen“ und mit „niedrigen Chancen“ eingespart werden soll. Die Förderung von politisch definierten Zielgruppen, insbesondere Jugendliche, Frauen, Menschen mit Behinderung und 50+ bleibt weiterhin aufrecht. Das heißt, diese Schwerpunktgruppen können unabhängig von den Arbeitsmarktchancen an entsprechenden Maßnahmen teilnehmen, sofern sie die dafür notwendigen Voraussetzungen erfüllen und genug Budget vorhanden ist.⁶ Der Art und Weise wie der AMS-Algorithmus gestaltet ist, liegen also bestimmte Ziele und Leitprinzipien zugrunde. Knappe Fördermittel für Fortbildung zur Erhöhung der Chancen am Arbeitsmarkt sollen so verteilt werden, dass eine möglichst hohe Vermittlungsquote erreicht werden kann.

Was wird optimiert?

Wie einleitend dargestellt, werden konkret drei Zielsetzungen schlagend: 1) die Steigerung der Effizienz des Beratungsprozesses durch eine Differenzierung der Kontakthäufigkeit und -intensität der drei Kund*innengruppen, 2) die Gewährleistung eines effektiven Mitteleinsatzes und damit eine Steigerung der Effektivität des Förderangebots und 3) eine Standardisierung der Vergabe von Fördermitteln, die einer etwaigen Willkür von Berater*innen beikommen soll. In informatischen Begriffen beschreiben diese drei Ziele unterschiedliche Optimierungsprobleme. Im ersten Fall sollen Arbeitszeit und Betreuungsaufwand von AMS-Mitarbeiter*innen optimiert werden; im zweiten Fall sollen die Ausgaben für Fördermaßnahmen und unterstützende Angebote hinsichtlich der Vermittlungschancen der drei Kund*innengruppen optimiert werden; im dritten Fall soll auf eine gerechte Vergabe von Fördermaßnahmen hin optimiert wer-

⁶ Beispielsweise fielen im Jahr 2018 von 61.190 Menschen mit gesundheitlicher Beeinträchtigung lediglich 12.421, die einen Feststellungsbescheid hatten, in die Zielgruppenförderung für Menschen mit Behinderung (Email Mag. Bruckner/Österreichischer Behindertenrat (Recht und Sozialpolitik), 3.10.2019).

den, indem mögliche (systematische) Bias bzw. Diskriminierungen durch AMS-Mitarbeiter*innen beseitigt werden sollen.

In diesen Optimierungsproblemen zeigt sich eine Zielheterogenität, die eine Reihe an Herausforderungen mit sich bringt. Im Grunde erfordert jedes dieser Ziele unterschiedliche Maßnahmen und muss mit einem ausreichenden Maß an Fokussierung und Sensibilität behandelt werden. Ziele wie die Bewältigung eines hohen Arbeitspensums, der „zielgenaue“ Einsatz von Förderressourcen und die gerechte Vergabe von Qualifizierungsmaßnahmen verfolgen allerdings unterschiedliche Rationalitäten. Jedes für sich ist vielschichtig und umfasst etwa Aspekte von Qualitätssicherung, Arbeitserleichterung oder Arbeitszufriedenheit, Sichtweisen auf Effektivität und Solidarprinzip (je nach arbeitsmarktpolitischen Ziel) oder Grundrechte wie Gerechtigkeit und Chancengleichheit. Vor diesem Hintergrund stellt sich die Frage, auf welches Ziel hin AMAS im technischen Sinn optimiert worden ist und auf Basis welcher Annahmen Verbindungen zwischen diesen Zielen hergestellt werden oder sich im System abbilden.

**Heterogenität
in der Zielsetzung**

2 ENTSTEHUNGSGESCHICHTE VON AMAS

Die ersten Anläufe des AMS zum „Profiling“, wie es ursprünglich genannt wurde [BER_11 2009], reichen bis ins Jahr 2008 zurück. Der Anstieg der Arbeitslosenzahlen aufgrund der Weltwirtschaftskrise wurde in unterschiedlichen Ländern als Anlass genommen, an Methoden des „Profiling und Targeting“ zu arbeiten [BER_11 2009]. Diesem internationalen Trend folgte auch das AMS und führte in Anlehnung an die Länder Deutschland, Schweiz und Dänemark von 2008 bis 2009 einen Pilotversuch durch (gemeinsam mit der Firma Synthesis Forschung GmbH). Wie in den Ländervignetten in Kapitel 6 näher ausgeführt, berechnet das deutsche System die Wahrscheinlichkeit der Vermittlungsfähigkeit in einer bestimmten Zielregion. In Dänemark wurde ein Profiling-System getestet, das das Risiko auf Langzeitarbeitslosigkeit prognostizierte ohne die Kategorisierung jedoch an den Zugang zu Maßnahmen zu knüpfen. In der Schweiz wurde kein Profiling sondern ein statistisches Targeting-System eingeführt, das Maßnahmen mit guter Wirksamkeit für die jeweilige Personen identifizierte, aber nach kurzer Zeit wieder eingestellt wurde.

Profiling
und Targeting

Der Pilotversuch 2008-2009 wurde in drei AMS Geschäftsstellen durchgeführt. Mit Grieskirchen, Zell am See und Wien/Schlosshoferstrasse wurden entsprechend der unterschiedlichen Zusammensetzung der regionalen Arbeitsmärkte in Österreich drei sehr unterschiedliche Regionen ausgewählt [BER_11 2009]. Einige Elemente des Systems weisen Ähnlichkeiten mit dem aktuellen System AMAS auf. So wurde bereits 2008 bei einer AMS-internen Präsentation zwischen kurz- und langfristigen Prognosen (30 und 90 Tage) unterschieden und die einfließenden Merkmale erinnern teilweise an aktuell benutzte Variablen: „Beschäftigungstage im Vorjahr, regionale Geschäftsstelle, Berufswunsch, Begünstigtenstatus, Alter der Person, Quartal des Vormerkungsbeginns, höchster Bildungsabschluss, Geschlecht, sowie Nationalität“ [PRÄS_5 2008, S. 7]. Die politische Zielsetzung der einzelnen Zuordnungen wurde zu diesem Zeitpunkt nicht klar definiert. Soweit aus den Präsentationsunterlagen ersichtlich, wurde lediglich grob zwischen vier „Orientierungslinien“ für die weitere Betreuung unterschieden: 1) keine Maßnahme notwendig, 2) Qualifizierung kann unterstützen, 3) Aktivierung kann unterstützen und 4) „intensivere“ Maßnahmen sind sinnvoll [PRÄS_5 2008, S. 3]. Die Gruppe der durch „Aktivierung“ zu unterstützenden ist im Zuge der weiteren Entwicklungen offenbar weggefallen. Das aktuelle System arbeitet wie beschrieben mit drei Gruppen, die sich weitgehend mit der ersten, der zweiten und der vierten der damaligen Kategorien decken. Im Endbericht zur Pilotphase 2008-2009 ist weiters zu lesen, dass für Arbeitssuchende mit geringen Integrationschancen mitunter mehr Budget aufzuwenden sei: „Umgekehrt ist eine als niedrig prognostizierte Integrationschance ein Signal, in der Anfangsphase der

erster Pilotversuch

Vormerkung nicht einfach zuzuwarteten. Es wird für den vermutlich notwendigen persönlichen Kontakt, um frühzeitig eine weitere detaillierte Erhebung potenzieller Integrationsbarrieren vorzunehmen, ein größeres Zeitbudget erforderlich sein.“ [BER_12 2011, S. 39] Dies deutet darauf hin, dass es in dieser ersten Experimentierphase noch keine explizit genannte Verknüpfung mit einer an die Segmentierung in vier Gruppen gekoppelte arbeitsmarktpolitische Stoßrichtung gab, die über die oben beschriebenen „Orientierungslinien“ hinausgingen. Ein ausdifferenziertes Angebotspaket an Förder- und Unterstützungsmaßnahmen für die unterschiedlichen Gruppen kristallisierte sich erst später heraus. Darüber hinaus ist anzumerken, dass bereits im Endbericht zur Pilotphase 2008-2009 der Treffsicherheit des Profiling große Bedeutung beigemessen wurde: „[D]ie Einschätzung der praktischen Nutzbarkeit des zugrundeliegenden statistischen Verfahrens wird letztlich in hohem Maße von seiner Treffsicherheit bestimmt“ [BER_12 2011, S. 22]. Sie wurde im Hinblick auf das damals verwendete System zusammenfassend als „ausreichend“ eingeschätzt, „um für praktische Anwendungen in Betracht gezogen zu werden“ [BER_12 2011, S. 39]. Auch wurde das System bereits damals als hilfreiches Werkzeug beschrieben, das zur Standardisierung der AMS Praxis „anhand greifbarer Indikatoren“ [BER_12 2011, S. 39] beitragen kann. Diese Aspekte haben die weitere Entwicklung des Systems begleitet. In den Jahren nach der Wirtschaftskrise sei das AMS mit seiner bisherigen Verteilung der Maßnahmen und Fördermittel – bei gegebenen Ressourcen – an seine Grenzen gestoßen [HAND_1 2019, S. 4]. Steigende Arbeitslosenzahlen, erhöhter Zugang an Asylberechtigten, sowie keine ausreichende Steigerung an Personal und Ressourcen würden eine „strategische Neuausrichtung“ [HAND_1 2019, S. 4] notwendig machen. Dementsprechend wurde 2014 eine strategische Grundsatzdiskussion im AMS begonnen im Hinblick auf Effektivität und Effizienz der eingesetzten, aber der Situation nicht ausreichend angepassten personellen und budgetären Ressourcen [BRIEF_3 2020].

Start des Projekts „Kundensegmentierung“

Im Februar 2015 wurde in einer Strategietagung vom AMS-Vorstand und allen Landesgeschäftsführer*innen der Beschluss gefasst, das Projekt „Kundensegmentierung“ zu starten, das eine neue Betreuungsstrategie und neue, maßgeschneiderte Dienstleistungen für differenzierte Kund*innengruppen beinhalten sollte [BRIEF_3 2020]. Im AMAS Handbuch [HAND_1 2019, S. 5] ist die strategische Zielsetzung des Projekts folgendermaßen formuliert:

1. „Einsparung von Mitteln bei Kund*innen mit hohen Arbeitsmarktchancen
2. Neues Angebot für Kund*innen mit niedrigen Arbeitsmarktchancen
3. Konzentration der Ressourcen auf Kund*innen mit mittleren Arbeitsmarktchancen mit dem größten Mehrwert (additiver Integrationserfolg), u. a. auch auf Basis unternehmensnaher Förderungen“

die Entwicklungsphase

Mit dieser arbeitsmarktpolitischen Ausrichtung wurde das Projekt im Juni 2015 in Angriff genommen. Laut Antwort auf unsere Datenanfrage [NOTES_8 2019] wurden zusätzlich zum AMS-Vorstand und den Landesgeschäftsführer*innen, sowie dem Verwaltungsrat und dem Strategieaus-

schuss, vier Arbeitsgruppen mit AMS-Mitarbeiter*innen aus den Landesorganisationen gebildet, die in den Entwicklungsprozess einbezogen wurden [BRIEF_3 2020]. Diese wurden mit der inhaltlichen Entwicklung des Modells und der neuen Betreuungsangebote (BBEN) betraut. Eine Arbeitsgruppe war primär mit der Modellentwicklung, sowie der Erstellung eines Pflichtenhefts beschäftigt, das als Grundlage eines europaweiten Ausschreibungsverfahrens herangezogen wurde und als Basis für die Angebote diente [AUSSCHR_1 2015]. Nachdem der Auftrag im Dezember 2015 (abermals) an die Firma Synthesis Forschung GmbH ging, wurde das Modell nach den Vorgaben des AMS durch Synthesis Forschung als externem wissenschaftlichen Auftragnehmer entwickelt. Während der Entwicklungsphase präsentierte Synthesis Forschung regelmäßig den Entwicklungsprozess vor dem Verwaltungsrat, den Landesgeschäftsführer*innen und den Arbeitsgruppen. Daran anschließende Diskussionen flossen in den Entwicklungsprozess ein [BRIEF_3 2020]. Die Inbetriebnahme des AMS-Algorithmus war ursprünglich für Ende 2016 vorgesehen, wurde damals jedoch „politisch gestoppt“ [BRIEF_3 2020].

Unter der türkis-blauen Regierung (2017-2019)⁷ wurde das Projekt dann wieder aufgenommen und ging im Herbst 2018 in ausgewählten Regionen Österreichs in Testbetrieb. Das damals zuständige *Ministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz* beschrieb die „personalisierte Arbeitsmarktbetreuung als Schlüssel zur effizienten Arbeitsmarktsteuerung“ [BER_13 2019, S. 3]. „Besonderes Augenmerk muss hier auf die Erhöhung der Schulungseffektivität gelegt werden, damit nur jene Fördermaßnahmen zur Anwendung kommen, die auch tatsächlich für eine Integration in den Arbeitsmarkt tauglich sind.“ [BER_13, S. 4] AMAS solle es ermöglichen, „auch bei reduziertem Ressourceneinsatz ... eine qualitative Beratungsleistung anzubieten“ [BER_13, S. 5]. Konkret wurde erwartet, dass die personalisierte Arbeitsmarktbetreuung zu „Einsparungspotenzial hinsichtlich Personal und Ressourcen“ [BER_13, S. 18] führe. Auch der Rechnungshof bezeichnete AMAS (...) als Instrument der Leistungssteigerung [BER_15 2017, S. 82]. Austeritätspolitik erscheint als einer der Hauptantriebe für die Entwicklung des Systems: sowohl der Rechnungshof, als auch das Sozialministerium forderten vom AMS Effizienzsteigerungen in Zeiten wirtschaftlicher Flaute. Während es anfangs das erklärte Ziel war, bei stagnierendem Budget eine wachsende Zahl an Arbeitssuchenden zu betreuen, wurde der AMS-Algorithmus in späteren Erwartungshaltungen des damals zuständige Sozialministeriums als Mittel für Budgetkürzungen gesehen [BER_13 2019, S. 5].

Rechtlich gesehen basiert AMAS auf dem Arbeitsmarktservice-Gesetz (AMSG), das regelt welche Daten für das System verwendet werden dürfen. Des weiteren ist für den AMS-Algorithmus die von der *Abteilung Ser-*

der Testbetrieb

⁷ Als türkis-blaue Regierung (2017-2019) wird die Koalition zwischen der Österreichischen Volkspartei (ÖVP, türkis) und der Freiheitlichen Partei Österreichs (FPÖ, blau) unter Bundeskanzler Sebastian Kurz bezeichnet, die im Dezember 2017 gebildet und im Mai 2019 beendet wurde.

vice für Arbeitskräfte erstellte Bundesrichtlinie „Arbeitskräfte unterstützen“, kurz KP 1 Richtlinie, zentral. Diese Richtlinie regelt und standardisiert die Beratungspraxis und bildet im Detail den Prozess ab, den AMS-Mitarbeiter*innen im Kund*innenkontakt durchlaufen sollen (siehe auch Kapitel 4). Konkrete Förderrichtlinien für Schwerpunktgruppen sind in der Gleichstellungspolitik und der Ausbildungsgarantie verankert. Die Datenschutzgrundverordnung (DSGVO) regelt schließlich die Datenschutzkonformität des Algorithmus.

Auswirkungen der Corona-Krise

Nach mehrmaligem Verschieben der landesweiten Implementierung – zuletzt wegen der Corona-Krise – war die Inbetriebnahme für Jänner 2021 vorgesehen. Die Corona-Krise stellte das Projekt vor große Herausforderungen und eine Reihe an Änderungsmaßnahmen wurden erforderlich (Telefonat mit Mag. Marte-Huainigg/AMS, 29.6.2020). Konkrete Maßnahmen während der Krise waren das Aussetzen und Ausarbeiten einer neuen KP 1 Richtlinie. Darüber hinaus wurde die Betreuung von Kund*innen auf telefonische Betreuung umgestellt und AMS-Mitarbeiter*innen wurden angewiesen das Beratungsgespräch weitgehend ohne Berücksichtigung der berechneten Integrationschancen (IC) zu führen. Nur wenn eine Förderung vergeben werden soll, sollte der vom System errechnete IC-Wert angesehen werden, um herauszufinden, in welcher Gruppe sich die entsprechende Person befindet. Sollte sich die Person in Gruppe „hoch“ oder „niedrig“ wiederfinden, für die keine Förderung vorgesehen sind, sollten die Mitarbeiter*innen entsprechend in Gruppe M umstufen, um die Förderung vergeben zu können.

Einschätzung der Österreichischen Datenschutzbehörde

Zuletzt bereitete ein Spruch der Österreichischen Datenschutzbehörde dem Projekt AMAS ein (vorläufiges) Ende [DSB 2020]. Grund dafür sei die fehlende gesetzliche Grundlage zum „Profiling“ auf Basis von persönlichen Daten, sowie der Umstand, dass sich Betroffene nicht rechtswirksam darauf berufen könnten von AMS Mitarbeiter*innen in eine höhere Kategorie umgestuft zu werden oder eine Kontrolle der getroffenen Entscheidung zu verlangen. Schließlich bemängelte die Datenschutzbehörde, dass angesichts der kurzen zur Verfügung stehenden Beratungszeiten (nicht mehr als 10 Minuten in manchen Geschäftsstellen) die errechnete Einstufung „routinemäßig“ übernommen werden könnte, auch wenn AMS-Mitarbeiter*innen formal die Letztentscheidung hätten [DSB 2020]. Auf diesen Aspekt geht Kapitel 4 näher ein. Auf Basis ihres Spruchs hat die Datenschutzbehörde den flächendeckenden Einsatz des Systems ab 1.1.2021 untersagt, sofern bis dahin keine geeignete gesetzliche Grundlage für die Datenverarbeitung geschaffen werde. Das AMS hat daraufhin eine Bescheidbeschwerde eingelegt und damit gegen den Spruch berufen.⁸ Der Ausgang des Verfahrens ist bis dato offen. Der Einsatz von AMAS ist zum momentanen Zeitpunkt ausgesetzt (November 2020).

⁸ <https://futurezone.at/netzpolitik/ams-algorithmus-ams-beruft-gegen-stopp-durch-datenschutzbehoerde/401042806>

3 SOZIOTECHNISCHE ANALYSE VON AMAS

Das vorliegende Kapitel widmet sich einer detaillierten Analyse des Systems AMAS in zwei Teilen: Einerseits eine technische Rekonstruktion des Systems basierend auf den uns zur Verfügung gestellten Dokumenten in deskriptiver Form, und andererseits eine kritische Analyse des Systems in Hinblick auf die Fragen von inhärenten Bias, Diskriminierung, Fehlerraten sowie Transparenz und Nachvollziehbarkeit („*Accountability*“).

3.1 TECHNISCHE REKONSTRUKTION DES SYSTEMS

Zum Verständnis des Systems AMAS ist eine Analyse der inneren Abläufe, Zusammenhänge sowie technischer Entscheidungen bei der Gestaltung und Implementierung notwendig. Dieses Kapitel widmet sich daher der Analyse der verwendeten Daten sowie deren Ursprüngen, der verwendeten Modelle und Variablen des Systems, sowie des potentiellen Bias, welchen es mit sich bringt.

Ursprungsdaten,
Modelle und Variablen

Das System AMAS besteht aus einer Reihe an ineinandergreifenden Komponenten, welche in diese technische Rekonstruktion miteingeschlossen werden. Diese inkludieren einerseits die Datenbasis und die statistischen Modelle, die dem System zugrunde liegen, und andererseits deren Operationalisierung im konkreten Einsatzfall als Assistenzsystem im Rahmen der alltäglichen Abläufe im AMS. Hierbei ist es wichtig zu betonen, dass dies eine ganzheitliche Betrachtung des AMAS als „algorithmisches System“ darstellt, welche absichtlich weit über die rein technischen Abläufe hinausgeht, und in seiner Gesamtheit stark von technischen und prozeduralen Entscheidungslogiken des Systems abhängt.

sowohl technische
Details als auch
Operationalisierung
wichtig für
ganzheitliche Sicht auf
das System

Jedes algorithmische System basiert auf einer Datenbasis, welche entweder bereits vor der Entwicklung als solche Bestand hatte oder eigens für die Verarbeitung durch das System aggregiert wird. AMAS bildet hier keine Ausnahme, und operiert auf prä-existenten, für die Verarbeitung aufbereiteten Daten über Arbeitssuchende, welche durch ein statistisches Stratifikationsverfahren zur Voraussage der „Integrationschance“ Arbeitssuchender herangezogen werden.

prä-existente Daten
werden eingesetzt

Die Herkunft dieser Daten ist zweigeteilt: Einerseits wird das Data Warehouse des AMS aus Stammdaten des Dachverbands der Sozialversicherungsträger gespeist, und andererseits fließen auch durch das AMS selbst gesammelte Daten (Dateneingabe durch Sachbearbeiter*innen, Selbstauskunft der Arbeitssuchenden) ein [DOK_2 2020, S. 11]. Diese Datenbasis

Daten aus dem
Data-Warehouse
und der SV

wird für die Verwendung im System AMAS einer Reihe an Transformationen unterzogen – so wird etwa das spezifische Alter einer Person nicht direkt betrachtet, sondern in eine von drei Alters-Klassen kategorisiert, bevor die resultierende Zuweisung in das algorithmische System einfließt [SPEZ_1 2019]. Soweit zum jetzigen Zeitpunkt bekannt, werden keine Daten explizit nur für die Verwendung im System AMAS gesammelt.

AMSG als rechtliche Basis?

Als rechtliche Basis für die Sammlung und Verarbeitung der Daten bezieht sich das AMS auf § 25 AMSG⁹ unter Verweis auf den „grundlegenden gesetzlichen Auftrag gem. § 29 AMSG“ [BRIEF_3 2020, S. 16]. Erst auf Nachfrage durch die Studienautor*innen wurde auch eine Datenschutzfolgenabschätzung gem. § 35 DSGVO¹⁰ genannt, deren Bericht zum jetzigen Zeitpunkt – also nachdem das System bereits beinahe 2 Jahre im österreichweiten Testbetrieb lief – noch ausständig war. Die gesammelten Daten inkludieren unter anderem auch sensible¹¹ personenbezogene Daten, wie etwa „gesundheitliche Beeinträchtigungen“, welche für die weitere Analyse des Systems noch von besonderer Bedeutung sein werden. Weiters ist erwähnenswert, dass das AMS innerhalb derselben gesetzlichen Rahmenbedingungen auch berechtigt wäre, Daten über Arbeitgeber zu sammeln, diese jedoch nach derzeitiger Kenntnis keine Verwendung im System AMAS finden.

3.1.1 DAS STRATIFIKATIONS-VERFAHREN, LOGISTISCHE REGRESSION UND DEREN IMPLIKATIONEN

Stratifikation und logistische Regression

Laut der vom AMS zur Verfügung gestellten Informationen liegen dem Assistenzsystem AMAS für die operative Anwendung sowohl ein simples Verhältnismodell empirischer Beobachtungen der „Integrationschance“ (IC), als auch statistische Analyseverfahren – das „Stratifikationsverfahren“ – sowie Modelle der logistischen Regression zugrunde. Das Stratifikationsverfahren wird eingesetzt, um die verfügbaren Personendaten in Konstellationen zusammenzufassen (bestehend aus Personengruppen gleicher oder ähnlicher Personeneigenschaften). Basierend auf diesen Konstellationen wird die „Integrationschance“ einzelner Personen im Vergleich mit den restlichen Personen innerhalb derselben Konstellation gebildet. Die logistische Regression erlaubt dann, die Einflüsse einzelner Variablenausprägungen auf die Integrationschance zu berechnen. Hierbei ist erwähnenswert, dass zum Zeitpunkt des erstmaligen Öffentlichwerdens des Systems im Oktober 2018 von Seiten des AMS bzw. der Entwickler*innen des

⁹ Arbeitsmarktservicegesetz (AMSG), siehe <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10008905>

¹⁰ Datenschutz-Grundverordnung (DSGVO), siehe <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32016R0679&from=DE#d1e3548-1-1>

¹¹ „besondere Kategorien personenbezogener Daten“, vgl. DSGVO Art. 9, <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32016R0679&from=DE#d1e2066-1-1>

Systems (Synthesis Forschung GmbH) in keinem der initial verfügbaren Dokumente von einem Stratifikationsverfahren zu lesen war, sondern ausschließlich die logistische Regression beschrieben wurde. Nunmehr ist klar, dass die empirischen Beobachtungen für die Integrationschance den eigentlichen Kern des Systems darstellen, auf dem auch die Operationalisierung als Assistenzsystem basiert, und dass die Regressionsmodelle im Allgemeinen primär illustrativen Charakter für die Analyse der Konstellationen- und Variableneinflüsse für die IC haben.

AMS/Synthesis Forschung konzeptualisieren die Gesamtheit des Systems als eine Kombination aus „Informationsbasis“ und „Assistenzsystem“. Abbildung 1 (auf der nächsten Seite) zeigt eine Illustration¹² dieser beiden Teile und ihrer Bestandteile. Die Informationsbasis inkludiert die ausgewählten Personenkriterien auf Basis der empirischen Beobachtungen, sowie die Berechnung verschiedener Wahrscheinlichkeiten mittels Stratifikation und logistischer Regression. Diese Informationsbasis stellt eine ausschließlich historische Sicht auf die beschriebenen Sachlagen dar: jegliche Berechnungen und Vorhersagen, die das System macht, folgen daher der Grundannahme, dass sich zukünftige Ereignisse und Zusammenhänge ähnlich verhalten wie vergangene Beobachtungen. Wie irreführend und potenziell problematisch diese Annahme sein kann, zeigt sich derzeit etwa gerade durch den radikalen Einfluss der COVID-19-Pandemie, deren Auswirkungen auf den Arbeitsmarkt signifikant und aus der Informationsbasis des AMAS unmöglich vorherzusagen waren. Das Stratifikationsverfahren stellt eine Vergleichbarkeit ähnlicher Konstellationen her, und die Ergebnisse der logistischen Regression scheinen nach derzeitigem Stand der Kenntnis nur als deskriptive Ergänzung des Datenmaterials bzw. zur Erstellung der Hilfetexte für die Sachbearbeiter*innen verwendet werden.

Das Assistenzsystem hingegen basiert auf den empirischen Beobachtungen vergangener Ereignisse, strukturiert durch die Konstellationen der Stratifikation, und operationalisiert die Integrationschance durch Segmentierung der berechneten IC-Werte. Somit umschließt das Assistenzsystem selbst den operativen Einsatz der gewonnen statistischen Daten, sowie die Prozesse, Rahmenbedingungen und Entscheidungen, die von den Sachbearbeiter*innen unter Mitwirken des Systems getroffen werden.

Die strukturelle Teilung des Gesamtsystems in Informationsbasis und Assistenzsystem ist insofern mit Vorsicht zu betrachten, als sie suggeriert, nur im Assistenzsystem selbst würden Entscheidungen getroffen und realweltliche Konsequenzen erwirkt. Die Gestaltung der Stratifikation der empirischen Datenbasis selbst – die Teilung in vier Populationen, die 12 zeitlichen Strata sowie die Merkmals-Konstellationen selbst haben aber selbstverständlich ebenso Auswirkungen auf das Endergebnis wie die spätere Segmentierung.

Informationsbasis
inkludiert empirische
Beobachtungen und
Stratifikationsanalyse
sowie logistische
Regression

das Assistenzsystem
basiert auf
Beobachtungen aus der
Vergangenheit

alle Teile beeinflussen
die Ergebnisse

¹² Diese Illustration wurde durch die Autorinnen auf Basis der verfügbaren Dokumente erstellt.

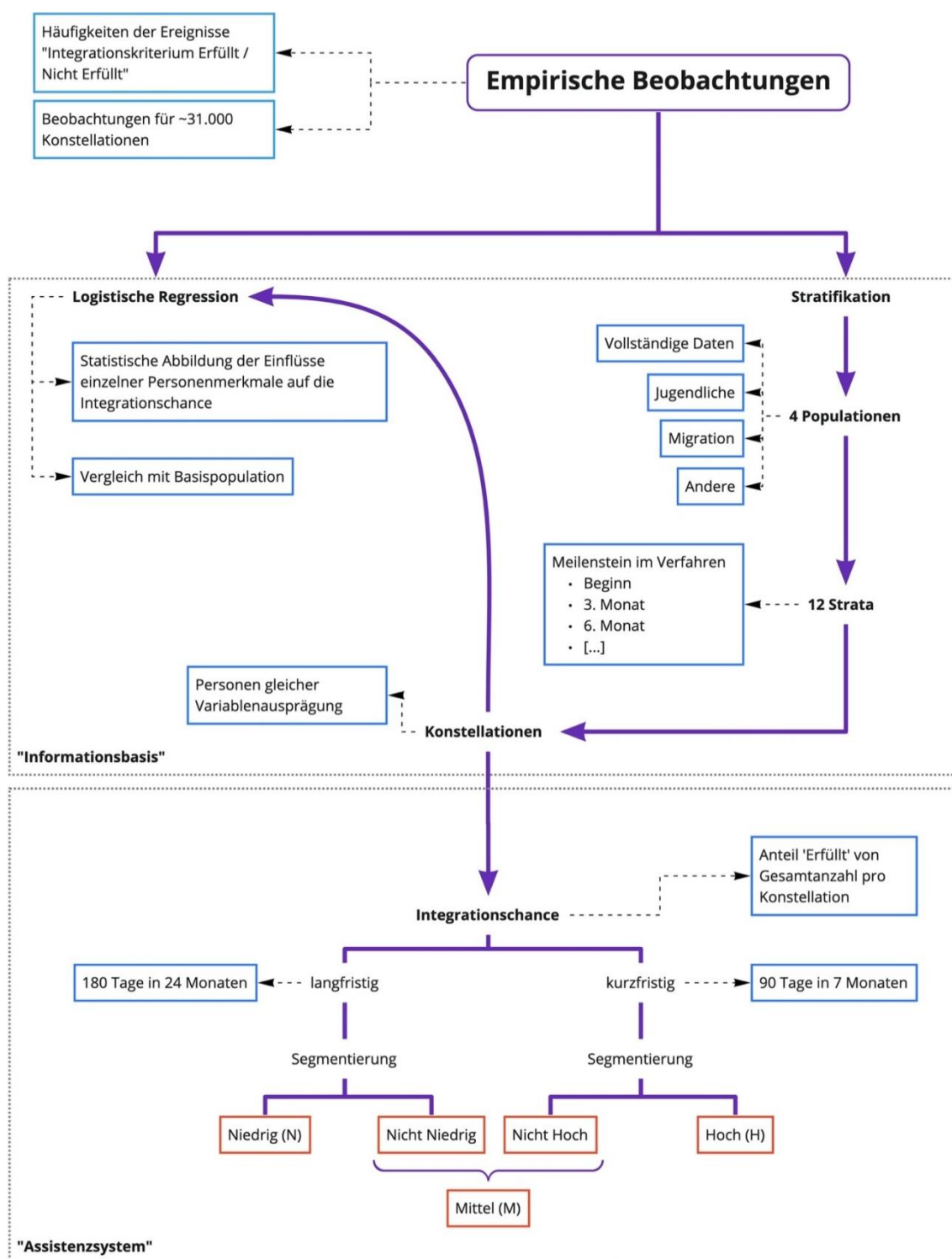


Abbildung 1: Der AMS-Algorithmus – Bestandteile, Quelle: ITA, 2020

Im Folgenden werden nun die grundlegenden Konzepte sowie die Verwendung der beiden statistischen Methoden im Rahmen des Systems erklärt.

Die Einschätzung der „Integrationschancen“ Arbeitssuchender basiert auf deren Zuteilung zu gewissen Gruppen sowie Konstellationen von Personeneigenschaften und anderen Umständen (wie etwa des „lokalen Arbeitsmarktgeschehens“). Grundlage der Einschätzung sind empirische Beobachtungen der Anteile von Personen mit gleichen Personeneigenschaften („Konstellationen“), welche ausgehend von einem zeitlichen Meilenstein im aktuellen Geschäftsfall jeweils eines von zwei Benchmark-Kriterien erfüllt haben: das „kurzfristige“ bzw. „langfristige Integrationskriterium“. Um das kurzfristige Kriterium zu erfüllen, muss eine Person zumindest 90 Tage ungeforderte Beschäftigung innerhalb von sieben Monaten ab Beobachtungs-Zeitpunkt (Meilenstein) aufweisen; für das langfristige Kriterium sind 180 Tage ungeforderte Beschäftigung innerhalb von 24 Monaten notwendig. Die „Integrationschance“ errechnet sich im System AMAS unmittelbar aus der Beobachtung dieser Anteile: Wenn beispielsweise innerhalb des Beobachtungszeitraumes der vorherigen vier Jahre von 100 männlichen Arbeitssuchenden über 50 Jahren, die unter gesundheitlichen Einschränkungen leiden, keine Betreuungspflichten haben und über einen Lehrabschluss verfügen, nur 30 Personen das kurzfristige Kriterium erfüllen, wird die kurzfristige „Integrationschance“ künftiger Personen derselben Konstellation (männlich, über 50 Jahre, mit gesundheitlichen Beeinträchtigungen, keine Betreuungspflichten, Lehrabschluss) mit „30 %“ beziffert¹³.

Das Stratifikationsverfahren erlaubt die Analyse und den Vergleich der aus den empirischen Beobachtungen gebildeten Gruppen und Konstellationen in unterschiedlichen Detailgraden. Während die Integrationschance selbst nur den direkten Vergleich einzelner Personenkonstellationen mit exakt gleichen Variablenausprägungen zulässt, können mittels Stratifikation auch größere Kombinationen – wie zum Beispiel die Verteilung der Integrationschancen auf Frauen und Männer allgemein – betrachtet werden. Diese Informationen dürften durch AMS/Synthesis Forschung hauptsächlich zur Verfeinerung des zugrundeliegenden Modells, wie etwa zur Auswahl der verwendeten Personeneigenschaften, verwendet worden sein, und könnten dem AMS Management als Analysetool hilfreich sein. Darüber hinaus werden einzelne Konstellationen durch Stratifikation aggregiert, wenn sich in den einzelnen Konstellationen keine ausreichende Anzahl an Beobachtungen finden. Der genaue Modus dieser Zusammenlegungen ist allerdings nicht bekannt.

Dem Stratifikationsverfahren gegenüber steht die logistische Regression, welche über den oben beschriebenen Informationsgewinn hinaus Einsicht in den Einfluss einzelner Merkmalsausprägungen in den jeweiligen Konstellationen geben soll. Basierend auf denselben, nominal skalierten Personen-Variablen in Abhängigkeit zu den beiden Zielvariablen „kurzfristiges“

die Integrationschance ist ein simples Verhältnis innerhalb einer Konstellation

Stratifikation analysiert auf unterschiedlichen Detailgraden

logistische Regression evaluiert den Einfluss einzelner Personenvariablen auf die Integrationschance

¹³ Das genannte Beispiel ist verkürzt, beinhaltet nicht alle vom System verwendeten Variablen und dient nur zur Illustration des Berechnungsschemas.

bzw. „langfristiges Integrationskriterium“ lassen sich somit Einflüsse einzelner Merkmalsausprägungen im Vergleich zu einer festgelegten Basisgruppe – Männer unter 30, höchstens Pflichtschulabschluss, österreichische Staatsbürgerschaft – feststellen. Die Ergebnisse sollen beschreiben, wie sich die „Integrationschancen“ zweier Personen, welche sich nur durch eine einzige Variablen-Ausprägung unterscheiden, zueinander verhalten [DOK_2 2020]. Daraus soll sich beispielsweise erkennen lassen, welchen Unterschied das Alter einer Person für den erwarteten Erfolg bei der Arbeitssuche macht, oder wie relevant das regionale Arbeitsmarktgeschehen für die „Integrationschance“ ist. Die logistische Regression ist vor allem zur Analyse der Ergebnisse des Stratifikationsverfahrens relevant; im operativen Betrieb basiert nur die Auswahl der Erklärungstexte auf den Erkenntnissen dieser Analyse. Hierbei ist anzumerken, dass es keine Hinweise für eine strukturierte, regel-basierte Erstellung dieser Erklärungstexte gibt – nur deren Einsatz (wann welcher Text angezeigt wird) ist in [SPEZ_2 2019; HAND_1 2019] beschrieben. Der genaue Zusammenhang zwischen den Ergebnissen der Regression und den Erklärungstexten ist bislang nicht bekannt. Das Erklärungs-Text-System wird im Kapitel zur Operationalisierung des Systems genauer beschrieben.

die logistische
Regression ist für die
Operationalisierung
weitgehend irrelevant

Somit steht fest, dass die logistische Regression für die operative Umsetzung der Integrationschancen-Berechnung nicht notwendig ist; die simple Auswertung der Häufigkeit der Erfüllung der Integrationskriterien innerhalb einzelner Konstellationen oder stratifizierten Aggregationen derselben ist nach derzeitigem Stand der tatsächliche Kern des Assistenzsystems.

3.1.2 VARIABLENAUSPRÄGUNGEN UND KONSTELLATIONEN

Um die Vergleichbarkeit dieser Konstellationen (für Stratifikation wie Regression gleichermaßen) zu gewährleisten, werden die verwendeten Daten zu gröberen Kategorien aggregiert; Tabelle 1, S. 29 zeigt eine Auflistung aller verwendeten Variablen sowie deren Ausprägungen im System.

nur für einen
kleinen Teil möglicher
Konstellationen gibt es
auch ausreichend
Beobachtungen

Laut [DOK_2 2020] ergeben sich aus diesen Variablen in allen möglichen Kombinationen etwa 81.000 Konstellationen. Die tatsächlich beobachteten, also unter den Arbeitssuchenden auftretenden Konstellationen, sind mit 31.000 weit geringer beziffert – es gibt also nicht für alle Kombinationen dieser Variablen-Ausprägungen auch tatsächlich Personen, die diese Variablenzuordnungen aufweisen. Auch von den 31.000 beobachtbaren Konstellationen gibt es nur für einen geringen Teil zehn oder mehr Beobachtungen (7.800 Konstellationen), und nur für 1.900 Konstellationen gibt es mehr als 50 Beobachtungen. Bei 504.000 Personen insgesamt und 310.000 Personen, deren Variablen-Konstellation mindestens von 50 anderen Personen geteilt wird, bedeutet dies, dass für knapp 39 % der Personen weniger als 50 andere Menschen als Vergleichsbasis für die Integrationschance zur Verfügung stehen; für etwa 12 % der Personen gibt es sogar weniger als 10 andere Personen mit derselben Konstellation, was die Aussagekraft der berechneten Integrationschance für diese Personenkreise massiv in Frage stellen würde. Dies ist auch den Entwickler*innen des Systems –

der Synthesis Forschung – offensichtlich bewusst: deklarieren sie doch nur jene 1.900 Konstellationen als „(statistisch) außerordentlich befriedigend“, und implizieren damit auch, dass weniger als 50 Beobachtungen für 39 % der Personen statistisch eigentlich nicht zufriedenstellen wären [DOK_2 2020, S. 13]. Um dieser Problematik zu begegnen und auch für diejenigen Personen, die von diesem Mangel an vergleichbaren Daten betroffen sind, einen IC-Wert berechnen zu können, bedient sich das System der Stratifikationsanalyse und aggregiert mehrere Konstellationen. Wie diese Aggregation durchgeführt wird – automatisch oder manuell – und nach welchen Gesichtspunkten Konstellationen zusammengelegt werden, ist nicht dokumentiert. Bekannt ist, dass so lange aggregiert wird, bis mindestens 10 Beobachtungen vorliegen. Warum bereits 10 Beobachtungen an eben jener Stelle als „ausreichend“ gesehen werden, bleibt seitens Synthesis Forschung jedoch unbeantwortet. Abbildung 2, S. 30 visualisiert die prozentualen Verteilungen der Konstellationen und Personendaten.

Tabelle 1: Variablen und deren Ausprägungen, Quelle: [SPEZ_2 2019]

Variable	Ausprägungen
Geschlecht	M/W
Altersgruppe	< 30/30-49/50+
Staatengruppe	Österreich/EU/Drittstaaten
Ausbildung	Pflichtschulabschluss/Lehre/Matura oder höher
Gesundheitlich Beeinträchtigt	Ja/Nein
Betreuungspflichten	Ja/Nein
Berufsgruppe	Produktion/Dienstleistung
RGS-Typ/Regionales Arbeitsmarktgeschehen	Typ 1 – 5
Beschäftigungsverlauf	< 75 %/> 75 % Beschäftigungstage in 4 Jahren
Frequenz Geschäftsfälle	0 GF in 4 Jahren/1 GF in 4 Jahren/2 GF in 4 Jahren (max. 1/Jahr)/3+ GF in 4 Jahren
Geschäftsfalldauer	0 GF > 6 Monate/1+ GF > 6 Monate
Maßnahmenteilnahme	0/1+ unterstützend/1+ qualifizierend/1+ beschäftigungsfördernd
Dauer der laufenden Arbeitslosigkeit/Erreichter Meilenstein	Beginn/3/6/9/12/15/18/21/24/30/36/48+ Monate

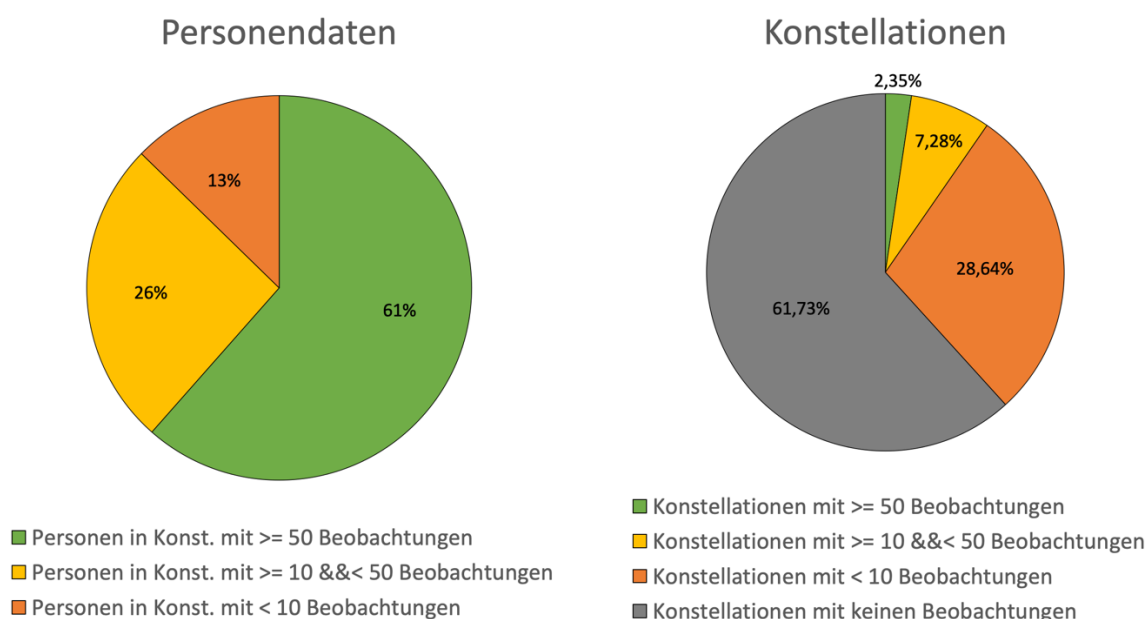


Abbildung 2: Personendaten und Konstellationen – Datenverfügbarkeit, Quelle: [DOK_2 2020, S. 13]

3.1.3 POPULATIONEN UND KATEGORISIERUNG IM ASSISTENZSYSTEM

Populationen clustern
nach
Datenverfügbarkeit

Basierend auf der Integrationschance einzelner Konstellationen werden im laufenden Betrieb Individuen Integrationschancen für das kurz- und langfristige Integrationskriterium zugewiesen. Wie in [DOK_2 2020, S. 23] ausgeführt, ist dies jedoch nicht gleichermaßen für alle Personen möglich: nicht für alle im AMS Data Warehouse erfassten Personen stehen Daten für den relevanten Zeitraum von vier Jahren in die Vergangenheit zur Verfügung. Um diesen Personen trotzdem in vergleichbaren Modellen eine Integrationschance zuordnen zu können, wird die Gesamtpopulation in vier Teilpopulationen aufgeteilt, welche jeweils eigene Modellberechnungen und Vergleiche basierend auf den vorhandenen Merkmalskonstellationen zugeordnet werden. Abgesehen von der Hauptpopulation mit vollständig verfügbaren Daten werden noch „in das Erwerbsleben einsteigende Jugendliche“, „in jüngerer Zeit (erneut) Zuwandernde“ sowie alle anderen Personen mit „fragmentierter Datenlage“ [DOK_2 2020, S. 23-24] separat behandelt.

Personen mit
fehlenden Daten sind
zuerst Migrant*innen,
dann Jugendliche

Die Zuordnung zu diesen Populationen erfolgt schrittweise: Von der Basispopulation werden all jene Personen ausgeschlossen, deren Erwerbsverlauf innerhalb der letzten 4 Jahre mindesten 150 Tage Versicherungslücken, keine Daten, Ausbildung oder sonstige erwerbsferne Positionen (mitversichertes Kind, Partner*in, sonstige Mitversicherung) aufweist, beziehungsweise diejenigen, deren beobachtete Versicherungszeiten unter 1.310 Tagen liegen [SPEZ_1 2019, S. 54; DAT_1 2019, S. 67-70, S. 74-75]. Von dieser Population werden jene Personen mit Migrationshintergrund (ausländische Staatsbürger*innenschaft, Einbürgerung oder ein ausländi-

sches Elternteil) in die Gruppe „partiell valide schätzbare Population mit Migrationshintergrund“ eingeteilt. Für diese Personen werden die beiden Merkmale Geschäftsfalldauer und Beschäftigungsverlauf nicht in die Berechnung einbezogen.

Von den restlichen Personen werden wiederum zuerst jene Personen, deren Alter unter 25 Jahren liegt, in die Gruppe „partiell valide schätzbare jugendliche Population“ eingeteilt, und für die Modellbildung zusätzlich zu Geschäftsfalldauer und Beschäftigungsverlauf auch die Staatengruppe ausgenommen (wohl da alle nicht-österreichischen Staatsbürger*innen bereits in der vorigen Population aufgenommen wurden). Weiters wird für diese Gruppe das Merkmal Altersgruppe neu definiert (unter/über 20-Jährige).

Die restlichen Personen – also jene, die weder vollständige Daten nach obigem Kriterium, noch Migrationshintergrund aufweisen und auch älter als 25 Jahre sind – werden der Gruppe „partiell valide schätzbare Population mit zuvor fragmentierter Erwerbskarriere“ zugeordnet, und für sie wird ebenfalls auf die Merkmale Staatengruppe, Geschäftsfalldauer und Beschäftigungsverlauf zur Auswertung verzichtet.

Zu dieser Aufteilung der Populationen mit unvollständigen Daten ist anzumerken, dass laut [DOK_2 2020, S. 17-20] zu Beginn des Geschäftsfalles etwa 31 % der Personen unvollständige Daten vorweisen. Die separate Behandlung von Jugendlichen mit unvollständigen Daten ist in Anbetracht der gesonderten Behandlung bei der Kategorisierung noch nachvollziehbar; der Grund für die explizite Ausklammerung von Personen mit Migrationshintergrund ist jedoch nicht klar, werden diese Personen doch genauso wie österreichische Staatsbürger*innen mit fragmentiertem Erwerbsverlauf behandelt. Auch die Reihenfolge der Aufteilung (Migrationshintergrund > Jugendliche > restliche Personen) scheint fragwürdig, führt sie doch dazu, dass Jugendliche mit Migrationshintergrund nicht zur Gruppe der Jugendlichen, sondern in diejenige mit Migrationshintergrund zugewiesen werden.

Für die Kategorisierung einer einzelnen Person im Assistenzsystem wird nun – in Abhängigkeit der oben beschriebenen vier Populationen – die Integrationschance berechnet und die Person je nach Ergebnis einer von 3 Gruppen zugeordnet [DOK_2 2020, S. 59]. Diejenigen, deren Integrationschance für das kurzfristige Kriterium 66 % oder mehr beträgt, werden der Gruppe für „hohe Chancen“ – kurz H – zugeordnet; jene, deren Integrationschance für das langfristige Kriterium 25 % oder weniger beträgt, der Gruppe N mit „niedrigen Chancen“. Zusätzlich kommen noch weitere Regeln zur Anwendung: Personen mit einer Einstellungszusage werden unabhängig von ihrer Chancenbewertung immer in die Gruppe H eingeordnet, und Personen, die aufgrund ihrer Integrationschance der Gruppe N zugewiesen wären, die aber zu „arbeitsmarktpolitischen Zielgruppen“ (wie etwa Jugendliche, s. o.) zählen, werden stattdessen zum „mittleren“ Segment M gezählt, welche sich ansonsten aus allen anderen Personen (weder in H noch N) zusammensetzt.

ein Drittel Kund*innen
ohne vollständige
Daten

Integrationschance
bestimmt
Segment-Zuweisung

Segment „hoch“ nur in den ersten drei Meilensteinen im Prozess	Aus den Definitionen der Integrationskriterien und -chancen ergeben sich weitere Beobachtungen für die Gruppenzuordnungen. Nachdem das kurzfristige Kriterium mit 7 Monaten beschränkt ist, kann eine Erfüllung innerhalb eines laufenden Geschäftsfalles nur für die Meilensteine 1 (Beginn des Geschäftsfalles) sowie nach 3 bzw. 6 Monaten prognostiziert werden. Ab dem 9. Monat eines Geschäftsfalles ist eine Zuweisung zu Gruppe H daher nicht mehr möglich.
Schwellenwerte optimiert nur für Treffsicherheit	Ebenso relevant ist die Frage nach der Wahl der beiden Schwellenwerte 66 % bzw. 25 %. Laut unterschiedlicher Aussagen seitens des AMS/Synthesis Forschung (vgl. etwa [DOK_2 2020, S. 60; PFLI_1 2016, S. 11]) wurden die Grenzwerte mit Hinblick auf die vom AMS geforderte „Treffsicherheit“ von mindestens 75 % gewählt. Bemerkenswert ist hierbei der Gegensatz zwischen der hohen Granularität der Voraussagen basierend auf Konstellationen mit Kleinstpopulationen unter 50 Beobachtungen, und der generalisierten Entscheidung für systemweit gleiche Grenzwerte.
Schwellenwerte ohne Graubereiche sind problematisch	Schlussendlich ist die Wahl von diskreten Werten ohne Graubereiche kritikwürdig. Die Aussagekraft einzelner Integrationschancen knapp um den Wert 25 % – also etwa der Unterschied zwischen 24,5 % und 25,2 %, ist unter dem Gesichtspunkt der unterschiedlichen Behandlung der Arbeitsuchenden der Gruppen N im Vergleich zur „mittleren“ Gruppe natürlich fragwürdig. Auch ist in den Handbüchern, der Dokumentation für Sachbearbeiter*innen und dem uns zur Verfügung gestellten Schulungsmaterial kein Hinweis auf besondere Vorsicht oder Aufmerksamkeit in solchen Grenzfällen zu finden.

3.1.4 MODELLANPASSUNGEN

jährliche Anpassungen am Modell	Die oben beschriebenen Modelldetails sind variabel und werden jährlich angepasst. Für das Jahr 2020 wurden etwa die Definitionen für das „vorangegangene Beschäftigungsausmaß“ um geförderte Beschäftigungsverhältnisse und Beschäftigung im Ausland erweitert; für die Erfüllung des Integrationskriteriums werden ebenfalls Beschäftigungen im Ausland einbezogen. Auch bei der Zuweisung der Merkmale „gesundheitliche Beeinträchtigungen“ und „Betreuungspflichten“ wurden kleinere Anpassungen vorgenommen (Zeitfenster, versicherungsrechtlicher Status).
Konjunkturfaktor basiert nicht auf Daten zur Wirtschaftsentwicklung!	Die größte Änderung ist jedoch die Einführung eines „Konjunkturfaktors“, welcher – anders als die Namensgebung vermuten ließe – keineswegs aus direkten Daten zur Wirtschaftsentwicklung, sondern nur aus dem Vergleich der aggregierten Integrationschancen unterschiedlicher Jahre generiert wird. Hierzu werden Trendentwicklungen der Integrationschancen der 81.000 möglichen Kombinationen über die vergangenen Jahre für eine Prognose der Entwicklung der Folgejahre eingesetzt. Das Vorgehen spiegelt somit die fragwürdige Grundhypothese wieder, eine rein empirische Beobachtung der Integrationschancen und deren Entwicklung gäbe direkten und verlässlichen Aufschluss über die Konjunkturentwicklung, ohne dass dafür externe Konjunkturdaten zum Einsatz kommen müssten. Da-

mit geht natürlich die Gefahr einher, etwaige Fehlannahmen im System nur zyklisch weiter zu tradieren, anstatt durch den Vergleich mit system-externen Datenquellen eine echte Evaluierung und Korrektur der Integrationschancen und deren Prognosen vorzunehmen. Eine detaillierte Beschreibung des statistischen Verfahrens, welches für diese Trend-Prognosen zum Einsatz kommt, fehlt in der Dokumentation ebenfalls. Bemerkenswert ist auch, dass die Verwendung des Faktors die Fehlerrate nur um wenige Prozentpunkte (-0,2 % für N, -1,8 % für H) verringert, gleichzeitig aber in den vorgelegten Beispielen [DOK_2 2020, S. 32ff] für das Jahr 2020 eine durchwegs pessimistischere Sicht auf die Chancenentwicklung präsentiert (-2 % Median der Zielerreichung kurzfristiges bzw. -3 % für das langfristige Kriterium); eine Kosten-Nutzen Analyse zur Entscheidung, diesen Faktor einzuführen, ist nicht ausgeführt.

3.2 IDENTIFIKATION KRITISCHER ASPEKTE

Die in Kapitel 3.1 vorgestellte technische Rekonstruktion des Systems AMAS wirft eine Reihe an Fragen auf. Im Folgenden diskutieren wir das dem System innewohnende Potential für Bias und Diskriminierung und analysieren es in Hinblick auf Fehlerraten und Treffergenauigkeit, Transparenz und Nachvollziehbarkeit („*Accountability*“).

3.2.1 BIAS UND DISKRIMINIERUNG

Die Einführung des AMS-Algorithmus wurde von Bedenken zu negativen Folgen durch dessen Einsatz begleitet, im Speziellen durch Fehlklassifikationen aufgrund von Bias und potenziell daraus folgender Diskriminierung (Lopez 2019; Fröhlich & Spiecker 2019). So wurde etwa das Risiko einer Benachteiligung, beispielsweise für Frauen, Menschen mit Behinderung und Frauen mit Betreuungspflichten, durch die Verwendung von sensiblen Variablen wie Geschlecht, „gesundheitliche Beeinträchtigung“ und anderen personenbezogenen Charakteristiken im zugrundeliegenden statistischen Modell thematisiert. Hierbei wurde hervorgehoben, dass die geringeren Chancen von marginalisierten Gruppen am Arbeitsmarkt bedeuten würden, dass diese Personen vermehrt als Gruppe M oder N klassifiziert werden. Das AMS konterte auf diese Kritik, dass die geplante Umsetzung sowie der Umgang mit persönlichen Attributen und produzierter Information nicht diskriminierend sei [BER_14 2019; NOTES_10 2019], da der Algorithmus einfach nur die Realität¹⁴ abbilde und dementsprechend

Einführung war
begleitet von Kritik zu
Bias und
Diskriminierung

¹⁴ Interview von „ZEIT Österreich“ mit Johannes Kopf: <https://www.zeit.de/2019/20/digitale-verwaltung-behoerden-aemter-effizienzsteigerung-probleme>.

die Zuteilungen zu Segmenten durchgeführt werde; eine Argumentation, die wir in Folge kritisch hinterfragen.

Organisatorische,
technische und
politische
Rahmenbedingungen
bestimmen die Modelle

In diesem Abschnitt zeigen wir zuerst, dass bei der Entwicklung des zugrunde liegenden statistischen Modells des AMS-Algorithmus nicht nur Treffergenauigkeit als Zielvorgabe eine Rolle gespielt hat, sondern auch organisatorische, technische und politische Rahmenbedingungen. Danach zeigen wir verschiedene Formen von Bias und Diskriminierung im AMS-Algorithmus auf und diskutieren soziale und politische Konsequenzen für Betroffene. Wir illustrieren anhand von Beispielen wie unscharfe, mehrdeutige Variablen zu verschiedenen Bias führen und diskutieren die Problematik rund um die vielen Konstellationen mit geringer statistischer Aussagekraft. Im nächsten Unterabschnitt widmen wir uns dem Bias der über die Zeit entsteht und zeigen anhand verschiedener disruptiver bzw. den Arbeitsmarkt verändernder Trends, wie z. B. COVID-19, auf, dass ein solcher Bias ein ernstes Problem für den Einsatz in der Praxis darstellt. Im Anschluss analysieren wir, wie durch die Inklusion von Variablen wie Geschlecht oder gesundheitliche Beeinträchtigung und die Aufteilung in die vier Subpopulationen anhand der Fragmentierung der Erwerbsgeschichte kumulative Benachteiligung in den Algorithmus eingebettet ist und wie dieser in Kombination mit der geplanten Ressourcenverteilungsstrategie strukturell diskriminierend wirken kann. Zum Abschluss diskutieren wir noch Konsequenzen von verschiedenen Formen der Fehlklassifikation, potenzielle Diskriminierung durch Fehlklassifikation und die in der Öffentlichkeit geäußerten Bedenken in Bezug auf Diskriminierung anhand unserer Studienergebnisse.

Die organisatorische Formung der Integrationschancen-Berechnung

Genauigkeit und
politische
Wertsetzungen

Wir zeigen in diesem Abschnitt, inwiefern während des Entwicklungsprozesses getroffene Entscheidungen über die technischen Konfigurationen des Systems die Segmentierung und die damit verbundenen sozialen Auswirkungen prägen. Das resultierende System entsteht nicht unweigerlich aus einem rein objektiven oder neutralen Handeln mit dem alleinigen Ziel, die Genauigkeit zu optimieren. Die Berücksichtigung verschiedener Abwägungen als Trade-Offs ist inhärenter Bestandteil der Entwicklung von Modellen, da sie als Abstraktionen der Welt nie deren vollständige Komplexität fassen können (Bowker & Star 1999). Entscheidungen darüber, welche Daten wie verarbeitet werden, prägen die Ausformung des Algorithmus und in Folge die Zusammensetzung der Segmentierung – und damit auch welche Bias das System aufweist. Jede dieser Entscheidungen hat somit Konsequenzen und ist, implizit oder explizit, eine Frage von politischer Wertsetzung.

keine neuen Daten für
das System erhoben

Bei der Auswahl von Variablen bzw. Merkmalen hatte das „AMS nicht völlig freie Hand“ (laut [BEGl_1 2019, S. 10]), sondern musste eine Reihe an organisatorischen Rahmenbedingungen und Anforderungen berücksichtigen. So müssen die Merkmale „auf eine gesamtösterreichisch ein-

heitliche Weise als Daten zur Verfügung stehen; und zwar sowohl nahe am aktuellen Zeitrand als auch über einen gewissen, einige Jahre zurückreichenden Zeitraum“ [BEG_1 2019, S. 10]. Damit sei es nicht möglich, einfach unmittelbar neu erhobene Daten und Kategorien zu integrieren oder Variablen explizit zu sehr lokalen Faktoren zu berücksichtigen. Dies zeigt, dass eine gewünschte österreichweite Standardisierung sowie vergangene Entscheidungen rund um Datensammlungs-Praktiken des AMS und seines Data-Warehouse die Variablenauswahl prägten und dieser nicht bloß rein technische Überlegungen zugrunde lagen. Im Umkehrschluss hat sich das AMS nicht dazu entschieden, neuartige Daten für die Segmentierung zu sammeln. Das ist insofern auch bemerkenswert, als ein erster Pilotversuch zu statistischem Profiling durch Synthesis Forschung bereits in den Jahren 2008 und 2009 durchgeführt wurde [BER_11 2009; BER_12 2011] und es somit genügend Zeit gegeben hätte, über Jahre ergänzende Daten – etwa zur Evaluation einzelner Maßnahmen – zu sammeln.

Variablen sollten weiters einen „hohen Wiedererkennungswert für die AMS-Beraterinnen/-Berater und ihre Kundinnen/Kunden haben [...] in Hinblick auf die Zielarchitektur des AMS und auf Zielgruppenbestimmungen von Förderprogrammen“ [BEG_1 2019, S. 10]. Auch diese Anforderung geht über eine technische Fokussierung auf Validität von Variablen/Kategorien und Erhöhung der Trefferquote hinaus, und zielt auf Verständlichkeit, Überzeugungskraft und Kompatibilität mit etablierten AMS Zielgruppen und Programmen ab. Das AMS setzte somit Grenzen wie „komplex“ Modelle sein konnten.

In [BRIEF_3 2020, S. 4] wurden weitere, den Algorithmus prägende Anforderungen beschrieben, wie etwa „Qualität“, „Validität“, eine „empirisch bestätigt [...] hohe Relevanz“ sowie Vertretbarkeit nach „ethischen arbeitsmarktbezogenen Gesichtspunkten“. Es wurde „bewusst darauf verzichtet, den Familienstand, Sanktionen wegen Fehlverhalten oder auch eine ehemalige ausländische Staatsbürgerschaft“ im AMS-Algorithmus zu berücksichtigen. Das bedeutet, dass die letztendlich ausgewählten Variablen, wie „Betreuungspflicht bei Frauen“ und „gesundheitliche Beeinträchtigung“, als ethisch vertretbar eingestuft wurden. Dies deutet darauf hin, dass ethische Überlegungen bei der Modellentwicklung bei Synthesis Forschung und dem AMS eine Rolle spielten, und dass diese auch die Entwicklung beeinflusst haben. Diese Überlegungen und die in diesem Abschnitt dargestellten Rahmenbedingungen und Anforderungen unterstreichen, dass nicht nur technische Überlegungen den Algorithmus und die Variablenausprägung bzw. Auswahl bestimmt haben. Die in den folgenden Abschnitten erläuterten eingebetteten Bias sind somit nicht bloß unweigerlich, sondern auch auf diverse Modellierungsentscheidungen und organisatorische Anforderung/Rahmenbedingungen zurückzuführen. Das AMS, Synthesis Forschung und andere involvierte Stakeholder haben somit eine politische Verantwortung in Bezug auf die Ausgestaltung und organisatorische Einbettung des Algorithmus.

Variablen sollen für Berater*innen „Wiedererkennungswert“ haben

auch „ethische Gesichtspunkte“ hätten Einfluss auf das Systemdesign nehmen sollen

Formen von Bias im Algorithmus

Als Bias bezeichnen wir in der vorliegenden Analyse inakurate oder verzerrende Einflussfaktoren, die zur Berechnung und Darstellung von Arbeitsmarktchancen innerhalb des AMS-Algorithmus beitragen. Wir untersuchen das System anhand verschiedener in der Fachliteratur definierter Formen von Bias in Softwaresystemen (Friedman & Nissenbaum 1996):

- *Technischer Bias*: Verkürzungen aufgrund von technischen Annahmen, Einschränkungen und Überlegungen, z. B. Probleme bei Quantifizierung von qualitativen Konstrukten wie menschlichen Werten oder Motivation.
- *Emergender Bias*: Verzerrungen, die beim Einsatz einer Software in der Praxis entstehen, z. B. erlangtes Wissen, welches nicht im technischen System berücksichtigt wurde.
- *Gegebener Bias*: Individuelle oder strukturelle Einbettung von in der Gesellschaft verankerten Werten durch Institutionen und deren Praktiken sowie Einstellungen, z. B. Gender Bias bei der Zielgruppensichtung für eine Technologie.

Anschließend diskutieren wir negative Konsequenzen von Bias und daraus potenziell folgende Diskriminierung.

Technischer Bias

Annahmen bei der
Modellierung
verursachen
technischen Bias

Jede Modellierung – das heißt, der Versuch, die komplexe Realität im Computer zu erfassen – macht Abstraktionen und Vereinfachungen notwendig, was eine Ursache von technischem Bias sein kann. Es ist jedoch wichtig festzuhalten, dass es hier Spielraum gibt und verschiedene Ausgestaltungen und Modellierungen unterschiedliche politische Konsequenzen mit sich bringen. Im Rahmen des AMS-Algorithmus wird versucht, die Komplexität des Arbeitsmarktes mit einer kleinen Anzahl an diskreten Variablen mit wenigen Variablenausprägungen zu modellieren. Mehrdeutigkeiten bzw. unscharfe Variablen wurden vereindeutigt (z. B. gesundheitliche Beeinträchtigung), fließende Variablen wurden in klar abgegrenzte Kategorien vereinfacht und schwer modellierbare Variablen ausgeklammert. Diese Modellierungsentscheidungen haben Auswirkungen auf die Zuordnung zu Konstellationen und folglich auf die berechneten Arbeitsmarktchancen und die Segment-Zuordnung. Weiters stehen diese angeführten Quellen von technischem Bias in Konflikt mit einer zentralen Annahme der Integrationschancen-Berechnung und zwar dass eine gewisse Chancenhomogenität innerhalb von Konstellationen vorliegt. Es wird also postuliert, dass Erwerbsarbeitslose in einer Konstellation gleiche bzw. sehr ähnliche Integrationschancen am Arbeitsmarkt haben und dass die gewählten Variablen dies durch eine hinreichende Differenzierung zwischen den Konstellationen ermöglichen. Die auf diesen Annahmen gestützte Integrationschancen-Berechnung führt a priori zu technischem Bias, da Erwerbsarbeitslose mit heterogenen Chancen zu den gleichen Konstellationen zugeordnet werden, wie wir in diesem Abschnitt anhand von Beispielen aufzeigen.

Beispiele für eine problematische Zusammenfassung unscharfer Variablen sind etwa „gesundheitliche Beeinträchtigung“ und „Berufsgruppe“. Bei ersterer werden unterschiedliche Beeinträchtigungen zusammengefasst, nicht aber die Auswirkungen der Beeinträchtigung für den gesuchten Beruf berücksichtigt: Rollstuhlfahrer*innen könnten etwa unterschiedliche Chancen haben, einen Bürojob oder einen Arbeitsplatz als Einzelhandelsverkäufer*in zu finden, obwohl sie derselben Konstellation zugeordnet wurden. Ähnlich verhält es sich mit den „Berufsgruppen“: Diese Variable wird in [DOK_1 2018] als dichotom mit den Ausprägungen „Produktions-“ und „Dienstleistungssektor“ beschrieben. Dabei wird eine AMS-interne Beschäftigungstaxonomie bestehend aus neun Gruppen referenziert, welche wahrscheinlich auf dem internationalen Standard ISCO-08 der *International Labor Organization* (ILO¹⁵) basiert, der auch auf der Website¹⁶ des AMS vermerkt ist. Laut dieser würde der Produktionssektor sechs separate Felder¹⁷ umfassen, während der Dienstleistungssektor weitere vier¹⁸ inkludiert. Die neunte Hauptgruppe der Taxonomie, „Hilfsarbeitskräfte“, fehlt in der Liste auffallend [SYN_1 2018]. Die grob vereinfachende Gruppierung in Produktion und Dienstleistung im System führt damit unweigerlich zur Zusammenfassung von Erwerbsarbeitslosen mit höchst unterschiedlichen beruflichen Hintergründen und steht damit im Widerspruch mit der Anforderung an Chancenhomogenität innerhalb von Konstellationen. So könnten etwa z. B. sowohl gelernte Köch*innen als auch IT-Fachkräfte als Teil des Servicesektors klassifiziert werden und in Folge Teil der gleichen Konstellation mit gleichem Chancenwert sein, obwohl diese Hintergründe ausdifferenziert in Form von zwei verschiedenen Konstellationen womöglich signifikant andere Anteils- bzw. Chancenwerte aufweisen würden.

unscharfe Variablen wie gesundheitliche Beeinträchtigung oder Berufsgruppe sind problematisch

Ein Beispiel für eine problematische Diskretisierung kontinuierlicher, realer Umstände ist etwa die Variable „RGS-Typ“, welche die Zugehörigkeit zum Arbeitsmarkt in der Umgebung der AMS-Stelle der Erwerbsarbeitslosen modellieren soll. Manche dieser Regionen liegen geographisch ausgesprochen nah beieinander und weisen grob unterschiedliche RGS-Typen auf, wie etwa die Bezirke Döbling (RGS-Typ 2) und Brigittenau (Typ 5) in Wien illustrieren. Erwerbsarbeitslose sind bei der Suche nach Beschäftigung natürlich nicht auf die im System definierte Region im Zuständigkeitsbereich einer AMS Stelle beschränkt, werden aber aufgrund eben dieser Zuweisung zur AMS Stelle bewertet. Ähnlich verhält es sich auch

Diskretisierungen wie beim RGS-Typ verursachen Fehler durch Überschneidungen

¹⁵ Für mehr Details: <http://www.ilo.org/public/english/bureau/stat/isco/isco08/> und International Labour Organization (2019).

¹⁶ Siehe <https://www.ams.at/bis/bis/IscostrukturBaum.php>.

¹⁷ Angehörige der regulären Streitkräfte, Führungskräfte, Akademische Berufe, Techniker*innen und gleichrangige nichttechnische Berufe, sowie Fachkräfte in Land- und Forstwirtschaft und Fischerei.

¹⁸ Bürokräfte und verwandte Berufe, Dienstleistungsberufe und Verkäufer*innen, Handwerks- und verwandte Berufe, sowie Bediener*innen von Anlagen und Maschinen und Montageberufe.

relevante
Informationen wurden
ausgeklammert

mit der Zuordnung zu den Altersgruppen: Die homogenisierten Chancenwerte innerhalb der Gruppen bedeuten, verbunden mit den harten Schwellenwerten zwischen Alterskategorien, dass im System zwei Personen derselben Konstellation im Alter von jeweils 30 und 49 Jahren ähnlicher bewertet werden als zwei Personen mit jeweils 29 und 30 Jahren.

Letztendlich kann das Ausklammern schwierig zu quantifizierender Umstände aus dem System zu weiteren Inhomogenitäten innerhalb einer Konstellation führen. So können unterschiedliche Motivation oder auch verschieden wahrgenommenes Auftreten (Sierminska 2015) für Erwerbsarbeitslose einen Vor- bzw. Nachteil bei der Arbeitsmarktintegration darstellen und somit signifikante Diskrepanzen in den tatsächlichen Chancen innerhalb einer Konstellation darstellen. Konkret im Fall der Gruppe der besonders Motivierten bedeutet dies tendenziell eine systematisch schlechtere Bewertung der Integrationschancen durch den Algorithmus.

Chanceninhomogenität
und wenige
Beobachtungen
innerhalb von
Konstellationen sind
problematisch

Viele der in diesem Abschnitt beschriebenen Problematiken rund um die Festschreibung von technischem Bias, sind immanente Probleme von algorithmischen Profiling-Systemen. Der Einfluss von manchen beschriebenen Formen von technischen Bias könnte zwar durch Ausdifferenzierung von Variablen reduziert werden, dies bedeutet aber wiederum, dass die Anzahl der Konstellationen erhöht werden müsste, da sich mehr Kombinationen aus Variablen und Ausprägungen ergeben. Dies hätte bei gleichbleibenden Erwerbsarbeitslosenzahlen zur Folge, dass sich eine größere Anzahl an Konstellationen mit „keine[n] Beobachtungen oder etwas weniger als zehn Beobachtungen“ ergäbe [BEGl_1 2019, S. 11], was wiederum die statistische Aussagekraft für eben diese Konstellationen begrenzen würde. Wie in Kapitel 3.1.2. ausgeführt, sind bereits jetzt etwa 12 % aller Erwerbsarbeitslosen Konstellationen mit weniger als 10 Beobachtungen zugeordnet. Es ist weiters anzunehmen, dass neue Erwerbsarbeitslose in der Praxis dann auch Konstellationen gänzlich ohne Beobachtungen zugeordnet werden, also Teil von Konstellationen werden zu denen überhaupt keine Daten aus der Vergangenheit bekannt sind. Dies könnte zum Beispiel der Fall sein, wenn Personengruppen mit gewissen Merkmalen am Arbeitsmarkt so erfolgreich sind, dass es in den letzten 4 Jahren zu keinen Beobachtungen von Arbeitslosigkeit kam und somit die neue erwerbsarbeitslose Person lediglich eine Ausnahme darstellt. Das Problem hier ist, dass das Modell nur Ein- und Ausgänge von Erwerbsarbeitslosen beim AMS berücksichtigt und nicht den ganzen Arbeitsmarkt. Dieses Phänomen könnte auch durch disruptive Trends (wie etwa die COVID-19 Pandemie) am Arbeitsmarkt auftreten, weil dann plötzlich vermehrt auch Personen mit in der Vergangenheit besonders guten Chancen, wenn ihre Branche betroffen ist, arbeitslos werden. In der Dokumentation wird überhaupt erst ab 50 Erwerbsarbeitslosen in einer Konstellation von einem „(statistisch) außerordentlich befriedigend[en]“ Resultat gesprochen (siehe Kapitel 3.1.2), womit nach den Evaluierungen 39 % aller Erwerbsarbeitslosen mit deutlich weniger statistisch gesicherten Klassifikationen konfrontiert sind. Um dieses Problem zu lösen, sollen laut Synthesis Forschung kleine Konstellationen zusammengefasst bzw. aggregiert werden. Uns ist

leider nicht bekannt, wie dies passieren soll, aber es ist wichtig festzuhalten, dass dies keine triviale Aufgabe mit einer eindeutigen Lösung ist, es voraussichtlich zu vielen Fehlklassifikationen führt und der Variablenausdifferenzierung, auf welcher das Modell basiert, widerspricht. So könnten zum Beispiel Berufsgruppen oder Ausbildungen zusammengefasst werden und damit dann einheitliche Integrationschancenwerte für Erwerbsarbeitslose mit eigentlich sehr unterschiedlichen Geschichten und Arbeitsmarktsituationen errechnet werden. Eine detailliertere Diskussion zu Auswirkungen von Fehlklassifikationen ist in Kapitel „Konsequenzen von Fehlklassifikation“ zu finden.

Emergender Bias

Emergender Bias entsteht im Laufe des Einsatzes eines Systems durch sich ändernde Rahmenbedingungen, wie etwa die Veränderungen gesellschaftlicher Werte, disruptiver Trends und außergewöhnlicher Ereignisse oder regionaler Schwankungen, auf die das System nicht im Detail reagieren kann.

Ein gutes Beispiel für die Wandelbarkeit gesellschaftlicher Werte ist etwa die dritte Geschlechtsoption, welche mittlerweile in Österreich gerichtlich anerkannt wurde (Gaigg & Simoner 2018). Diese ist im AMS-Algorithmus momentan nicht berücksichtigt. Dieses Fehlen stellt bereits eine Diskriminierung dar und birgt Potential für Bias: Derzeit ist völlig unklar, wie Personen mit drittem Geschlecht klassifiziert werden. Weiters wäre selbst bei einem sofortigen Einführen dieser Option im System aufgrund fehlender historischer Daten mit signifikanten Ungenauigkeiten zu rechnen.

Auch arbeitsmarktverändernde oder disruptive Trends haben das Potential, die Performance des Systems stark zu beeinträchtigen. Ein Beispiel sind etwa die Folgen außergewöhnlicher Ereignisse, wie etwa die Rezession im Jahr 2008 oder die durch COVID-19 ausgelösten Krise. Die Folgen der COVID-19 Pandemie im Speziellen etwa bewirk(t)en, dass am Arbeitsmarkt lange sehr nachgefragte Jobs in der (Stadt-)Hotellerie und Gastronomie plötzlich schwer vermittelbar wurden. Dieses Beispiel illustriert, dass viele der historischen Daten, auf denen der AMS-Algorithmus derzeit basiert, nun während und für Jahre nach der Pandemie wohl kaum brauchbar sind.

Nicht jedes Auftreten von emergentem Bias ist jedoch auf solche Einzelergebnisse zurückzuführen. Langfristige Trends, wie etwa die zunehmende Digitalisierung vieler Arbeitsbereiche oder die Globalisierung, können eine Verzerrung mit sich bringen, auf die das System AMAS erst mit einer deutlichen Verzögerung ‚reagieren könnte‘. Diese beiden Trends sind anschauliche Beispiele dafür, wie sich ganze Wirtschaftszweige innerhalb relativ kurzer Zeit, durch Ausgliederung mit massiven Arbeitsplatzverlusten oder Veränderungen in Anforderungsprofilen an Arbeitssuchende, wandeln können und Prognosen des Systems damit verzerrt werden.

Gesellschaft ist
laufend
im Wandel

arbeitsmarktverändern
de Trends führen zu
Bias

Veränderungen am
Arbeitsmarkt sind Teil
moderner Wirtschaft

Die Annahme, dass die „Umstände“ [DOK_2 2020, S. 9] am Arbeitsmarkt stabil seien, wodurch von der Vergangenheit auf die Gegenwart geschlossen werden könne, wird nicht nur durch globale Katastrophen, sondern auch durch regionale oder lokale Veränderungen in Frage gestellt. Diese Veränderungen sind immanente Bestandteile einer modernen Wirtschaft, die jedoch kaum explizit im AMS-Algorithmus berücksichtigt werden und damit auch das Potential für Bias in sich tragen. In den Arbeitsmarkt intervenierende Gesetzesänderungen, wie z. B. die „Aktion 20.000“ oder punktuelle Veränderungen, wie etwa der Konkurs eines größeren Unternehmens, beeinflussen individuelle Integrationschancen beträchtlich und haben aufgrund der Abhängigkeit von historischen Daten und der geringen Berücksichtigung von lokalen Faktoren erst verspätet Auswirkungen auf die Prognosen des Systems.

Gleichzeitig ist es schwierig, die Auswirkungen solcher Vorkommnisse im Sinne einer statistischen Prüfung der Qualität der Prognosen zu evaluieren. Die in den Dokumenten angeführten Evaluierungen des Algorithmus beziehen sich zum Beispiel lediglich auf 2016 bis 2018 – ein Zeitraum, der im Angesicht sowohl langfristiger, globaler disruptiver Trends als auch lokaler Einflüsse auf den Arbeitsmarkt keine robuste Evaluierung erlaubt.

emergentem Bias
muss in der
Entwicklung
vorgebeugt werden

Angemessene Gegenmaßnahmen zur Bewältigung emergenter Bias bedürfen sorgfältiger Planung vor und während des Entwicklungsprozesses. Die Strategie des AMS, auf diese emergenten Verzerrungen des Systems zu reagieren, indem die Problembehandlung rein in den Zuständigkeitsbereich der Sachbearbeiter*innen gestellt wird, ist wenig erfolgversprechend. So sollen die Berater*innen in „persönlichen Interaktion gegebenenfalls eingetretene Veränderungen (zum Besseren/zum Schlechteren) explizit berücksichtigen“ [DOK_2 2020, S. 9]. Ob dies auch externe Veränderungen wie die oben beschriebenen miteinschließt, bleibt dahingestellt. Um die fehlende Flexibilität des Systems auszugleichen, bräuchten Sachbearbeiter*innen als menschliches Korrektiv aber jedenfalls zusätzliche Ressourcen und Training sowie den nötigen Freiraum, dem Algorithmus zu widersprechen um dementsprechend reagieren zu können. Gleichzeitig könnte diese Flexibilität aber gerade durch den vermehrten Einsatz von Algorithmen in gewissen Bereichen reduziert werden. Denn mit der Automatisierung geht oft auch eine Rationalisierung von Personal einher und – in Folge – eine Reduktion von Kund*inneninteraktionen. Dadurch kann auch durch Erfahrung erlangtes Wissen („tacit knowledge“) verloren gehen (Hodosi et al. 2017). Die derzeit einzige andere relevante Maßnahme des AMS stellt der kürzlich eingeführte Konjunkturfaktor [BER_7 2019, S. 26] dar, welcher system-externe Prognosen einbezieht. Jedoch ist hier noch unklar, wie diese Prognosen erstellt werden oder wie ein einzelner Faktor akkurat die Auswirkungen kontextueller Veränderungen für 81.000 verschiedenen Konstellationen modellieren soll.

Wie in den vorherigen Absätzen ausgeführt, halten wir emergenten Bias für eine ernstzunehmende Herausforderung, deren Bewältigung komplexe Strategien und sorgfältigere Planung braucht. Die derzeitigen Maßnahmen des AMS halten wir daher für unzureichend.

Gegebener Bias

Nach Friedman und Nissenbaum (1996) entspringt gegebener Bias gesellschaftlichen Vorurteilen, Praktiken, Werten, Einstellungen oder sozialen Umständen, die unabhängig von der Entwicklung des Softwaresystems vorliegen. Dieser Bias wird von Individuen oder Institutionen entweder bewusst oder ohne Absicht perpetuiert. In diesem Abschnitt erläutern wir Ursachen für diese Form des Bias im AMS-Algorithmus und zeigen verschiedene Ausprägungen auf. Wir argumentieren, dass gewisse durch das Entwicklungsteam ausgewählte Variablen (siehe Abschnitt „Organisatorische Formung der Chancenberechnung“) und Subpopulationen sowohl direkt als auch indirekt verschiedene Ungleichbehandlungen am Arbeitsmarkt dekontextualisiert in den Algorithmus einbetten.

Wie in den letzten Abschnitten aufgezeigt, stellt der AMS-Algorithmus keine „einfache“ oder gar realitätsgetreue Abbildung des Arbeitsmarktes dar, sondern eine vielfach verzerrte. Dennoch spiegelt er auch einen „hohen Grad an Ungleichheit unter den AMS-Kundinnen/-Kunden [...] in Hinblick auf [...] Beschäftigungsintegration“ [DOK_2 2020, S. 21] am Arbeitsmarkt zu einem gewissen Grad wider. Diese Ungleichheit ist auch auf vorherrschende und historische strukturelle Ungleichbehandlungen am Arbeitsmarkt zurückzuführen, welche der Algorithmus nicht explizit ausweist. Es wird etwa z. B. nicht unterschieden, ob eine Person aufgrund von „personenbezogenen“ Charakteristika Diskriminierung oder Bevorzugung bei Einstellungsgesprächen erfährt, oder ob in einer Branche, wo Personen mit ähnlichen Charakteristika häufig tätig sind, kaum Neuansetzungen durchgeführt werden. Beide Umstände resultieren in geringeren Integrationschancen für diese Person, obwohl die Ursachen unterschiedlich sind.

Besonders gravierend ist gegebener Bias für marginalisierte Gruppen, also Personengruppen mit gewissen Merkmalen die an den gesellschaftlichen Rand gedrängt sind und damit von sozio-ökonomischer Benachteiligung betroffen sind. Diese können im AMS-Algorithmus durch personenbezogene Variablen wie Geschlecht, Staatsgruppe und gesundheitliche Beeinträchtigung abgebildet sein. Viele marginalisierte Gruppen unterliegen teils historisch langanhaltender und damit „anhäufender“, sogenannter kumulativer, Benachteiligung (Gandy 2016), welche auch in Variablen und Ausprägungen bzw. Kategorien innerhalb des AMS-Algorithmus erfasst werden. Ein Beispiel ist der langanhaltende erschwerte Zugang zu Bildung für Menschen mit Behinderung. Dieser trägt zur kumulativen Benachteiligung der marginalisierten Gruppe „Menschen mit gesundheitlichen Beeinträchtigungen“ bei, da Bildung die Chancen am Arbeitsmarkt stark beeinflussen kann und in Folge dessen auch die sozio-ökonomische Situation. Diese kumulative Benachteiligung schlägt sich auch in der gleichnamigen Variable im Algorithmus nieder. Benachteiligungen können auch intersektional wirken, sodass etwa Frauen aus Drittstaaten besondere Hürden im Vergleich zu Männern aus Drittstaaten erleben (Weichselbaumer 2016). Durch die Zusammenfassung von Erwerbsarbeitslosen mit gewissen Merkmalen zu Konstellationen wird kumulative Benachteiligung in den

gesellschaftliche
Ungleichheiten

strukturelle
Ungleichbehandlungen
werden durch
Variablen eingebettet

kumulative
Benachteiligung
marginalisierter
Gruppen

Algorithmus eingebettet. Jene Konstellationen, die marginalisierte Gruppen mit gewissen „personenbezogenen“ Charakteristika repräsentieren, werden dabei wahrscheinlicher als Teil des M oder N Segments klassifiziert. Dies wird z. B. dadurch deutlich, dass Frauen als Teil der Gesamtpopulation zu Beginn ihrer Arbeitslosigkeit ca. doppelt so oft wie Männer dem Segment N (Frauen zu 6 % und Männer zu 3 %) und halb so oft dem Segment H (Frauen zu 18 % und Männer 31 %) zugeordnet werden, wie in Abbildung 3 zu sehen ist [BER_7 2019, S. 21].

	Zielfunktion 1 (3 in 7)		Zielfunktion 3 (6 in 24)		Anteil Geschäftsfälle		Trefferquote ⁵ im Bereich	
	Zahl GF ¹	M ²	Zahl GF ¹	M ²	»N« ³	»H« ⁴	»N«	»H«
Alle Populationen zusammengefasst	617.390	45%	627.467	71%	5%	26%	84%	81%
Männer	368.710	50%	374.991	75%	3%	31%	84%	81%
Frauen	248.680	38%	252.476	65%	6%	18%	83%	80%

Abbildung 3: Merkmalsprofil der gesamten Population zu Beginn des Geschäftsfalls:

Anzahl, Median, Anteil im unteren/oberen Bereich, Trefferquote nach Geschlecht,

Quelle: Ausschnitt aus Tabelle aus [BER_7 2019, S. 21]

Während manche Personengruppen durch diese Effekte weniger Zugang zu besonders geförderten Unterstützungsmaßnahmen bekommen, führt ein ähnlicher Mechanismus dazu, dass gesellschaftlich bevorteilte Gruppen, wie z. B. autochthone Männer, deren Kompetenzen am Arbeitsmarkt eher überschätzt werden und die folglich höhere Vermittlungschancen haben können, eher falsch dem Segment H oder M zugeordnet werden.

Gegebener Bias durch Proxy-Variablen

indirekte Einbettung von Ungleichheit

Während das Potential für die Reproduktion von Ungleichbehandlung anhand von Variablen wie Geschlecht und gesundheitlicher Beeinträchtigung naheliegend und intuitiv leicht nachvollziehbar ist, können andere Variablen auch indirekt kumulative Benachteiligung von marginalisierten Gruppen einbetten. Wir erläutern dies anhand der Variable „regionaler Arbeitsmarkt“, welche die Situation des Arbeitsmarkts rund um eine AMS Geschäftsstelle modelliert: Diese beschreibt die „Zugänge von Geschäftsfällen und Geschäftsfall-Anfangsbestand“ in Relation zu den „Aufnahmen von Beschäftigung“ (voll versicherungspflichtige, ungeforderte Beschäftigung) [BER_7 2019, S. 9]. Die Variable birgt jedoch die Gefahr als Proxy (d. h. stellvertretend) für historische sozio-ökonomische Benachteiligung in AMS Regionen aufgrund deren demographischer Zusammensetzung zu fungieren. Damit kann sie implizit die kumulative Benachteiligung marginalisierter Gruppen repräsentieren. In Abbildung 4 ist durch die Wiener Bezirke farblich die Verteilung der 5 RGS-Typen dargestellt. Die Farben und „Problemdruck“-Quotienten „Q“ repräsentieren in der folgenden Reihung die Situation des Arbeitsmarkts in einem Bezirk, beginnend mit dem geringsten „Problemdruck“:

- Grün (RGS-Typ 1: Q = 0.56)

- Gelb (RGS-Typ 2: $Q = 0.68$)
- Blau (RGS-Typ 3: $Q = 0.78$)
- Grau (RGS-Typ 4: $Q = 0.89$)
- Rot (RGS-Typ 5: $Q = 1.09$)

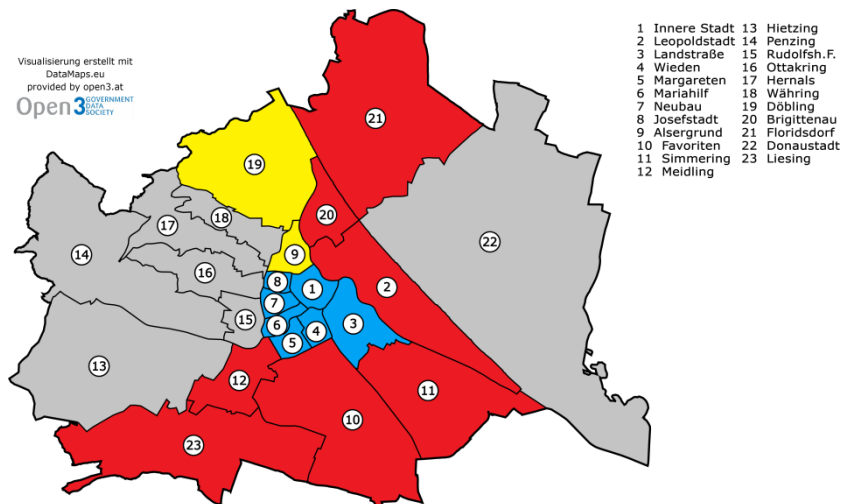


Abbildung 4: RGS-Typen Visualisierung in Wien im Jahr 2016,
Quelle: [BER_7 2019, S. 11])

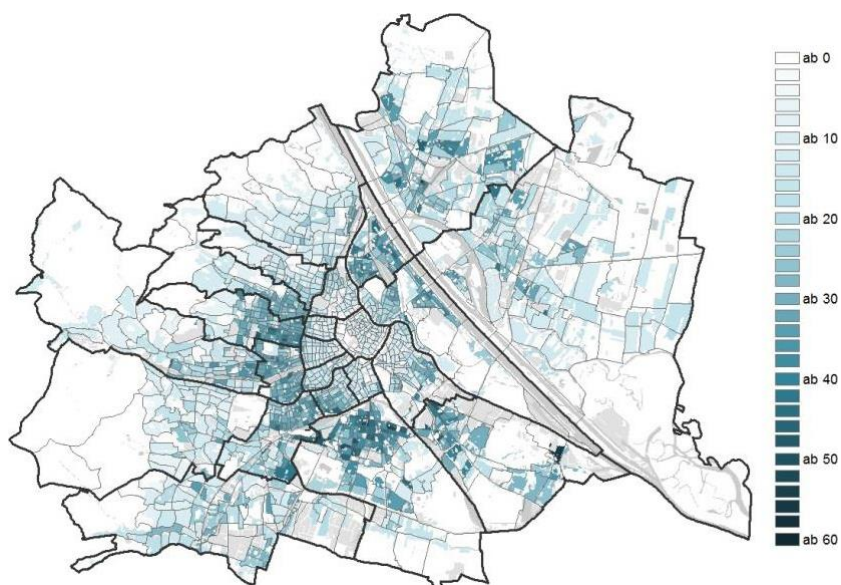


Abbildung 5: Anteil Personen mit max. Pflichtschule in Wien im Jahr 2016
Quelle: <https://www.wien.gv.at/stadtentwicklung/studien/b008541.html>

RGS-Typ als Proxy für Bildungsungleichheit

Beim Vergleich der RGS-Typen Visualisierung (Abbildung 4, S. 43) mit der Grafik zur Verteilung von Personen mit max. Pflichtschulabschluss (Abbildung 5¹⁹) zeigt sich, dass die Bezirke innerhalb des Gürtels tendenziell sowohl eine bessere Arbeitsmarktsituation aufweisen (RGS-Typ 2 und 3) als auch einen geringeren Anteil an Personen mit maximal Pflichtschulabschluss. Dieser Umstand wird im Vergleich mit dem RGS-Clustering im Jahr 2017 [BER_7 2019, S. 12] noch eindeutiger, da Bezirke mit RGS-Typ 4 (grauer Farbe) zu RGS-Typ 5 (rote Farbe) abgestuft werden. Diese Vergleiche deuten somit darauf hin, dass in den Algorithmus gewisse Bildungsunterschiede zwischen den Wiener Bezirken anhand der Variable „regionaler Arbeitsmarkt“ implizit eingebettet werden. Es würde jedoch noch detaillierte statistische Untersuchungen benötigen, um definitive Aussagen zum Ausmaß der Korrelationen mit sozioökonomischen Faktoren treffen zu können.

Gegebener Bias durch Unterteilung in Subpopulationen

Subpopulationen als Proxy für Ungleichheit

Benachteiligungen am Arbeitsmarkt werden auch implizit durch die Separation in drei Subpopulationen mit unvollständigen Datensätzen in den Algorithmus eingebettet (siehe Kapitel 3.1). Diese Populationen haben eine besonders hohe Wahrscheinlichkeit, eher dem Segment M oder N zugeordnet zu werden (Lopez 2019), da eine unvollständige Beschäftigungshistorie einen Nachteil am Arbeitsmarkt darstellen kann. Aus [DAT_3 2020] ist etwa ersichtlich, dass Personen zu Beginn ihrer Arbeitslosigkeit mit unvollständigen Daten zu 7.43 % dem Segment N zugeordnet werden. Dies ist mehr als doppelt so viel wie bei den Personen mit vollständigen Datensätzen, welche zu 3.68 % Gruppe N zugeordnet werden.

Besonders Frauen mit „Migrationshintergrund“ sind betroffen

Eine Ausnahme bilden „junge Erwachsene“ mit einem Alter unter 25, welche automatisch zur Gruppe M zugeordnet werden sollen, gleichgültig welche errechneten Chancen sie aufweisen [NOTES_10 2019]. Im Gegensatz dazu ist eine Diskrepanz besonders bei der Subpopulation mit „Migrationshintergrund“ ersichtlich, was bereits an und für sich eine problematische Kategorie ist (siehe z. B. Will 2019). Erwerbsarbeitslose dieser Gruppe sind zu Beginn ihrer Arbeitslosigkeit zu 10 % Segment N zugeteilt (siehe Abbildung 6). Besonders hervorzuheben sind auch Frauen mit „Migrationshintergrund“, welche zu 16 % dem Segment N zugeordnet werden. Bei Männern mit „Migrationshintergrund“ sind es vergleichsweise nur 7 %.

¹⁹ Bericht: Segregationstendenzen in Wien? Sozioökonomische Durchmischung 1981-2016, <https://www.wien.gv.at/stadtentwicklung/studien/pdf/b008541.pdf>

	Zielfunktion 1 (3 in 7)		Zielfunktion 3 (6 in 24)		Anteil Geschäftsfälle		Trefferquote ⁵ im Bereich	
	Zahl GF ¹	M ²	Zahl GF ¹	M ²	»N« ³	»H« ⁴	»N«	»H«
Migrationshintergrund	102.452	29%	104.596	55%	10%	10%	82%	75%
Männer	60.207	34%	61.269	59%	7%	12%	79%	74%
Frauen	42.245	24%	43.327	50%	16%	8%	83%	77%
Alter								
bis 29 Jahre	57.980	26%	58.674	52%	10%	6%	81%	75%
30 bis 49 Jahre	39.645	35%	41.282	60%	9%	16%	82%	75%
50 Jahre u. älter	4.827	37%	4.640	55%	25%	17%	86%	75%

Abbildung 6: Population mit Migrationshintergrund und ihr Merkmalsprofil zu Beginn des Geschäftsfalles: Anzahl, Median, Anteil im unteren/oberen Bereich, Trefferquote nach Geschlecht und Alter, Quelle: Ausschnitt aus Tabelle aus [BER_7 2019, S. 23].

Im Detail zeigt Abbildung 6, dass Personen „mit Migrationshintergrund“ zu 10 % dem Segment N und nur zu ebenfalls 10 % dem Segment H zugeordnet werden (verglichen mit 5 % respektive 26 % der Gesamtpopulation). Auch hier sind Frauen deutlich häufiger im Segment N klassifiziert als Männer. Ebenso ist hier ersichtlich, dass junge Erwachsene unter 25 „mit Migrationshintergrund“ dieser und nicht der Jugendlichen-Population zugewiesen werden (anders ist die hohe Zahl von beinahe 60.000 Personen nicht erklärbar). Im nächsten Abschnitt erläutern wir die sozialen Auswirkungen der verschiedenen Formen von Bias, die in den letzten Abschnitten diskutiert wurden.

Gegebener Bias und Stereotypisierung in Textbausteinen und Empfehlungspraktiken

Auch in den Erklärungstexten und deren algorithmischer Auswahl konnten wir gegebenen Bias identifizieren. Alle sieben Textbausteine für die „Niedrig“-Gruppe weisen auf Defizite hin. Das System ist also dazu programmiert, diese Gruppe lediglich über negative stereotype Ursachen zu informieren. Dabei bestimmen oft Alter, Betreuungspflichten und Beschäftigungsverlauf, welcher Textbaustein angezeigt wird – Merkmale, die durch die Betroffenen nicht oder nur begrenzt beeinflussbar sind. Es ist auch bemerkenswert, dass die Zuordnungsregeln für Personen im „Niedrig“-Segment, die nur über einen Pflichtschulabschluss verfügen, je nach Staatszugehörigkeit unterschiedlich ausfällt: Personen aus Drittstaaten werden in diesem Fall „Kompetenz-Schwächen“ attestiert, EU-Staatsbürger*innen jedoch „keine Berufsausbildung“. Der Erklärungstext für Drittstaatenangehörige erwähnt in diesem Fall noch (potentiell) mangelnde Deutschkenntnisse, obwohl diese im AMS-Algorithmus nirgends abgebildet sind und folglich auch kein Kriterium für die Zuordnung der Textbausteine sein können. Es wird somit einfach angenommen, dass Personen aus Drittstaaten mit Pflichtschulabschluss nicht ausreichend Deutsch sprechen. Personen mit einer Zugehörigkeit zu einem beliebigen EU Land sind von dieser Regel jedoch ausgenommen.

Textbausteine
reproduzieren
Stereotype

Nur Frauen werden
nach
Betreuungspflichten
gefragt

Auch im offiziellen Handbuch für Berater*innen finden sich illustrative Beispiele zu problematischen Praktiken, anhand derer die Verwendung des neuen Systems erklärt werden soll. Eine Analyse solcher Illustrationen ist relevant, da sie ein potenzielles Muster für die Praxis in den AMS Geschäftsstellen darstellen. Im Dialogfenster zur Aufstufung (siehe auch Kapitel 4) von Kund*innen etwa ist hier der Grund für das Aufstufen – nämlich, dass die Kundin ihre Betreuungspflichten „geregelt“ habe und „uneingeschränkt“ zur Verfügung stehe – besonders illustrativ. Nur bei Frauen werden Betreuungspflichten explizit im System berücksichtigt; und – wenn sie dem Segment N zugeordnet sind – wird Berater*innen der folgende Textbaustein angezeigt: „Sie haben Betreuungspflicht in ihrem familiären Umfeld“ [SPEZ_2 2019]. Dieser Dialog suggeriert, Betroffene müssten sich nun für ihre Betreuungspflichten bzw. deren „Regelung“ rechtfertigen, um aufgestuft zu werden, wenn der Algorithmus eine Zuteilung zu Segment N errechnet. Diese Praxis ist als eine algorithmisch verankerte, vergeschlechtlichte Mikroaggression (Costanza-Chock 2020, S. 44f) zu verstehen, also eine normalisierte, oft als klein wahrgenommene Äußerung mit der Marginalisierte (meist) alltäglich konfrontiert sind und die irritierend bzw. abwertend wirkt. Frauen sind z. B. auch am Arbeitsmarkt immer wieder mit dieser Frage der Betreuungspflichten konfrontiert; das AMS reproduziert mit der einprogrammierten Annahme dass Frauen* für die Kinderbetreuung zuständig sind diese Problematik, anstatt z. B. auch Männern diese Frage zu stellen und damit Betreuung durch Väter zu normalisieren, oder diese Variable und Frage als programmierten Standard abzuschaffen.

Soziale Auswirkungen von Bias und Diskriminierung

Fehlklassifikation
kann negative Folgen
für Betroffene haben

Die Falschzuteilungen zu einem der Segmente H, M oder N hat potenziell negative Konsequenzen für Betroffene. Sie kann aufgrund einer Zuteilung zu einem Segment mit „niedrigeren“ Chancen als dem tatsächlichen Segment bzw. dem Segment, in dem sich Kund*innen selbst sehen, negative psychologische Effekte mit sich bringen. Die Folge kann eine selbst-erfüllende Prophezeiung sein, bei der die Segmentierung zu Entmutigung führt, und damit die Arbeitslosigkeit der Betroffenen zementiert. Eine Zuteilung zum Segment M kann dem Ziel der Arbeitsmarktintegration durch unpassende Angebote auch entgegenwirken: Synthesis Forschung erklärt etwa, dass paradoxerweise gerade die Teilnahme an Schulungen dazu führen kann, dass die Jobsuche länger dauert: „Für diese Geschäftsfälle tritt ein „Maßnahmenzeiteffekt“ (Stichwort: Lock-in-Effekt) auf. Die betreffenden Personen haben zwar „brutto“ aber nicht „netto“ sieben Monate Zeit, um 90 Tage ungeforderte Beschäftigung zu erzielen.“ [NOTES_1 2017, S. 5] Anders formuliert verbringen die Arbeitssuchenden ihre Zeit mit Schulungen statt mit der Jobsuche und der Wiedereinstieg wird damit verzögert – auch wenn dies nur ein kurz- bis mittelfristiger Effekt sein dürfte

[NOTES_1 2017, S. 5].²⁰ Vor diesem Hintergrund ist die Praxis, nicht-Teilnahme an Schulungen in bestimmten Fällen zu sanktionieren²¹ (ohne, dass Rechtsanspruch auf Schulungen besteht), besonders problematisch.

Aus den historischen Beobachtungen ergeben sich weitere paradoxe Effekte. So werden etwa „Wiedereinsteigerinnen“ nach Karenz anfangs „unterdurchschnittliche [kurzfristige] Integrationschancen“ [NOTES_1 2017, S. 6] ausgewiesen, nach 90 Tagen aber bereits hohe kurzfristige Integrationschancen. Zu Beginn werden sie also dem „mittleren“ Segment, nach 90 Tagen dann aber dem „hohen“ Segment zugeordnet. Erklärt wird diese Beobachtung damit, dass Maßnahmen „erst nach einer gewissen Dauer ausgewählt, akzeptiert und umgesetzt“ werden [NOTES_1 2017, S. 6]. Der positive Effekt der Maßnahmenteilnahme wird also erst nach 90 Tagen bei der Prognose sichtbar. Was aber, wenn binnen 90 Tage keine Maßnahme gestartet wurde, etwa weil in diesem Zeitraum keine passende Schulung zur Verfügung stand? Durch das Aufrücken in das „hohe“ Segment könnte dann in Zukunft der Zugang zu den möglicherweise hilfreichen Maßnahmen erschwert sein.

Fehlklassifikation
kann zu längerer
Arbeitslosigkeit führen

Generell kann eine Fehlklassifikation zu Segment H oder N zur Verwehrung des Zugangs zu gewissen AMS-Förderressourcen führen, obwohl diese nötig wären. Im Fall einer Fehlklassifikation in Segment H muss entweder gewartet werden, bis eine Klassifikation in Segment M erfolgt, oder auf die Einsicht der Berater*innen gesetzt werden. Mögliche negative Konsequenzen einer solchen Fehlklassifikation werden etwa durch eine Studie aus der Schweiz ersichtlich (Schiprowski 2017), denn es zeigte sich, dass ein verpasster, früher Termin mit Berater*innen im Durchschnitt eine 10 Tage längere Arbeitslosigkeit bedeutet. Im Fall einer Fehlklassifikation zu Segment N kann anhand der Einschätzung von Berater*innen, der BBEs und dem Perspektiven-Check eine Aufstufung erreicht werden. Diese Korrekturen können trotzdem zu Verzögerungen führen und laufen Gefahr Betroffene zu entmutigen. Fehlklassifikationen können für individuelle Schicksale entscheidend sein (siehe auch Kapitel 5).

Die Segmentierung und die geplante Ressourcenaufteilung (kostenintensive Angebote für das „mittlere“ Segment, kostengünstigere Angebote für das „niedrige“ Segment) stellt eine Budgetkürzung für Personen mit als schlecht prognostizierten Chancen dar. Da historisch benachteiligte Gruppen häufiger Gefahr laufen, negativere Chancen ausgewiesen zu bekommen, kann diese geplante Praxis eine strukturelle Diskriminierung darstellen. Auch in der Fachliteratur werden statistisch angeleiteten Praktiken

²⁰ Eine Frage, die sich hier stellt ist, ob eine „blockierende“ längere Schulung zwar kurzfristig den Wiedereinstieg verzögert, in weiterer Folge aber möglicherweise zu stabileren und besseren Arbeitsverhältnissen führen kann. Solche Überlegungen sind im Übrigen in der Zieldefinition des AMS-Algorithmus (Anzahl der Tage in Beschäftigung) ebenfalls nicht abgebildet.

²¹ Kurse müssen aus rechtlicher Sicht „zumutbar“ sein – auch wenn sie für die betroffenen Arbeitslosen sinnlos erscheinen: <https://www.arbeiterkammer.at/beratung/arbeitsrecht/Arbeitslosigkeit/Zumutbarkeit.html>.

zur „Effizienzsteigerung“ ähnlich des AMS-Algorithmus als eine Form von Diskriminierung²² (Gandy 2016) verstanden. Im Fall des AMS-Algorithmus werden Informationen über Benachteiligungen am Arbeitsmarkt für das AMS als Grundlage für Einsparungen objektiviert „nutzbar“ gemacht. Potenziell höhere Kosten teurer, kreativer Unterstützung, die einen Ausgleich der Ungleichheiten für besonders benachteiligten Erwerbsarbeitslosen darstellen könnten, werden somit „eingespart“.

Eine gegebene strukturelle Diskriminierung erhöht somit die Wahrscheinlichkeit für einzelne, einem niedrigeren Segment zugeordnet zu werden. Sie laufen damit verstärkt Gefahr, die beschriebenen potenziellen negativen Auswirkungen einer herabsetzenden Fehlklassifizierung aufgrund von kumulativen Benachteiligungen zu erfahren. In Folge kann es auch häufiger zu Diskriminierung kommen, d. h. zu einer Schlechterstellung aufgrund persönlicher Merkmale. Bei der geplanten Umsetzung des AMS-Algorithmus würde dies bedeuten, weniger Zugang zu adäquaten Unterstützungsmaßnahmen zu erhalten, obwohl sie gewünscht und/oder nützlich wären, oder trotz „hoher Chancen“ Kurse besuchen zu müssen.

Die Kategorie des
Langzeitarbeitslosen

Der AMS-Algorithmus läuft Gefahr, die Kategorie des „Langzeit-Erwerbsarbeitslosen“ (Segment N) einzuzementieren. Die Klassifizierung von Arbeitssuchenden als Segment N kann einen Prozess in Gang setzen, in dem weniger Unterstützung durch das AMS zur Realisierung und somit Bestätigung der Vorhersage führen kann. Zum Beispiel können die Kürzungen für Segment N negative Auswirkungen auf Arbeitsmarktchancen von marginalisierten Gruppen haben, welche in den Algorithmus zurückgespeist werden, da dieser regelmäßig durch neu gesammelte Daten aktualisiert werden soll. Diese Feedback-Schleife kann mit der Zeit die Arbeitsmarktchancen von marginalisierten Gruppen immer weiter schwächen. Eine Reihe von Studien zu ähnlich gelagerten Systemen zeigen die von uns festgestellten diskriminierenden Mechanismen ebenfalls auf (Ensign et al. 2017; Eubanks 2018; Gandy 2016; O’Neil 2016).

Praxis ist nuanciert,
aber
Gegenmaßnahmen sind
limitiert

In der Praxis ist jedoch davon auszugehen, dass die beschriebenen Effekte des Profiling-Systems variiert und nuancierter sind, da Berater*innen z. B. von algorithmischen Empfehlungen abweichen können (siehe Kapitel „Spannungsverhältnisse auf der Ebene der Berater*innen“, S. 75). Eine Zuteilung zu Segment N kann im Einzelfall durch reduzierten Druck auf Erwerbsarbeitslose Freiräume eröffnen, die letztendlich für die Arbeitssuche förderlich sind. Auch Förderrichtlinien für Schwerpunktgruppen können die beschriebenen Effekte beeinflussen. So könnte etwa eine festgesetzte volle Unterstützung von ausgewählten Gruppen – laut AMS gilt dies etwa

²² Häufig wird diese Form der Diskriminierung in den USA am Beispiel von People of Color diskutiert, die historisch und strukturell auf vielfältige Weise benachteiligt werden. Wird diese strukturelle Benachteiligung, die sich u. a. in niedrigerem Bildungsniveau niederschlägt, bei datengetriebenen Entscheidungsfindungen (z. B. bei der Kreditvergabe) ignoriert, führt dies zu einer Diskriminierung, welche wiederum zur strukturellen Benachteiligung beiträgt und diese damit verstärkt.

für junge Erwachsene unter 25 – manche der benachteiligenden Auswirkungen der beschriebenen Feedback-Schleifen abschwächen oder diesen sogar entgegenwirken. Obwohl das AMS argumentiert, etwa durch eine etablierte Gender-Budgeting-Richtlinie speziell Frauen vor potenzieller Diskriminierung durch den Algorithmus zu schützen, ist festzuhalten, dass später die Abschaffung dieser Richtlinie angekündigt wurde (Lopez 2019). Dies zeigt die Volatilität von Richtlinien auf und unterstreicht, dass diese kein stabiles Mittel gegen diskriminierende Effekte durch Algorithmen bieten. In den nächsten Abschnitten behandeln wir verschiedene Formen von Diskriminierung anhand unterschiedlicher Merkmale.

Diskriminierung durch Diskrepanz in Fehlerraten

Viele der bisher diskutierten Arten von Fehlklassifikation (aufgrund von Vereinfachung usw.) sind innerhalb der Logik des AMS-Algorithmus nicht greifbar: Mit den verwendeten Daten lässt sich nicht herausfinden, ob eine dieser Fehlklassifikationen vorliegt oder nicht – immerhin „kennt“ AMAS beispielsweise nur zwei Berufsgruppen. Doch auch innerhalb der algorithmischen Logik und nach den Maßstäben des Algorithmus können Fehler entstehen: Wenn etwa eine Person im Sinne der Zieldefinition erfolgreich eine Arbeit gefunden hat, aber AMAS es anders prognostiziert hatte, und vice versa. Das schlägt sich in den Kennzahlen „Fehlerraten“ und „Treffer-sicherheit“ wieder. Im Abschnitt „Fehlerraten und andere Kennzahlen“ gehen wir noch detaillierter auf diese Kennzahlen ein.

Obwohl Modellierungspraktiken wie etwa die beschriebene Unterteilung der Gesamtpopulation aufgrund unvollständiger Datensätze die Treffsicherheit des Systems erhöhen können (siehe Kapitel „Fehlerraten und andere Kennzahlen“), ist festzuhalten, dass Fehlerraten zwischen Subpopulationen und Konstellationen variieren können (Barocas & Selbst 2016). Solche Diskrepanzen können auch diskriminierend wirken, da gewisse Gruppen durch systematisch höhere Fehlerraten benachteiligt werden. Das Problem der Ungleichbehandlung durch variierenden Fehlerraten wird oft auch durch „Simpson’s Paradox“ charakterisiert (Bickel et al. 1975). Es besagt, dass Qualitätsmessungen wie Präzision oder Fehlerrate auf einem Level der Populationsgranularität über verschiedene Subpopulationen genau gleich sein können, während die Qualitätsmessungen auf einem anderen Level variieren. Es gibt zahlreiche Studien, wie solche Diskrepanzen in Algorithmen identifiziert (im Englischen oft als ‚Unfairness‘ oder ‚Injustice‘, dt. „Ungerechtigkeit“ bezeichnet), untersucht und adaptiert werden können, sodass Algorithmen die gleiche oder sehr ähnliche Qualität über verschiedene Subpopulationen bieten (siehe z. B. Friedler et al. 2016; Martin Wattenberg et al. 2016; Gajane & Pechenizkiy 2017). Aus den uns vorliegenden Unterlagen geht nicht hervor, dass solche Strategien umfassend aufgegriffen wurden.

auch innerhalb der Logik des Algorithmus entstehen Fehler

Unterschiedliche Betroffenheit von Fehlklassifikationen

Niederlande: Bei einem niederländischen statistischen Profilingssystem von Erwerbsarbeitslosen (vgl. Kapitel 6.3.1, S. 89f), ähnlich dem AMS-Algorithmus, wurde etwa gezeigt, dass bei Optimierung des Systems auf Treffsicherheit, Erwerbsarbeitslose mit „Migrationshintergrund“ bis zu 2,5 mal öfter falsch klassifiziert werden, als mit dem lange etablierten regelbasierten Ansatz (Desiere & Struyven 2020). Diese Abwägung von häufiger Fehlklassifikation von marginalisierten Gruppen und Maximierung der Trefferquote über die Gesamtpopulation bedarf, wie die Autor*innen festhalten, Beachtung von Entscheidungsträger*innen und eine Entscheidung zur Priorisierung dieser Werte. Eine solche gesellschaftspolitische Debatte ist aus unserer Sicht für den AMS-Algorithmus noch ausständig.

Meilensteine	Integrationskriterium 1 (3 in 7)		Integrationskriterium 3 (6 in 24)		Anteil Geschäftsfälle		Trefferquote ⁵ im Bereich	
	Zahl GF ¹	M ²	Zahl GF ¹	M ²	»N« ³	»H« ⁴	»N«	»H«
Zu Beginn	504.054	50%	505.936	77%	4%	35%	84%	82%
Nach 3 Monaten	244.576	33%	261.531	61%	8%	14%	84%	80%
Nach 6 Monaten	143.974	21%	154.018	47%	17%	1%	84%	74%
Nach 9 Monaten	100.049	14%	107.115	38%	29%	0%	85%	69%
Nach 12 Monaten	75.708	11%	82.752	33%	36%	–	85%	– ⁶
Nach 15 Monaten	60.002	9%	66.744	30%	42%	–	86%	–
Nach 18 Monaten	49.795	8%	55.406	27%	45%	–	88%	–
Nach 21 Monaten	41.360	7%	46.758	25%	48%	–	89%	–
Nach 24 Monaten	36.009	6%	40.241	23%	51%	–	88%	–
Nach 30 Monaten	27.873	5%	31.770	17%	65%	–	90%	–
Nach 36 Monaten	22.025	4%	25.495	11%	79%	2%	92%	77%
Nach 48 Monaten	15.870	1%	13.242	8%	92%	–	94%	–

Abbildung 7: Voll valide beobachtbare Population im Verlauf des Geschäftsfalles im Überblick:
Zahl der Geschäftsfälle, Median, Anteil im unteren/oberen Bereich, Trefferquoten,
Quelle: [DOK_2 2020, S.64]

politische Debatte zu
Diskriminierung durch
Fehlerraten notwendig

In den analysierten Dokumenten werden Trefferquoten für manche Subpopulationen angegeben, z. B. für alle Frauen in einer Population, oder für einen bestimmten Meilenstein. Dabei konnten wir eine deutliche Variation feststellen: Während manche Subpopulationen bei gewissen Meilensteinen Trefferquoten bis zu 94 % erreichen [DOK_2 2020, S. 67], kommen andere nur auf 69 % [DOK_2 2020, S. 64] (siehe Abbildung 7) und liegen damit sogar unter der in der Ausschreibung geforderten Treffsicherheit von 75 % (siehe Kapitel „Fehlerraten und andere Kennzahlen“ für eine tiefergehende Analyse). In [BER_7 2019] sind Trefferquoten für die Zuteilung zu Gruppe N von sogar nur 65 % vermerkt, in diesem Fall für die unter 25-Jährigen in Tirol. Diese sind jedoch nur eine kleine Subpopulation und stellen etwa 45 Erwerbsarbeitslose dar. Bis zu 100 % Treffsicherheit haben im Vergleich die unter 25-jährige in Salzburg in Gruppe N. Diese Merkmalskombination umfasst etwa 30 Erwerbsarbeitslose. Diese großen Unterschiede in der Treffsicherheit zeigen Klassifikationsfehler auf, die besonders Subpopulationen und Konstellationen mit wenigen Beobachtungen betreffen. Im Abschnitt „Technischer Bias“ (S. 36) haben wir diese Problematik bereits behandelt und in Zusammenhang zu der gewählten Mo-

dellkomplexität gesetzt. Nach einem Verständnis von Diskriminierung als Schlechterstellung aufgrund gewisser Merkmale, sind diese teilweise deutlichen Unterschiede in der Genauigkeit als eine solche Diskriminierung zu bewerten. Es benötigt eine politische Debatte und Wertentscheidung, ob und inwieweit Diskrepanzen in der Genauigkeit zwischen verschiedenen Gruppen akzeptabel sind. Aufgrund der großen Menge an Konstellationen und dem fehlenden Zugang zu Modell und Daten, können wir keine vollständige Auflistung solcher Diskrepanzen anhand diverser Merkmale geben und haben stattdessen nur Beispiele aufgeführt. Folgestudien könnten einen detaillierteren Blick auf die beschriebenen Diskrepanzen werfen und sie politisch abwägen.

Diskriminierung aufgrund „ethnischer Herkunft“ oder von „Frauen“

Öffentlich wurde Kritik zu Diskriminierung aufgrund „ethnischer Herkunft“ laut, woraufhin AMS-Vorstand Johannes Kopf konterte, dass diese nicht explizit als Variable im Algorithmus berücksichtigt wird und dass das neue System somit nicht gegen Anti-Diskriminierungsrichtlinien verstoße [NOTES_10 2019]. Wie wir bereits erläutert haben (siehe Kapitel 3.1.3.), werden Erwerbsarbeitslose als „mit Migrationshintergrund“ klassifiziert, wenn sie eine unvollständige Beschäftigungsgeschichte aufweisen und entweder ausländische Staatsbürger*innenschaft, Einbürgerung oder ein ausländischer Elternteil vorhanden sind. Diese Merkmale stehen in starkem Zusammenhang mit „ethnischer Herkunft“ und widersprechen somit der Aussage des AMS, dass der Algorithmus diese Gruppen nicht berücksichtige. Dadurch, dass Modelle für diese Gruppen explizit errechnet werden, sind sie im Algorithmus als eigene Kategorie mit Chancenwerten zu verstehen. Die explizite Einbettung von Merkmalen zu Einbürgerung und ausländischen Elternteilen widerspricht auch Aussagen in [BRIEF_3 2020, S. 4] wonach „ehemalige ausländische Staatsbürgerschaft“ im Algorithmus nicht berücksichtigt werde. Weiters werden Erwerbsarbeitslose unter 25 Jahren mit „Migrationshintergrund“ der Subpopulation „Migrationshintergrund“ zugeordnet. In der Logik des Algorithmus werden somit Personen in erster Linie durch ihren „Migrationshintergrund“ und nicht durch ihr Alter definiert. Bei der Ressourcenverteilung macht dies laut Auskunft von Synthesis Forschung [NOTES_5 2020] keinen Unterschied, da aufgrund von Förderrichtlinien alle Personen mit einem Alter von unter 25 Jahren Zugang zu allen arbeitsmarktpolitischen Instrumenten des AMS bekommen sollten. Die Erstellung von Statistiken, die Berechnung von Integrationschancen und automatisierten Generierung von Empfehlungen können jedoch sehr wohl beeinflusst sein.

Auf eine strukturelle Diskriminierung von Erwerbsarbeitslosen „mit Migrationshintergrund“ deutet die hohe Repräsentation dieser Gruppe im Segment N hin (siehe Abschnitt „Gegebener Bias durch Unterteilung in Subpopulationen“). So sind z. B. Frauen der Subpopulation „mit Migrationshintergrund“ zu 16 % in Segment N während Frauen mit vollständig bekannter Erwerbsgeschichte nur zu 6 % in Segment N vertreten sind. Wir

Hohe Repräsentation von Migrant*innen im Segment N ist problematisch

nehmen weiters an, dass der Kategorie „Migrationshintergrund“ auch viele Geflüchtete zugeordnet sind, da diese als Neuankommende zwangsläufig einen fragmentierten Erwerbsverlauf aufweisen und ein hoher Anteil an Geflüchteten in Gruppe N bereits durch das AMS bestätigt wurde [NOTES_10 2019, Frage 7]. Damit stimmen die vorgeschlagenen Segmentierungen des AMS-Algorithmus auch teilweise mit Plänen der türkisch-blauen Regierung (2017-2019) überein, „nach Jahren der starken Fokussierung auf Förderung und Integration von Flüchtlingen“ [BER_13 2019, S. 3] nun andere in den Zielvorgaben genannte Gruppen zu stärken.

Frauen zu Beginn der
Arbeitslosigkeit
doppelt so oft in
N-Segment wie Männer

Ein weiteres großes Diskussionsthema ist die Kritik an potenzieller Diskriminierung von Frauen. AMS-Vorstand Johannes Kopf argumentierte, dass Frauen im Segment M überrepräsentiert und im Segment N unterrepräsentiert seien.²³ Jedoch, wie in Abschnitt „Gegebener Bias durch Unterteilung in Subpopulationen“ dargestellt, sind Frauen mit vollständiger Erwerbsgeschichte doppelt so oft dem Segment N zugeordnet wie Männer mit vollständiger Erwerbsgeschichte (siehe Abbildung 6, S. 45). Eine ähnliche Diskrepanz ist bei den Subpopulationen Frauen und Männer mit „Migrationshintergrund“ ersichtlich. Wir erklären den Unterschied zu den Aussagen von Johannes Kopf damit, dass diese sich möglicherweise auf ein anderes Modell beziehen als jenes zu dem wir die Dokumentation erhalten haben, andere Zeitspannen bzw. Zeitpunkte für den Vergleich herangezogen werden oder die Verteilung über alle (!) Frauen und Männer eine andere ist als für die zwei gerade aufgelisteten Subpopulationen. Damit wird klar, dass es bei einer Diskussion zu Diskriminierung essenziell ist, auf welchen Granularitätsebenen Ungleichheit untersucht wird, und dass ohne Modell und Daten eine breitere politische Diskussion zu Diskriminierung in der Öffentlichkeit auf Augenhöhe kaum möglich ist. In den beiden angesprochenen Subpopulationen der Frauen (i.e. Frauen in Segment N und Frauen der partiell valide schätzbaren Population mit „Migrationshintergrund“) kann jedenfalls von struktureller Diskriminierung gesprochen werden, da ihnen systematisch der Zugang zu kostenintensiven Ressourcen der Gruppe M erschwert wird.

Verzicht auf Variablen
keine Lösung

Wir haben in diesem Kapitel verschiedene Formen von Bias und Diskriminierung in AMAS aufgezeigt und soziale und politische Konsequenzen für Betroffene diskutiert. Wir haben analysiert, wie durch die Inklusion von Variablen wie „Geschlecht“ oder „gesundheitliche Beeinträchtigung“ kumulative Benachteiligung in das System eingebettet wird und wie dieses in Kombination mit der geplant Ressourcenverteilungsstrategie strukturell diskriminieren kann. Es ist jedoch auch wichtig festzuhalten, dass ein Verzicht auf Variablen, Ausprägungen oder Kategorien im System die genannten Probleme nicht unbedingt verbessern würde; im schlimmsten Fall könnte es sogar eine Verschlechterung bewirken, da die Genauigkeit des Systems beträchtlich sinken könnte. Es geht also nicht darum, einzelne Kategorien zu entfernen, sondern den Umgang mit Kategorien zu verbessern und zu überprüfen, ob etwa bestehende Benachteiligung am Ar-

²³ <https://www.johanneskopf.at/2018/11/14/wie-ansicht-zur-einsicht-werden-koennte>.

beitsmarkt systematisch in geringere Unterstützungswürdigkeit übersetzt und damit strukturell vorhandene Ungleichbehandlung weiter verschärft wird.

Fehlerraten und andere Kennzahlen

Treffsicherheit und Fehlerraten spielen sowohl bei der Entwicklung als auch bei der Debatte über den AMS-Algorithmus eine entscheidende Rolle. Bereits bei der Ausschreibung des Systems wurde eine gewisse Treffsicherheit (von mindestens 75 %) des Systems verlangt [AUSSCHR_1, 2015]. Dies wohl, weil diese (und weitere) Kennzahlen häufig zur Evaluierung von Klassifikationssystemen herangezogen werden. Entsprechend wurde diesen Kennzahlen auch in der Dokumentation umfangreich Platz eingeräumt und sie fanden ebenso Eingang in Diskussionen über das System, etwa in internationalen Vergleichen. In diesem Abschnitt werden wir diese Kennzahlen und deren Funktion in der Entwicklung und Dokumentation einer genauen Analyse unterziehen.

Im Kern handelt es sich beim AMS-Algorithmus um ein Klassifikationssystem: Erwerbsarbeitslose werden als einer von drei Gruppen zugehörig klassifiziert. Unabhängig von der Methode der Klassifikation sind in der Informatik und Statistik eine Reihe von Kennzahlen geläufig, welche der Beurteilung des Klassifikators dienen. Denn eine Vorhersage kann sich als richtig oder falsch herausstellen: Richtig im Sinne von AMAS wäre zum Beispiel, wenn sich herausstellt, dass ein*e Erwerbsarbeitslose*r des Segments mit „niedriger Integrationschance“ wie erwartet binnen 24 Monaten weniger als 180 Tage ungeförderter Erwerbsarbeit aufweist. Eine falsche Klassifikation wäre demnach, wenn dieselbe Person wider Erwarten mehr als 180 Tage ungeförderter Erwerbsarbeit binnen 24 Monaten erreicht. Tabelle 2 stellt dies anhand weiterer fiktiver Beispiele anschaulich dar:

75 % Treffsicherheit
als Anforderung

AMAS ist ein
Klassifikationssystem

Tabelle 2: Fiktive Beispiele für korrekte und inkorrekte Segmentzuordnungen und Prognosen

Tage ungeförderter Erwerbsarbeit in ...		Segment	Vorhersage	Korrektheit
7 Monaten	24 Monaten			
120 Tage		H	H	korrekt
100 Tage		H	M	inkorrekt
20 Tage	100 Tage	N	N	korrekt
	200 Tage	M	N	inkorrekt
80 Tage		M	H	inkorrekt
0 Tage	150 Tage	N	M	inkorrekt
95 Tage		H	N	inkorrekt
85 Tage	185 Tage	M	M	korrekt

Die Tabelle zeigt für fiktive Beispielpersonen die tatsächlich erreichte, ungeförderter Erwerbsarbeit und daraus folgende Segment-Zuordnung, eine Segment-Vorhersage für diese Person sowie die Korrektheit dieser Vorhersage. Die erste Zeile etwa stellt ein Beispiel für eine Person dar, welche in 7 Monaten 120 Tage ungeförderter Erwerbsarbeit erreicht hat. Damit

erfüllt sie die kurzfristige Perspektive (mindestens 90 Tage in 7 Monaten) und sollte dem Segment H zugeordnet werden. Die Vorhersage H ist damit korrekt. Die Person in der zweiten Zeile hat zwar nur 100 Tage erreicht, erfüllt damit aber dennoch die Definition für Segment H. Da die Vorhersage aber M ist, ist sie inkorrekt.

Begriffsklärung

Treffergenauigkeit ist der Anteil der korrekt klassifizierten Datenpunkte

Leider sind die Definitionen der Kennzahlen nicht konsistent in allen (Teil-)disziplinen der Informatik und Statistik. Um Missverständnisse und Unklarheiten zu vermeiden, seien hier einige Begriffe definiert. Im Bereich der algorithmischen Klassifikation beschreibt die Treffergenauigkeit (engl. *Accuracy*) den Anteil der Datenpunkte, der korrekt klassifiziert wird. Das Gegenstück zur Treffergenauigkeit ist die Fehlerrate (engl. *Error Rate*), also der Anteil der inkorrekt klassifizierten Datenpunkte.

Definition von „Relevanz“ und „Sensitivität“

Bei einer Klassifikation können Fehler unterschiedlicher Art auftreten. Betrachtet man etwa das „Niedrig“-Segment, dann können Personen fälschlich diesem Segment zugeordnet werden, obwohl sie eigentlich zum Segment „hoch“ oder „mittel“ gehören. In diesem Fall spricht man von Fehlern 1. Art (engl. *False Positives*). Andererseits können Personen des „Niedrig“-Segments fälschlich den Segmenten „hoch“ und „mittel“ zugeordnet werden. Hier handelt es sich dann um sogenannte Fehler 2. Art (engl. *False Negatives*). Die Kennzahl „Relevanz“ (engl. *Precision*) gibt Auskunft darüber, wie viele der Datenpunkte eines Segments korrekterweise dort sind, z. B. wie viele dem Segment „niedrig“ zugeordnete Personen tatsächlich die Definition des Segments erfüllen. Dem gegenüber bezeichnet die Kennzahl „Sensitivität“ (engl. *Recall* oder *True Positive Rate*), wie viele der Datenpunkte, die die Definition für ein Segment erfüllen, korrekterweise diesem zugeordnet wurden. Anders gesagt gibt die Relevanz Aufschluss über die Häufigkeit von *False Positives*, wohingegen die Sensitivität Aufschluss über die Häufigkeit von *False Negatives* zulässt.

KENNZAHLEN AUF EINEN BLICK

Treffergenauigkeit: Der Anteil aller Datenpunkte der korrekt vorhergesagt wurde (über alle Gruppen hinweg).

Fehlerrate: Der Anteil aller Datenpunkte, der fehlerhaft vorhergesagt wurde (über alle Gruppen hinweg).

Relevanz: Der Anteil eines Segments, der korrekt vorhergesagt wurde, z. B. wie viele dem Segment „niedrig“ zugeordnete tatsächlich die Definition des Segments erfüllen.

Sensitivität: Der Anteil einer Gruppe der korrekt als dieser zugehörig vorhergesagt wurde, z. B. wie viele jener, die die Definition des Segments „niedrig“ erfüllen wurden tatsächlich diesem Segment zugeordnet.

Kennzahlen lenken die Entwicklung

Bereits in den Ausschreibungsunterlagen für den AMS-Algorithmus hat das AMS Gütekriterien für jedes Modell (kurz- oder langfristig) vorge-schrieben. Konkret müssen die Modelle so entwickelt werden, „dass eine Treffsicherheit des Modells im Durchschnitt von mindestens 75 % erreicht wird“ [AUSSCHR_1 2015, S. 6]. Weiters solle die Treffsicherheit auch „pro Bundesland und [Regionalem Arbeitsmarktgeschehens-Typ] nach Geschlecht, pro Bundesland und Altersgruppen als auch nach den Kund/-innen-segmenten („hohe“, „mittlere“, „niedrige“ Integrationschance) für die Zugänge des Jahres 2014“ angeben [AUSSCHR_1 2015, S. 7]. Erst im späteren Pflichtenheft wird dieser Begriff genau definiert [PFLI_1 2016, S. 29] und entspricht der oben definierten Treffergenauigkeit. Bei der geforder-ten Treffsicherheit pro Segment muss es sich hingegen entweder um die Relevanz oder um die Sensitivität handeln. Auch die Definition im Pflich-tenheft ist hier nicht vollständig eindeutig: „Basierend auf der Treffsicher-heit von Einzelpersonen, kann die Treffsicherheit von Teilpopulationen durch die Summierung von Personen mit richtiger und falscher Prognose berechnet werden.“ [PFLI_1 2016, S. 61] Diese Ausführung erlaubt unter-schiedliche Interpretationen, sodass sowohl Relevanz als auch Sensitivität zutreffen können. Mehrere Definitionen für die „Trefferquote“ der Seg-mente H und N in der Dokumentation entsprechen der Relevanz und es ist davon auszugehen, dass es sich bei segmentspezifischen Werten i.d.R. darum handelt (z. B. [DOK_2 2020, 62; DOK_1 2018, 12]), was für [DOK_1 2018] auch in einem Telefonat mit Synthesis GmbH bestätigt wurde.

Wie bereits erwähnt, haben einzelne Kennzahlen (Relevanz, Sensitivität, Treffergenauigkeit, etc.) für sich genommen begrenzte Aussagekraft. Erst in Kombination ergibt sich ein detailliertes Bild. Wird etwa nur die Re-levanz für ein Segment angegeben, kann dieser Wert über verschiedene Wege hergestellt werden, wie Abbildung 8 (S. 56) veranschaulicht. In bei-den Vorhersagen ist die Relevanz für Gruppe H 80 %, gleichzeitig variiert aber die Sensitivität deutlich: In einem Fall (b) ist die prognostizierte Gruppe H sehr groß und es wurden nur wenige Personen fälschlicher-weise als nicht-H klassifiziert. Im anderen Fall (c) ist die Gruppe H klein und ein beträchtlicher Teil ist fälschlich als nicht-H klassifiziert. Während die Relevanz für Gruppe H jeweils 80 % beträgt, unterscheidet sich die Relevanz der Gruppe nicht-H beträchtlich. Daher ist es kritisch zu sehen, dass seitens der Synthesis Forschung GmbH durchgängig nur Werte für die Gruppen H und N ausgewiesen werden.

Diese Praxis hat auch weitere Auswirkungen. Bei der Entwicklung der Modelle wurde zwischen unterschiedlichen Fehlertypen unterschieden, wie eine Antwort des AMS an die Gleichbehandlungsanwaltschaft deut-lich macht: Die Grenzwerte wurden, bei vergleichbarer „Treffsicherheit“ (d. h. Relevanz) so gewählt, dass Fehler tendenziell eher dem Segment M als H oder N zugeordnet werden [NOTES_10 2019, Frage 13]. Der Fokus auf die Kennzahl Relevanz der Segmente H und N eröffnete hier einen Spielraum während der Entwicklung, der es bei gleichen Kennzahlen er-laubte Segmentzuordnungen zu variieren.

Gütekriterien für
verschiedenen
Merkmale bei
Ausschreibung
festgesetzt

Treffsicherheit für
Gruppe M nicht
bekannt

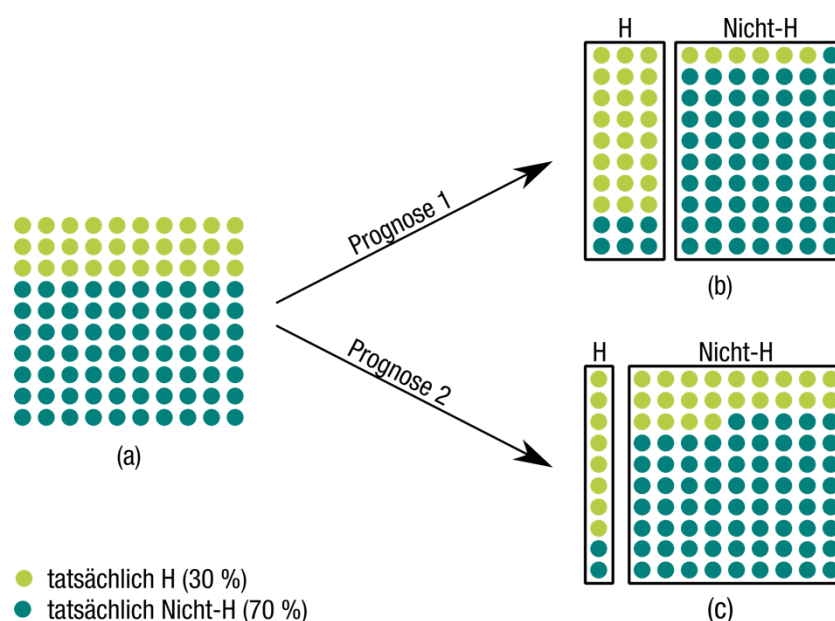


Abbildung 8: Begrenzte Aussagekraft einzelner Kennzahlen, Quelle: ITA, 2020

Abbildung 8 veranschaulicht, dass bei einer bestimmten Verteilung von H und nicht-H Personen (a) unterschiedliche Prognosen der IC-Werte zu verschiedenen Ergebnissen führen können (b und c). Dabei zeichnen sich die beiden hier dargestellten Möglichkeiten dadurch aus, dass sowohl (b) als auch (c) eine Relevanz von 80 % für die Gruppe H aufweisen. Wie deutlich zu erkennen ist, sind die Fehler in der Gruppe nicht-H stark unterschiedlich. Daher hat die Kennzahl Relevanz, vor allem wenn nur für 2 der 3 Segmente ausgewiesen, nur begrenzte Aussagekraft über die Art und Häufigkeit der Fehler.

Nicht alle Personen werden bei der Berechnung der Indikatoren berücksichtigt

mangelnde
Aussagekraft
veröffentlichter
Kennzahlen

Neben der beschränkten Aussagekraft der ausgewiesenen Indikatoren ist eine weitere Entscheidung von AMS und Synthesis Forschung zu hinterfragen. So werden nur jene Konstellationen für die Berechnung der Treffsicherheit herangezogen, für die mindestens zehn Beobachtungen vorliegen [DOK_2 2020, S. 62].²⁴ Anders gesagt, die Korrektheit bzw. Inkorrekt-

²⁴ An dieser Stelle ist es interessant, dass bereits in der Ausschreibung festgehalten wurde, dass Personen mit lückenhaften Daten („z. B. keine Beschäftigungskarriere; bei Jugendlichen, Berufseinsteiger, ev. Wieder-Einsteigerinnen, Auslandszugänge, etc.“, [AUSSCHR_1 2015, S. 7] von der Modellgüterechnung ausgenommen werden können. Nach unserem Verständnis findet das allerdings nicht statt: Diese Personen werden stattdessen den drei Populationen mit fragmentiertem Erwerbsverlauf zugeordnet. Sollten Personen derart von der Modellgüterechnung ausgenommen werden, dann seien „entsprechende Vorkehrungen in der Modellentwicklung einzubauen. Die Personen in dieser Gruppe „unvoll-

heit von 12 % aller Prognosen wurde aus der Berechnung der Indikatoren ausgenommen. Vermutlich liegt es daran, dass Synthesis Forschung Konstellationen mit 10-49 Beobachtungen als „gerade ausreichend“ [DOK_2 2020, S. 13] und erst jene mit 50 oder mehr Beobachtungen als „(statistisch) außerordentlich befriedigend“ betrachtet. Ebenfalls korrekt stellt die Synthesis Forschung fest, dass bei Merkmalskombinationen mit wenigen Beobachtungen zufällige Ereignisse „die zu erwartende Arbeitsmarktintegration [...] in hohem Ausmaß“ beeinflussen [DOK_2 2020, S. 13]. Wie bereits erwähnt, werden daher zur Berechnung der IC-Werte mehrere Konstellationen zusammengelegt bis mindestens 10 Beobachtungen vorliegen. Das Vorgehen zur Berechnung der Trefferquote (ignorieren von Konstellationen mit wenigen Beobachtungen) weicht also vom Vorgehen der Berechnung der IC-Werte ab und stellt die Aussagekraft der dokumentierten Indikatoren für die tatsächliche Praxis in Frage.

Modellevaluation

Die Synthesis Forschung GmbH führte auch eine Evaluation des Modells, das AMAS zugrunde liegt, durch [DOK_2 2020]. Das heißt, es wurde evaluiert, ob die Variablen und deren Ausprägungen aussagekräftig gewählt wurden. Dazu wurde wiederum die logistische Regression herangezogen und für alle Schwellenwerte zwischen 0 und 100 % errechnet, wie oft die Prognose richtig ist, dass Personen die kurz- bzw. langfristige Integration schaffen. Ähnlich wie oben dargestellt, kann man nun den Blick auf zwei Kennzahlen richten: Wie hoch ist der Anteil der Personen, für die korrekterweise vorhergesagt wurde, dass sie die Integration schaffen (*True Positive Rate*)? Und wie hoch ist der Anteil der als erfolgreich prognostizierten Personen, die aber die Integration nicht schaffen (*False Positive Rate*)?

Ins Extreme getrieben kann die Prognose sein, dass *alle* die Integration schaffen (dann sind beide Kennzahlen 100 %), oder aber es kann prognostiziert werden, dass *keine*r* die Integration schafft (dann sind beide Kennzahlen 0 %). Ein perfektes Modell würde gleichermaßen 100 % *True Positives* und 0 % *False Positives* erreichen. Setzt man beide Kennzahlen für alle Schwellenwerte zueinander ins Verhältnis und stellt sie in einem zwei-dimensionalen Koordinatensystem dar, so entsteht eine Kurve. Je größer die Fläche unter dieser Kurve ist (*Area under the Curve*, AUC), desto besser ist das Modell wie in **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellt. Dabei werden für das AMAS-Modell AUC-Werte von 0,83 für das kurzfristige und 0,82 für das langfristige Modell angegeben [DOK_2 2020, S. 48-49].

Kompromiss bei der Wahl des Anteils richtig zu falsch klassifizierter Erwerbsarbeitsloser

ständige Informationen“ sind auszuweisen (eigene Kategorie, Markierung, etc.).“ [AUSSCHR_1 2015, S. 7] Eine analoge Ausweisung von Personen für deren Merkmalskombinationen wenige Beobachtungen vorliegen gibt es unseres Wissens nicht.

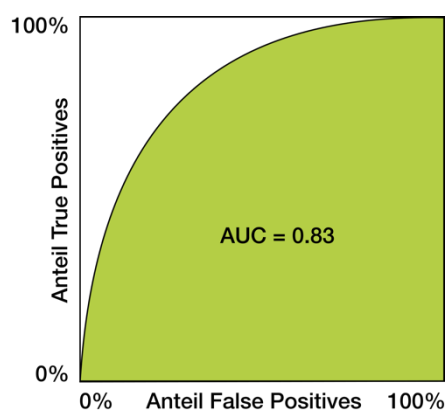


Abbildung 9: Evaluation der Logistischen Regression,

Quelle: reproduziert nach [DOK_2 2020, S. 48]

Abbildung 9 bildet für alle Schwellenwerte die Prognosekraft des Modells für die kurzfristige Integrationsperspektive ab. Je größer die grüne Fläche, desto besser. Leider ist diese Evaluierung nur begrenzt als Evaluierung von AMAS geeignet. Erstens errechnet sich die Segmentierung aus zwei Integrationschancen-Werten. Zweitens wurde hier mit logistischer Regression (ohne Interaktionsterme²⁵) gearbeitet, die aber wie bereits ausgeführt, nicht zur Berechnung der Integrationschancen herangezogen wird.

Fragliche Eignung von Kennzahlen als (primäres) Qualitätsmerkmal

Eignung der
gewählten Kennzahlen
ist zu hinterfragen

Wie wir deutlich gemacht haben, bieten technische Indikatoren wie Treffergenauigkeit und Relevanz einige Fallstricke und sind selten so klar und eindeutig sind wie ihre mathematische Grundlage suggeriert. Doch auch darüber hinaus sind diese quantitativen Metriken problematisch. Erstens dienen sie häufig als Güteindikator und sollen die Qualität der Methode vermitteln: Nicht ohne Grund wurde eine minimale Treffergenauigkeit von 75 % im Pflichtenheft festgehalten. Unter Verweis auf die hohe Treffergenauigkeit wurde außerdem wiederholt die Einführung des AMS-Algorithmus verteidigt – und auch die Volksanwaltschaft schließt sich implizit diesem Narrativ an, wenn sie die hohe Treffergenauigkeit als Gegenargu-

²⁵ Bei logistischer Regression gibt es sogenannte „Haupteffekte“ von Variablen, die durch die Koeffizienten der einzelnen Variablen wie in [DOK_1 2018, 11] dargestellt werden. Es kann aber auch sein, dass zwei Variablenausprägungen wie zum Beispiel „über 50“ und „gesundheitlich beeinträchtigt“ zusammen einen anderen Effekt haben als sich nur durch die zwei Haupteffekte darstellen lässt. Das sind sogenannte Interaktionseffekte und können explizit durch weitere Variablen, sogenannte Interaktionsterme, dargestellt werden. Je mehr Variablenkombinationen berücksichtigt werden, desto näher kommen die durch die logistische Regression berechneten IC-Werte an jene durch das Stratifikationsverfahren heran. Das wurde teilweise in [DOK_3 2020] dokumentiert. Laut [NOTES_5 2020] wurden für die dokumentierte Modellevaluation in [DOK_2 2020] keine Interaktionsterme verwendet.

ment zu Zweifeln an der Aussagekraft des Modells ins Feld führt [BER_14 2019, 100]. Dabei ist jedoch unklar, ob und weshalb Treffergenauigkeiten im (weiten) Bereich von ca. 75-90 % als „gut“ eingestuft werden sollen, besonders wenn hunderttausende Personen²⁶ betroffen sind. Ähnlich verhält es sich mit den Ergebnissen der Modell-Evaluierung. Zweitens werden Indikatoren häufig zum (quantifizierbaren) Vergleich verschiedener Systeme herangezogen, was ihre Bedeutung in öffentlichen Debatten noch weiter verstärkt. Beispielhaft war dies etwa in einem OECD-Report über statistisches Profiling im Bereich der arbeitsmarktpolitischen Einrichtungen von OECD-Staaten der Fall (Desiere, Langenbucher & Struyven 2019). Wie in den Ländervignetten in Kapitel 6 dargestellt, unterscheiden sich die verschiedenen Systeme teilweise beträchtlich. Dennoch wurde für alle Systeme der Indikator „accuracy“ ausgewiesen, der (ver)einfach(t)e und intuitive Vergleiche dieser Systeme ermöglicht. Dies geht so weit, dass aus dieser Kennzahl „Lob“ für den AMS-Algorithmus abgeleitet wurde (Szigetvari 2019²⁷). Darüber hinaus ist anzuzweifeln, ob die Zahl für AMAS tatsächlich dem englischen Begriff „Accuracy“ entspricht, oder doch korrekt als „Precision“ zu übersetzen wäre – und damit die Vergleichbarkeit noch weiter in Frage gestellt wird.

Konsequenzen von Fehlklassifikation

Die Unterscheidung der Fehlertypen ist im Kontext des AMS-Algorithmus essentiell, da je nach Fehlertyp für AMS-Kund*innen andere Ressourcen verfügbar werden oder nur noch erschwert zugänglich sind. So sind zum Beispiel die Angebote für das Segment M nur noch erschwert zugänglich, wenn ein*e Kund*in fälschlich dem Segment H oder N zugeordnet wird. Andererseits sind die Angebote für Gruppe M plötzlich leichter zugänglich, wenn eine Person fälschlich der Gruppe M statt H oder N zugeordnet wird. Gerade deshalb ist die oben beschriebene lückenhafte Dokumentation der Fehlerraten zu kritisieren.

Schlussendlich muss die Treffsicherheit des Klassifikationssystems auch unter dem Gesichtspunkt der direkten Auswirkungen einer Fehlklassifikation im Einzelfall betrachtet werden. Vonseiten der Synthesis Forschung wird hierzu präventiv eine Argumentation für die Unbedenklichkeit von Fehlklassifikationen geliefert: Fehlklassifikation in das Hoch-Segment ziehe „nur“ eine längere Arbeitslosigkeit nach sich, die vom System durch das Eintreten späterer Meilensteine wieder korrigiert werde; im „Niedrig“-Segment entstünde Personen durch die früher gelungene Arbeitsaufnahme ebenso kein Schaden, da sie ja dadurch ohnehin den Wirkungsbereich des Assistenzsystems verließen [DOK_2 2020, S. 61]. Beide Argumente sind

Unterscheidung der
Fehlertypen ist
essentiell

direkte Auswirkungen
einer Fehlklassifikation
im Einzelfall
einbeziehen

²⁶ Für 500.000 Personen, die pro Jahr betroffen sind, wären bei eine Treffergenauigkeit von 75% bis zu 125.000 Personen fehlerklassifiziert; selbst bei 90% sind dies immer noch 50.000 Personen.

²⁷ <https://www.derstandard.at/story/2000096564832/teilt-arbeitslose-nach-chancen-ein-oecd-lobt-ams-algorithmus>

nur schwer haltbar, ignorieren sie doch den Einfluss der Segment-Zuordnung auf die verfügbaren Maßnahmen. Eine Person der „hohe“ Chancen attestiert werden, hat einen erschwerten Zugang zu potentiell relevanten (Weiterbildungs-)Maßnahmen, zumindest bis es zu einer Abstufung kommt. Im „Niedrig“-Segment hingegen ist der psychologische Effekt der Fehlzurordnung von besonderer Relevanz: Für Personen in ohnehin schwieriger Lebenslage mag das autoritative Attestieren „niedriger“ Chancen durch das System zu besonderer Demotivation oder gar dem Aufgeben der Bemühungen führen. Eine Fehlklassifikation mag in der Beschreibung von Synthesis zwar willkommen sein, da ja dann eine Anstellung gefunden wurde – aber viel wahrscheinlicher scheint es, dass das System durch den Akt der Klassifikation alleine bereits Einfluss auf das Ergebnis nimmt, und somit eine selbsterfüllende Prophezeiung darstellt. Diese Überlegungen finden in den gelieferten Unterlagen keinerlei Erwähnung und bedürfen eines ernsthafteren Umgangs mit menschen-zentrierter Evaluation des Konzeptes für das System AMAS.

3.2.2 NACHVOLLZIEHBARKEIT UND TRANSPARENZ

Wie bei allen algorithmischen Systemen, deren Einsatz direkte Auswirkungen auf große Personenkreise und potentielle negative Folgen mit sich bringen kann, sind Fragen der Transparenz und Nachvollziehbarkeit („*Accountability*“) des Systems von großer Bedeutung. Hierbei muss zwischen verschiedenen Dimensionen unterschieden werden, welche im Folgenden explizit gemacht werden sollen.

zwei Formen der Transparenz

Transparenz als Forderung an die Gestaltung und den Einsatz algorithmischer Systeme kann zwei Formen annehmen: Einerseits die Dokumentation des Systems selbst – die Publikation systeminterner Prozesse, der verwendeten Daten und ihrer Herkunft, der relevanten Designentscheidungen im Entwicklungsprozess, Maßnahmen der Evaluation usw. – und andererseits die Nachvollziehbarkeit von Outputs des Systems, wie etwa im Fall des Systems AMAS die Kategorisierung im Einzelfall. Beide Aspekte haben im Zuge des Öffentlichwerdens des Systems einen gewissen Wandel durchlaufen, etwa durch die Veröffentlichung zusätzlicher Dokumente oder das Hinzufügen der Segmentzusatzinformationen als Erklärungsmodell für die Einzelentscheidungen des Systems.

mangelnde Transparenz bei Bekanntmachung von AMAS

Die Transparenz des Systems war zum Zeitpunkt der erstmaligen Veröffentlichung im Oktober 2018 dürftig, was zu einer Reihe an Fehlinterpretationen über die Abläufe und Modelle des Systems, die direkt auf die veröffentlichte Dokumentation zurückzuführen sind, führte. So lag der Fokus in [DOK_1 2018] etwa noch ausschließlich auf der Beschreibung des logistischen Regressionsmodelles und erklärte das Zustandekommen der Integrationschance im Einzelfall nicht im Detail, was die Vermutung nahelegte, die Integrationschance wäre direkt auf das Regressionsmodell zurückzuführen. Ebenso waren zu den Trefferquoten nur beispielhafte Auszüge verfügbar, die keine zuverlässigen Schlüsse auf die Performance des Systems zuließen.

Durch eine Reihe an langwierigen Anfragen unterschiedlicher Akteur*innen (z. B. Watchdog-Organisationen wie Epicenter.Works oder der Gleichbehandlungsanwaltschaft) nach dem Informationsfreiheitsgesetz konnten, Stück für Stück, diese und andere essentielle Details des Systems langsam aufgearbeitet werden. Erst nach dem Projektstart zur vorliegenden Studie war es letztendlich möglich, eine umfangreichere Anzahl an Dokumenten zu erhalten, welche teilweise (z. B. [DOK_2 2020]) erst mit Jänner 2020, also mehr als ein Jahr nach dem ersten Ausrollen des Systems im Testbetrieb, verfasst wurden, wie die Zeitstempel zeigen. Selbst zum aktuellen Zeitpunkt jedoch sind die Dokumente, auf denen diese Studie fußt, nicht öffentlich verfügbar und damit nur einem kleinen Kreis an Forscher*innen zugänglich, was die Diversität der möglichen Analysen massiv einschränkt.

Dieses zögerliche und wenig proaktive Veröffentlichen ist problematisch, da eine Funktion dieser Systemtransparenz ja unter anderem auch das Einbringen von Kritik und möglichen Anpassungsvorschlägen zum System wäre. Ist das System jedoch bereits seit über einem Jahr konzipiert und (wenn auch nur im Testbetrieb) ausgerollt, sinkt die Wahrscheinlichkeit, dass fundamentale Änderungen auch nur in Erwägung gezogen werden können. Wie zahlreiche internationale Beispiele zu Empfehlungen und Direktiven in Bezug auf algorithmische Audits – also Untersuchungen von algorithmischen Systemen und deren Folgen – zeigen, ist volle Transparenz bereits in frühen Stadien der Entwicklung und jedenfalls vor dem produktiven Einsatz nötig, um fundierte und wissenschaftliche Kritik in die Entwicklungen miteinbeziehen zu können (siehe etwa Treasury Board of Canada 2019; Crawford et al. 2019).

Dem Aspekt der Transparenz des gesamten Systems und seiner Abläufe gegenüber steht die Erklärbarkeit von Einzelentscheidungen. Während Systemtransparenz einem Anspruch an Vollständigkeit gerecht werden muss, um Expert*innen die Möglichkeit zu Feedback und Kritik zu geben, steht für die Erklärbarkeit die Verständlichkeit und der Nutzen für vom System Betroffene im Mittelpunkt. Im Fall des Systems AMAS bezieht sich dies etwa auf die Klassifikation von Arbeitssuchenden zu einer der drei Gruppen, oder die Erklärung zur Aussagekraft und Bedeutung der „Integrationschance“. Wie in Kapitel 4 ausführlicher beschrieben, sollen „Segmentzusatzinformationen“ Betroffenen und Sachbearbeiter*innen gleichermaßen Aufschluss über das Zustandekommen der Klassifikation durch das Ausweisen des Einflusses einzelner Personenattribute geben. Diese Information ist jedoch, nach heutigem Stand, nur unzureichend umgesetzt, wird sie doch ausschließlich für die Gruppe mit vollständig verfügbarer Datenbasis angezeigt: Personen mit unvollständiger Erwerbshistorie (Migrationshintergrund, Jugendliche, fragmentierter Erwerbsverlauf) bekommen keine Erklärung für ihre Klassifikation.

Darüber hinaus ist der positive Nutzen der möglichen Text-Bausteine für das Verständnis der Betroffenen fraglich. Alle der möglichen sieben Textbausteine für die „Niedrig“-Gruppe weisen auf Defizite hin, ermutigende Textbausteine werden nur der Gruppe mit „hohen Chancen“ angezeigt. Von diesen sieben Texten beziehen sich außerdem drei auf schwer bis

Aufarbeitung durch
Anfragen
verschiedener
Akteur*innen

keine Erklärungen von
AMAS-Entscheidungen
für Personen mit
fragmentiertem
Erwerbsverlauf

Nutzen der
Textbausteine fraglich

nicht beeinflussbare Faktoren (Betreuungspflichten, gesundheitliche Einschränkungen, fortgeschrittenes Alter); Personen, die diese Texte als Erklärung für die Zuweisung erhalten, bekommen somit in schriftlicher Form scheinbar objektive Argumente für die Aussichtslosigkeit ihrer Situation vorgelegt, haben aber keine weiteren Handlungsmöglichkeiten, um die Bewertung durch das System zu beeinflussen.

Unterschied zwischen
Korrelation und
Kausalität oft nicht
bewusst

Schlussendlich sind die vorliegenden Formulierungen irreführend, suggerieren sie doch einen Kausalzusammenhang zwischen Personeneigenschaften und den tatsächlichen Chancen am Arbeitsmarkt. Das System AMAS, wie derzeit beschrieben, operiert jedoch ausschließlich auf Korrelationen zwischen Personenattributen und Beobachtungen. Der Einleitungssatz „Auf Basis der uns zur Verfügung stehenden Daten waren folgende Faktoren für die Berechnung ihrer Arbeitsmarktchancen ausschlaggebend“ [SPEZ_1 2019], der den Textbausteinen vorangestellt wird, spricht zum Beispiel von „Berechnung“ statt etwa „Schätzung“, „Prognose“ oder ähnlichen Begrifflichkeiten, die der statistischen Natur des Systems besser Rechnung tragen würden. Ob dieser feine, aber signifikante Unterschied den Betroffenen bewusst wird, ist fraglich – fehlen doch im Kontext der Erklärungen jegliche Hinweise auf Fehlerraten des System, oder etwa die Größe der Vergleichspopulation. Letzteres wäre gerade für Personen, deren Merkmalskombination nicht mit statistisch ausreichenden Beobachtungen anderer Personen übereinstimmt und die daher mit anderen, merkmalsunterschiedlichen Konstellationen zusammen bewertet werden (weniger als 10 andere Personen derselben Merkmalskombinationen werden aggregiert geschätzt) eine wichtige Information für die Aussagekraft der berechneten Integrationschance.

Mangel an Transparenz
für Betroffene und
Expert*innen

Entgegen den wiederholten Beteuerungen seitens des AMS, höchstmögliche Transparenz beim System AMAS anzustreben, zeigen die beiden beschriebenen Aspekte schwere Transparenz-Mängel, die es sowohl den Betroffenen als auch Expert*innen schwer machen, das System bzw. seine Ergebnisse einzuschätzen. In Hinblick auf die Möglichkeiten, das System zur Rechenschaft („*Accountability*“) zu ziehen, bieten sich dadurch auch nur eingeschränkte Handlungsoptionen. Stimmen Betroffene den ihnen zugewiesenen Klassifikationen nicht zu, müssen sie im Dialog mit ihren Sachbearbeiter*innen eine Lösung suchen. Dabei stehen ihnen jedoch nur höchst rudimentäre Informationen in Form der Textbausteine zur Verfügung, durch deren scheinbare Objektivität nur wenig Gegenargumente möglich sind. Der Hinweis auf die statistische Schwankungsbreite oder eine etwaige mangelhafte Datenlage ist den Betroffenen hier durch die fehlende Information unmöglich. Damit wird wohl für die meisten negativ Betroffenen keine Alternative übrigbleiben, als sich in der ohnehin schon schwierigen Situation der Jobsuche mit einer psychisch belastenden Einschätzung des Systems und den daraus folgenden Konsequenzen abzufinden.

4 AMAS IN DER BERATUNGSPRAXIS

Das folgende Kapitel widmet sich der Einführung von AMAS in die Arbeitspraxis des AMS und stellt dar, wie das System in den Beratungsprozess eingebettet ist. Es zeigt, wie der AMS-Algorithmus eingeführt und den Berater*innen näher gebracht werden soll und wie er im EDV-System des AMS implementiert wurde. Daraus ergibt sich die Frage, wie AMAS die bestehende Praxis der Berater*innen verändert und welche konkreten Informationen und Vorgaben sie bekommen. Auf diese deskriptive Darstellung der neuen Aspekte in der Beratungspraxis folgt eine Analyse der Spannungen und Ambivalenzen, die das System für AMS-Mitarbeiter*innen, für Arbeitssuchende aber auch für die Organisation AMS aufwirft.

4.1 AMAS IM KERNPROZESS „ARBEITSKRÄFTE UNTERSTÜTZEN“

Das zentrale Dokument, das die Einführung und Einpassung von AMAS in die AMS Praxis definiert, ist die Bundesrichtlinie Kernprozess „Arbeitskräfte unterstützen“, kurz auch KP 1 Richtlinie genannt [RICHT_1 2019]. Vor COVID-19 war vorgesehen, dass die Richtlinie mit 1.1.2020 in Kraft tritt, die Übergangsfrist mit 1.7.2020 endet und die Richtlinie ab dann verpflichtend umzusetzen ist. Die Vollinbetriebnahme des Systems in allen Geschäftsstellen des AMS, wurde aufgrund des erhöhten Arbeitspensums im Zusammenhang mit den Arbeitsmarktmaßnahmen aufgrund der Corona-Krise auf Jänner 2021 verschoben und mittlerweile durch einen Bescheid der Datenschutzbehörde gestoppt [DSB 2020] (siehe Kapitel 2). Das AMS legte Bescheidbeschwerde ein und der Ausgang des Verfahrens ist bis dato offen.

Die KP 1 Richtlinie stellt den Prozess dar, der dem AMS als Grundlage dafür dient, „Arbeitskräfte bei der Suche nach Beschäftigungsmöglichkeiten und bei der Anpassung an die Arbeitskräftenachfrage [zu] unterstützen“ [RICHT_1 2019, S. 5]. Der Prozess ist in vier Phasen untergliedert, die in Abbildung 10 dargestellt sind, und integriert für die Einführung von AMAS eine Reihe an neuen Regelungen.

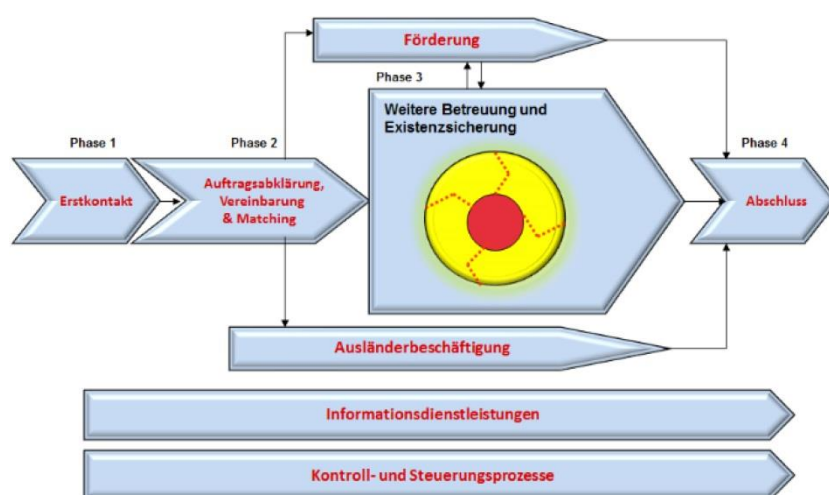


Abbildung 10: Vier Phasen des Kernprozesses „Arbeitskräfte unterstützen“,
Quelle: [RICHT_1 2019, S. 5]

vier Phasen der Betreuung

Die Betreuung durch das AMS umfasst Beratung, Vermittlung, finanzielle Unterstützung im Rahmen der Arbeitslosenversicherung, Förderung der Beschäftigung, der Qualifizierung, Maßnahmen zur Unterstützung dieser Förderung sowie Unterstützung im Rahmen der sogenannten Ausländerbeschäftigung. Schon beim Erstkontakt (Phase 1) werden notwendige Daten erhoben, um eine Grundlage für die Prognose der Arbeitsmarktchance der Kund*innen zu schaffen. AMAS spielt dann insbesondere in der Phase 2 im Rahmen der Abklärung und Vereinbarung des Betreuungsbedarfs eine wichtige Rolle und wirkt sich durch eine Einteilung der Arbeitssuchenden in „Servicekund*innen“ mit „hohen Arbeitsmarktchancen“ (H), „Beratungskund*innen“ mit „mittleren Arbeitsmarktchancen“ (M) und „Betreuungskund*innen“ mit „niedrigen Arbeitsmarktchancen“ (N) auf die konkrete Ausgestaltung der Phase 3 im Sinne unterschiedlicher Betreuungsintensitäten aus. AMAS ist ein wesentlicher Teil des Bereichs Förderung, das heißt jener Aktivitäten, die der Integration von Arbeitssuchenden in den Arbeitsmarkt dienen oder der Sicherung einer Beschäftigung in Form von Maßnahmen und Beihilfen zu Maßnahmen, die vom AMS finanziert werden. Die unterschiedlichen Phasen des Kernprozesses „Arbeitskräfte unterstützen“ werden in der KP 1 Richtlinie graphisch in Form eines Modells abgebildet, das die konkreten Schritte in Phase 2 etwa durch die Darstellung in Abbildung 11 erklärt.

Prozessschritte

Die Elemente des Modells bezeichnen Funktionen, Aktivitäten, Prozessschritte, Vorgänge wie etwa das Element „Gesetzliche Voraussetzungen für Vormerkung prüfen“; Ereignisse, die auslösende Elemente für Prozesse sind, wie etwa „Gesetzliche Voraussetzungen für Vormerkung erfüllt“; und Schnittstellen zu vor- und nachgelagerten Prozessen, wie etwa „Ausländerbeschäftigung“ oder „Förderung“. Die Arbeitsmarktchancen werden erstmals in Phase 2 bewertet und führen hier beispielsweise zur „Entscheidung ob sofort vermittelt werden kann“ und/oder zu einem Leistungsantrag für „weitere Betreuung und Existenzsicherung“. Im Laufe der Pha-

sen 3 und 4 des Prozesses werden die Arbeitsmarkchancen immer wieder bewertet und führen dann zum Abschluss einer Betreuungsvereinbarung.

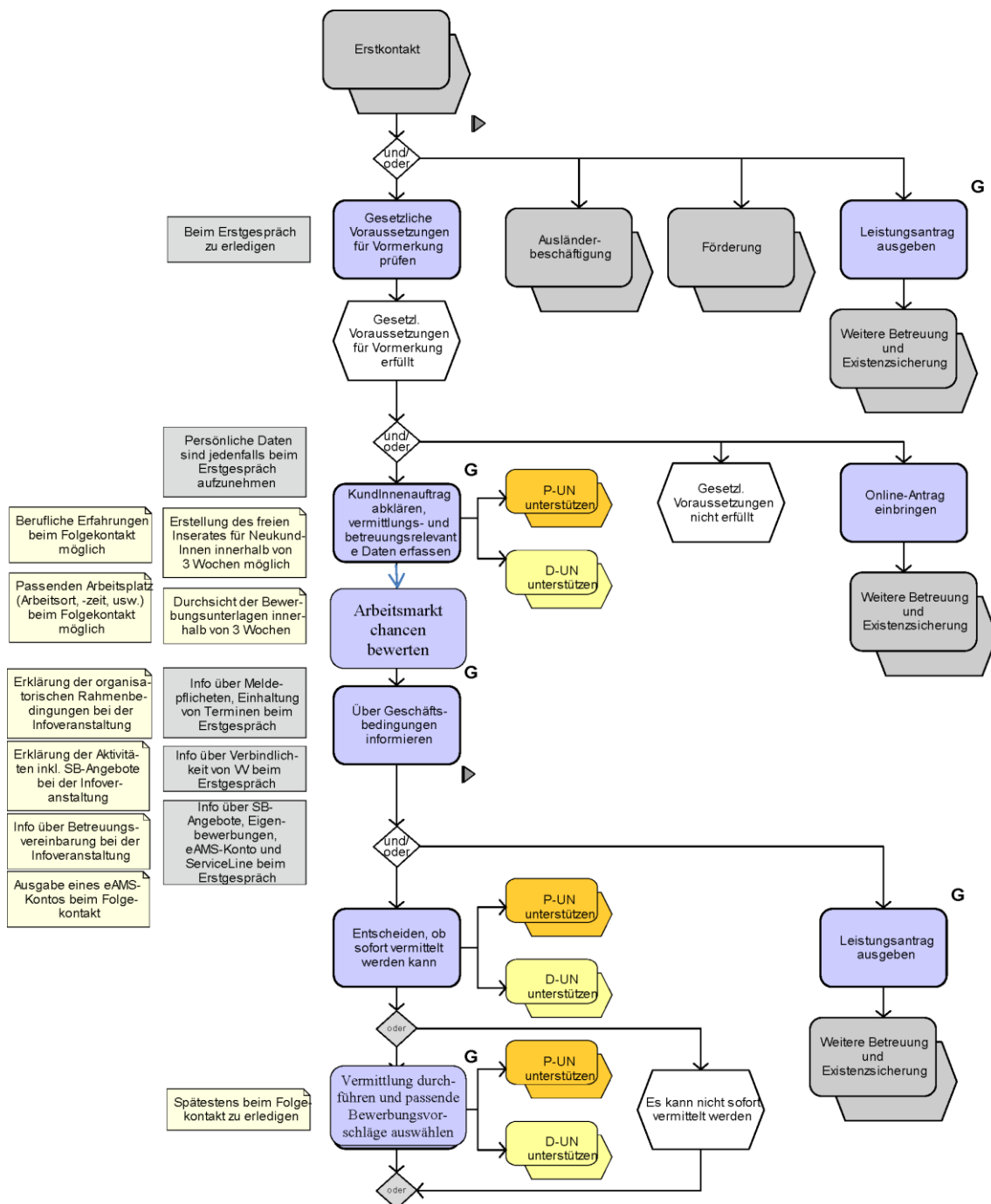


Abbildung 11: Ausschnitt aus Detail „Phase 2 Auftragsabklärung, Vereinbarung & Matching“, Quelle: [RICHT_1 2019, S. 152]

4.1.1 EINFÜHRUNG VON AMAS IN DIE AMS PRAXIS

Auf Basis der KP 1 Richtlinie und anderer Dokumente wie etwa dem „AMAS Handbuch“ [HAND_1 2019] wurden eine Reihe von Initiativen rund um die Einführung des Systems gestartet und Schulungsunterlagen ausgearbeitet. Zunächst wurde eine „Informationsplattform“ mit dem Tool „Connections“ im Intranet des AMS eingerichtet. Dort finden sich alle relevanten Dokumente zu AMAS wie die KP 1 Richtlinie oder das Handbuch. Darüber hinaus wurde eine E-learning Plattform aufgesetzt, in der Informations- und Schulungsunterlagen zu finden sind, und eine interne Kommunikationsplattform über eine eigene AMAS Seite zur Verfügung gestellt, über die alle Mitarbeiter*innen erreicht werden können [BRIEF_3 2020].

Schulungen,
Multiplikator*innen
und e-Learning

Neben diesen schriftlichen Unterlagen sind Informationsveranstaltungen und Workshops geplant, um AMS-Mitarbeiter*innen persönlich in das System einzuschulen [BRIEF_3 2020]. Bereits im Herbst 2018 wurden in allen Landesstellen „AMAS Ansprechpersonen“ definiert, die als „Multiplikator*innen“ dienen sollten. Diese werden über Inhalte, Funktionsweise und Entwicklungen zu AMAS laufend informiert. Diese „AMAS Ansprechpersonen“, sowie weitere AMS-Mitarbeiter*innen in Führungspositionen, wurden bei einer ganztägigen Informationsveranstaltung vertiefend über „das Assistenz-System“ informiert. Diese Veranstaltungen hatten bis März 2020 (Zeitpunkt der Beantwortung unserer Datenanfrage) in sieben Bundesländern stattgefunden. Ausgehend von den zentralen, bundesweiten Richtlinien und Unterlagen wurden in den Landesorganisationen weitere Schulungspläne ausgearbeitet. Die Schulungen in den Landesgeschäftsstellen und in regionalen Geschäftsstellen werden durch interne Trainer*innen bewerkstelligt, wodurch ein Praxisbezug hergestellt werden soll. In diesem Prozess wurden die bundesweiten Richtlinien und Informationen für Praktiker*innen „übersetzt“ und allfällige „landesspezifische Schwerpunktsetzungen“ vorgenommen (Email Dr. Ernst Haider/AMS, 18.3.2020). Für die Zukunft sind zusätzliche „Train the Trainer Werkstattgespräche“ geplant, um externe Trainer*innen, mit denen das AMS bereits zusammenarbeitet, zu schulen und das Wissen zu AMAS in laufende Weiterbildungsangebote zu integrieren. Laut Auskunft des AMS gibt es in den Landesgeschäftsstellen rund 160 „AMAS Ansprechpersonen“. Insgesamt sollen alle rund 4000 AMS-Mitarbeiter*innen in den „Beratungs-, Info- und Servicezonen“ persönlich eingeschult werden und schriftliche Unterlagen und e-Learning Tools etc. zur Verfügung gestellt bekommen. Mitarbeiter*innen aus anderen Bereichen, wie z. B. aus dem Bereich „Service für Unternehmen“, werden im Rahmen einer Informationsveranstaltung oder bei Dienstbesprechungen über AMAS informiert, da sie das System kennen, aber nicht anwenden müssen. AMS-Mitarbeiter*innen aus den „BerufsInfo-Zentren“ werden ebenfalls entsprechend informiert (Email Dr. Ernst Haider/AMS, 18.3.2020). Zusammenfassend bedeutet das, dass der Schulungsprozess von der Bundesgeschäftsstelle hin zu den regionalen Geschäftsstellen in einem Top-Down Verfahren bewerkstelligt wird und dass versucht wird, etwas differenziert auf die jeweiligen Kontexte der unterschiedlichen Bereiche und Regionen, in denen die Mitarbeiter*innen tätig sind,

einzugehen. Dieser Prozess wurde aufgrund der Corona-Krise Mitte März bis auf weiteres ausgesetzt, weshalb noch nicht alle Mitarbeiter*innen eingeschult wurden, obwohl AMAS bereits seit Herbst 2018 im Testbetrieb lief.

4.1.2 IMPLEMENTIERUNG DER GESCHÄFTSFUNKTION „ARBEITSMARKTCHANCE“

Die AMS-Mitarbeiter*innen arbeiten mit einem EDV-System in das alle Daten über Arbeitssuchende eingetragen werden. Das System wird intern kurz als PST (Personenstammdaten) bezeichnet. Wie in Abbildung 12 ersichtlich, gibt es im PST verschiedene Register wie etwa PER, PLV, BW1 usw. (Leiste links oben) in die Personendaten wie das Beginn-Datum des Geschäftsfalls, der Status der Deutschkenntnisse der Kund*innen, Beruf oder Ausbildung, möglicher Berufsantritt, Schlüsselwörter für Kenntnisse und Fähigkeiten, gewünschtes Arbeitsausmaß usw. eingetragen werden.

das EDV-System
des AMS

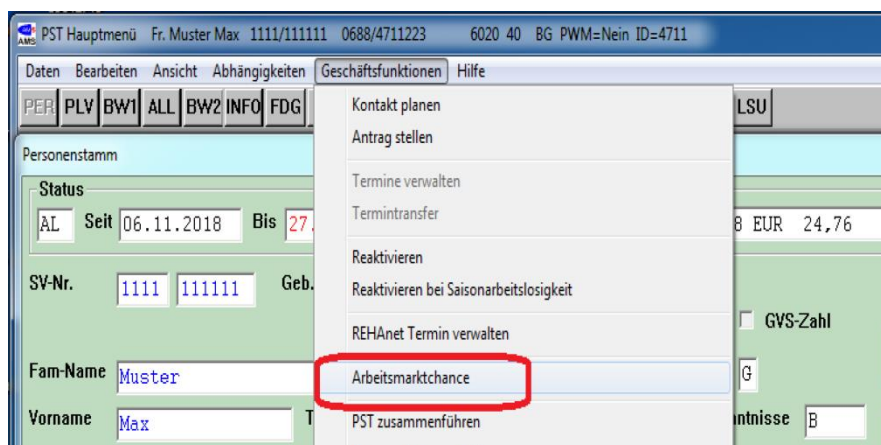


Abbildung 12: Grafische Benutzungsoberfläche des EDV-Systems „Personenstammdaten“,
Quelle: [SCHU_7 2020, S. 37]

Der Screenshot zeigt auch die Einbettung der neuen Geschäftsfunktion „Arbeitsmarktchance“. Falsche Eintragungen in den unterschiedlichen Segmenten des Systems können zu einer falschen Prognose der Arbeitsmarktchancen führen [SCHU_7 2020]. Das heißt, die Treffsicherheit der Prognose hängt auch mit der Eintragungs- und Datenqualität im Gesamtsystem zusammen.

Abbildung 13 zeigt die Benutzungsoberfläche der Geschäftsfunktion „Arbeitsmarktchance“ im Detail. Rechts oben sehen AMS-Mitarbeiter*innen die Prognose der kurzfristigen und langfristige Arbeitsmarktchancen der jeweiligen Kund*innen in Prozent. Angezeigt werden immer die aktuellsten Berechnungen der Arbeitsmarktchance anhand eines Wertes zwischen 0 und 100 Prozent und diese Prozentwerte sind nicht veränderbar. Links oben findet sich die computergestützte Arbeitsmarktchance (CAM). Der CAM wird im Data-Warehouse des AMS bestimmt und in den PST eingespielt.

Anzeige der
computergestützten
Arbeitsmarktchance
(CAM)

Er kann durch die Berater*innen nicht gelöscht werden [SCHU_7 2020, S. 39]. Der letzte zum genannten Datum berechnete CAM ist in diesem Fall CAMN also CAM „niedrig“. In einem AMS-internen Informationsworkshop wird auch betont, dass der CAM eine Prognose (Wahrscheinlichkeit) darstellt und keine Aussage über das Individuum trifft [SCHU_7 2020].

Abbildung 13: Geschäftsfunktion „Arbeitsmarktchance“ im PST,
Quelle: [SCHU_7 2020, S. 38]

Eintragen der Berater*innen- Arbeitsmarktchance (BAM)

Unter dem CAM befindet sich das Feld Berater*innen-Arbeitsmarktchance (BAM), in das die AMS-Mitarbeiter*innen ihre aus dem Beratungsgespräch hervorgehende eigene Einschätzung der Arbeitsmarktchancen eintragen können. Dafür muss auch ein Eintrag im Pflichtfeld Begründung erfolgen, während bei der Einspielung des CAM keine Begründung angezeigt wird [SCHU_7 2020, S. 43]. Der BAM kann auch gesetzt werden, wenn vom System noch kein CAM eingespielt wurde [SCHU_7 2020, S. 40].²⁸ In diesem Beispiel (Abbildung 13) ordnet AMAS die arbeitssuchende Person im Segment „niedrige Arbeitsmarktchance“ ein, während der*die Betreuer*in sie im Segment „mittlere Arbeitsmarktchance“ einordnet. Damit verbunden ist eine Einstufung der Arbeitssuchenden als „Beratungskundin“ oder als „Betreuungskundin“, wovon sich die weitere Betreuungsstrategie, konkret die Förderung und internen Dienstleistungen sowie das Kontaktintervall ableitet (siehe Abbildung 18, S. 73).

²⁸ Ohne Einspielung eines CAM ist eine Eingabe eines BAM nur möglich, wenn der PST aktuell ist oder sich die Person in einer Maßnahme befindet [SCHU_7 2020, S. 40].

Gemäß der Auslegung seitens des AMS, dass AMAS keine Entscheidungen trifft, sondern die Entscheidungsfindung nur unterstützt, treffen die AMS-Mitarbeiter*innen über die Felder „Ändern“ oder „CAM übernehmen“ schließlich die jeweils aktuelle Letztentscheidung über die Einschätzung der Arbeitsmarktchancen der arbeitssuchenden Person. Ein Klick auf „Ändern“ übernimmt den BAM, ein Klick auf die Funktion „CAM übernehmen“ löscht den BAM Deskriptor. Ein CAM kann durch einen BAM überschrieben aber nicht gelöscht werden. Wird der CAM übernommen, können sich AMS-Mitarbeiter*innen die sogenannten Segmentzusatzinformation, also eine systemgenerierte Begründung für die Prognose der Arbeitsmarktchance anzeigen lassen; zumindest für jene Personen mit einem vollständigen Datensatz (siehe Abbildung 16 weiter unten).

CAM ändern
oder übernehmen

Eintragung eines BAM, weil die Arbeitsmarktchance durch die Beraterin oder den Berater besser oder schlechter bewertet werden	
Kriterium wird im Arbeitsmarktchancen Assistenz-System nicht berücksichtigt	<ul style="list-style-type: none"> • Motivation und Selbsthilfepotenzial der Kundin/des Kunden • Wohnungssituation • Sucht • Schulden • fehlender Leumund, Haft • vorhandene/fehlende Qualifikation einer Person (nicht die höchste abgeschlossene Ausbildung) • aufgrund der Kompetenzen der Person sind alternative Berufe möglich • Sonstige Kriterien
Kriterien des Arbeitsmarktchancen Assistenz-System haben sich im Laufe der Vormerkung verändert	<ul style="list-style-type: none"> • Betreuungspflichten wurden geregelt/sind entstanden • regionale Mobilität hat sich verbessert/verschlechtert • gesundheitliche Einschränkungen • der bisherige Beruf kann nicht mehr ausgeübt werden, weil es sich um einen „veralteten“ Beruf handelt oder gesundheitliche Einschränkungen dies nicht mehr zulassen • Beschäftigungszeiten im Ausland • Arbeitsmarktlage wird aktuell besser/schlechter eingeschätzt
Arbeitsmarktchance verbessert sich durch den Einsatz einer Förderung	<ul style="list-style-type: none"> • Für Personen mit CAMN ist von der Beraterin oder vom Berater zu überlegen, ob durch den Einsatz einer Förderung die Chance der Arbeitsmarktintegration steigt. Mit der Kundin oder dem Kunden ist abzustimmen, ob das Angebot für sie/ihn bewältigbar ist. Das Ergebnis von Einstiegstests bei Trägern kann die Entscheidung der Beraterin oder des Beraters unterstützen. • Nach dem Einsatz einer Förderung (z.B. BBEN) hat sich die Arbeitsmarktchance verbessert. Der Ergebnisbericht eines Trägers kann die Entscheidung der Beraterin oder des Beraters unterstützen. • Zusage eines Unternehmens, dass Person nach einer Qualifizierung beschäftigt wird.

Abbildung 14: Kriterien für die Eintragung eines BAM, Quelle: [RICHT_1 2019, 82]

Für die Eintragung eines BAM gibt es eine Reihe an Kriterien (siehe Abbildung 14). Grundsätzlich wird ein BAM eingetragen, wenn ein*e Berater*in die Arbeitsmarktchance der arbeitssuchenden Person besser oder

Kriterien für die
Eintragung eines BAM

schlechter bewertet als AMAS [SCHU_7 2020, S. 41]. Die AMS-Berater*innen sind hier aufgefordert als „soziales Korrektiv“ zu fungieren und falsche oder unzureichende Einstufungen mittels ihrer eigenen Erfahrungen und Einschätzungen zu berichtigen.

Wird die Arbeitsmarktchance einer Person durch den*die AMS-Mitarbeiter*in von „hoch“ oder „mittel“ auf „niedrig“ heruntergestuft, muss ein Perspektivencheck angeboten werden und im entsprechenden Feld im unteren Bereich der Eingabemaske wird ein Protokoll davon im System angelegt. Die Daten dafür werden aus dem Teilnahmen-Administrationssystem (TAS) übernommen [SCHU_7 2020, S. 43].

BAM wird gelöscht

Ein weiteres Beispiel, in dem eine bessere Segmentzuordnung durch den BAM vorgenommen wurde als jene, die vom CAM eingespielt wurde, ist in Abbildung 15 zu sehen. Obwohl die Begründung für den BAM positive Arbeitsmarktchancen in Aussicht stellt, dient dieser Fall im Informationsworkshop als Beispiel für eine Übernahme des schlechteren CAM und eine entsprechende Beendigung und Löschung des BAM.

The screenshot shows the 'Arbeitsmarktchance' application window. It is divided into three main sections: 'Computergestützte Arbeitsmarktchance', 'Beraterinnen-Arbeitsmarktchance', and 'Protokoll'.

- Computergestützte Arbeitsmarktchance:** Contains fields for 'CAM' (set to 'CAMN'), 'erstellt' (13.12.2019), and 'S001'. To the right, 'Arbeitsmarktchance in %' shows 'kurzfristige' at 6 and 'langfristige' at 19.
- Beraterinnen-Arbeitsmarktchance:** Contains fields for 'BAM' (set to 'BAMM'), 'erstellt' (16.12.2019), 'bis' (empty), and 'geändert' (empty). Below these is a 'Begründung' field with the text: 'Kundin hat Betreuungspflichten geregelt und steht der Vermittlung uneingeschränkt zur Verfügung'.
- Protokoll:** A table with columns 'Code', 'ab', 'bis', 'Ben.', and 'Begründung'. The table is currently empty.
- Buttons:** At the bottom, there are buttons for 'Ändern', 'CAM übernehmen' (highlighted with a red rectangle), 'Segmentzusatzinformation anzeigen', 'Schließen', and 'Hilfe'.

Abbildung 15: Aufstufung von Computer-Arbeitsmarktchance „niedrig“ (CAMN) auf Berater*innen-Arbeitsmarktchance „mittel“ (BAMM),
Quelle: [SCHU_7 2020, S. 42]

Erklärung der
Einstufung

Über die Schaltfläche „Segmentzusatzinformation anzeigen“ können Details zur Einstufung der computerunterstützten Arbeitsmarktchance angezeigt werden (siehe Abbildung 16). Die aus AMAS generierten Erklärungstexte sollen den Kund*innen im Beratungsgespräch mitgeteilt werden und dienen dazu, die Einschätzung von Berater*in und Kunde*in abzugleichen.

Segmentzusatzinformation

Erstellung: S001 / 03.11.2019 Bearbeitung: S001 / 03.11.2019

Dokument

Betreff: Arbeitsmarktchance von mittel auf niedrig

Auf Basis der uns zur Verfügung stehenden Daten waren folgende Faktoren für die Berechnung Ihrer Arbeitsmarktchance ausschlaggebend

Neue Arbeitsmarktchance: CAMM
 Segmentzusatzinformation:
 Sie haben keine über das Pflichtschulniveau hinausgehende Berufsausbildung absolviert.
 Ihre erworbene Qualifikation ist auf dem Arbeitsmarkt nur wenig verwertbar.

kurzfristige Arbeitsmarktchance: 10 %
 langfristige Arbeitsmarktchance: 25 %

Bisherige Arbeitsmarktchance: CAMM
 Sie waren über einen sehr langen Zeitraum in den letzten Jahren beschäftigt.
 Sie haben keine über das Pflichtschulniveau hinausgehende Berufsausbildung absolviert.

kurzfristige Arbeitsmarktchance: 10 %
 langfristige Arbeitsmarktchance: 35 %

Drucken Schließen Hilfe

Abbildung 16: Segmentzusatzinformation zum CAM, Quelle: [SCHU_7 2020, S. 45]

Die Segmentzusatzinformation wird nur angezeigt, wenn es einen CAM gibt und dieser auf voll validen Datensätzen beruht. Aus diesem Grund erscheint bei rund 30 % der Kund*innen keine Information [SCHU_7 2020, S. 44]. Die Segmentzusatzinformation richtet sich nach dem CAM, auch wenn ein BAM eingetragen wurde, und stellt eine „Auflistung der wichtigsten positiven und hemmenden Einflussfaktoren mittels vorgefertigter Textbausteine“ sowie Informationen zu AMS Regeln [SCHU_7 2020, S. 44]. Im Informationsworkshop wird aber auch angeführt, dass die Segmentzusatzinformation „keine Erklärung für die Einstufung“ liefert [SCHU_7 2020, S. 44]. Welche Bedeutung diese Unterscheidung zwischen Einflussfaktoren und Erklärung für die Einstufung aus Sicht der Betreuer*innen und aus Sicht der Kund*innen genau haben soll, bleibt unklar. In der Praxis scheint eine Gleichsetzung naheliegend. Die Segmentzusatzinformation wird für Neuzugänge täglich und ansonsten monatlich auf Basis des CAM generiert und ist ein Ergebnis der Regressionsanalyse (siehe Kapitel 3). Im konkreten Beispiel in Abbildung 16 wird eine Arbeitssuchende vom Segment „mittlere Arbeitsmarktchance“ ins Segment „niedrige Arbeitsmarktchance“ aufgrund einer Verringerung der langfristigen Arbeitsmarktchance von 35 % auf 25 % heruntergestuft. Während in der früheren Prognose eine lange Beschäftigungsdauer innerhalb der letzten Jahre als ausschlaggebend ausgewiesen wird, wird die erworbene Qualifikation in der neuen Prognose vermutlich durch die Dauer des Geschäftsfalls als wenig verwertbar bezeichnet. Aus der Dauer der Arbeitslosigkeit werden hier offenbar Rückschlüsse auf die Verwertbarkeit der Qualifikation gemacht. Die Segmentzusatzinformation kann ausgedruckt werden und wird dabei seit Dezember 2019 automatisch mit der PST-Datenschutzauskunft ergänzt. Abbildung 17 zeigt eine Auflistung von Faktoren, die für die Berechnung von Arbeitsmarktchancen als ausschlaggebend ausgewiesen werden:

Erklärungstexte
 nur für
 Arbeitssuchende mit
 vollständigen Daten

Kurzbezeichnung (kommt im PST nicht vor)	Text im Segment „Arbeitsmarktchancen“
„Ermutigend“	
1. Hohe Beschäftigungskontinuität	Sie waren über einen sehr langen Zeitraum in den letzten Jahren beschäftigt.
2. Erfahrung mit arbeitsmarktpolitischen Maßnahmen	Sie haben Bereitschaft gezeigt, an Förderangeboten des AMS teilzunehmen.
3. Bewerbungserfahren	Sie haben wiederholt erfolgreich zu arbeiten begonnen und verfügen somit über Beschäftigungserfahrung auf dem Arbeitsmarkt.
4. Etablierte Berufsausbildung	Sie verfügen über eine Berufsausbildung mit formalem Abschluss zumindest auf Facharbeiterniveau.
„Besonders herausfordernd“	
5. Kompetenz-Schwächen	Sie weisen nur Pflichtschulbildung und/oder geringe Deutschkenntnisse auf, was die Arbeitsuche erschwert.
6. Keine Berufsausbildung	Sie haben keine über das Pflichtschulniveau hinausgehende Berufsausbildung absolviert.
7. Erhebliche Beschäftigungslücken	Sie waren nur über kurze Zeiträume in den letzten Jahren beschäftigt.
8. Geringe Maßnahmeninanspruchnahme *)	Sie waren in den letzten Jahren in keine Förderangebote des Arbeitsmarktservice einbezogen.
9. Betreuungspflichten	Sie haben Betreuungspflichten in ihrem familiären Umfeld.
10. Gesundheitliche Einschränkungen	Sie weisen gesundheitliche Einschränkungen auf, die die Arbeitssuche erschweren.
11. Fortgeschrittenes Alter	Für Menschen wie Sie mit einem höheren Alter ist es schwieriger eine neue Beschäftigung zu finden.

Abbildung 17: Faktoren für Segmentzusatzinformationen, Quelle: [SCHU_7 2020, S. 46]

4.1.3 SYSTEMINFORMATIONEN UND VORGABEN FÜR AMS-MITARBEITER*INNEN

detaillierte Vorgaben

Die KP 1 Richtlinie, und die darauf aufbauenden Dokumente und Schulungsunterlagen, definieren sehr genau wie AMAS in der Praxis durch AMS-Mitarbeiter*innen gehandhabt werden soll. Folgende Beschreibung kann als Leitsatz für den Einsatz des Systems begriffen werden:

„Das Arbeitsmarktchancen Assistenz-System berechnet, wie es anderen Personen in der Vergangenheit in einer ähnlichen Situation ergangen ist.“ AMAS ist „ein wissenschaftlich-empirisch fundiertes computerunterstütztes Rechenmodell, das persönliche Eigenschaften der Person (Alter, Geschlecht, gesundheitliche Beeinträchtigung etc.) in Beziehung zum Arbeitsmarkt setzt und auf dieser Basis die zukünftigen Integrationschancen prognostiziert. Dieses „neue Personenmerkmal“ (Deskriptor im PST und Geschäftsfunktion Arbeitsmarktchance) ist somit eine Unterstützung bzw. „2. Meinung“ im Betreuungsprozess. Die Letztentscheidung über die Zuordnung zu einer KundInnengruppe treffen aber immer die BeraterInnen, deren Erfahrung und persönliche Beurteilung dazu wichtig sind.“ [RICHT_1 2019, S. 13]

Letztentscheidung soll mit Arbeitssuchenden getroffen werden

Hier finden sich die relevanten Elemente, die in allen Unterlagen in der einen oder anderen Form wiederholt werden. Insbesondere die Beschreibungen von AMAS als „2. Meinung“ und die Formulierung, dass die „Letztentscheidung“ bei den Berater*innen und deren Erfahrungen liegt, findet sich in nahezu allen Dokumenten und Schulungsunterlagen wieder. Ein „Entscheidungsautomatismus“ durch AMAS wird entschieden abgelehnt. Weiters wird durchwegs betont, dass AMS-Mitarbeiter*innen die Arbeitsmarktchance mit den Arbeitssuchenden besprechen sollen, sie mit der Selbsteinschätzung der Arbeitssuchenden abgleichen sollen, allfällige Bedenken und Ängste thematisieren und gemeinsam mit den Arbeitssuchenden die „Letztentscheidung“ treffen sollen. Konkret wird das so formuliert:

„Die Ergebnisse des Arbeitsmarktchancen Assistenz-Systems sind eine **errechnete Grundlage zur Einschätzung der Chancen am Arbeitsmarkt**. Die Ergebnisse sollen ein Ausgangspunkt sein, um gemeinsam mit der Kundin oder dem Kunden eine Einschätzung der Potenziale und gegebenenfalls Hindernisse bei der Arbeitsmarktintegration zu treffen. Aufgrund dieser Auseinandersetzung ist entsprechend der Möglichkeiten für das jeweilige KundInnensegment die optimale Betreuungsstrategie (Förderungen und Betreuungsdienstleistungen) zu definieren. Mögliche Bedenken und Ängste der Kundinnen und Kunden müssen besprochen und von den Beraterinnen und Beratern ernst genommen werden. Hat die Kundin bzw. der Kunde dezidiert eine andere Einschätzung zu den Arbeitsmarktchancen als die Beraterin bzw. der Berater, ist dies in der Betreuungsvereinbarung zu dokumentieren.“

[RICHT_1 2019, S. 14].

Darüber hinaus wird festgehalten, dass auf Anfrage von Kund*innen erklärt werden soll, dass AMAS berechnet wie es „anderen Personen in der Vergangenheit in einer ähnlichen Situation ergangen ist“ und dass der errechnete Chancenwert eine Wahrscheinlichkeit aufzeigt. Das Kund*innen-Gespräch soll die „Akzeptanz“ von AMAS erhöhen, wie in der KP 1 Richtlinie zu lesen ist [RICHT_1 2019, S. 80].

Erhöhung der
Akzeptanz von AMAS

Die Richtlinie regelt die Betreuungsvereinbarung inklusive der Kontaktintervalle etc. sehr genau: „die Häufigkeit der Kontakte und die Intensität der Betreuung orientieren sich am KundInnensegment“ [RICHT_1 2019, S. 91]. Auch wie und was in der Betreuungsvereinbarung festgehalten werden soll, ist genau festgelegt, wie Abbildung 18 zeigt:

Einstufung legt
Kontakthäufigkeit fest

Hohe Arbeitsmarktchancen	Mittlere Arbeitsmarktchancen	Niedrige Arbeitsmarktchancen
Ein Folgekontakt spätestens innerhalb von 8 Wochen ab Vormerkung (keine Verpflichtung für weitere persönliche Kontakte)	Monatliches Kontaktintervall (Maximalintervall, d.h. Kontakt kann auch öfter als monatlich erfolgen)	Mindestens 1 persönlicher Kontakt im Jahr bei Personen in der BBEN, sofern die Dienstleistung „Bewerbungsunterstützung einschließlich Reflexion des Bewerbungsverhaltens und der Unterstützung bei der Rückmeldung an das AMS betreffend Vermittlungsvorschläge“ der BBEN mindestens alle 2 Monate in Anspruch genommen wird. Maximal 4 Monate bei <ul style="list-style-type: none"> - Personen in der BBEN, die die Mindestanspruchnahme von 2 Monaten bzgl. der Dienstleistung „Bewerbungsunterstützung“ der BBEN nicht einhalten - sowie bei Personen, die NICHT an der BBEN teilnehmen.

Abbildung 18: Kontaktintervalle nach Segmenten, Quelle: [SCHU_7 2020, S. 60]

Schließlich wird definiert, wie und wann Berater*innen die berechnete Chance ändern können (BAM) und aufgrund welcher Kriterien (siehe Abbildung 15, S. 70). Des Weiteren wird festgelegt, wann Berater*innen keinen BAM eintragen dürfen und welche Vorkehrungen notwendig sind, wenn sie Arbeitssuchende auf „niedrig“ umstufen wollen. Dann nämlich muss ihnen ein Perspektivencheck durch eine externe Einrichtung angeboten werden.

Standardisierung der Interaktion mit Kund*innen

Das standardisierte Vorgehen im Beratungsgespräch, das in der KP 1 Richtlinie definiert ist, wurde darüber hinaus in einen Fragenkatalog mit fiktiven Kund*innenanfragen übersetzt, um den AMS-Mitarbeiter*innen ganz konkrete Hilfestellung für die Interaktion mit Kund*innen an die Hand zu geben. Hier ist zu lesen, dass im Kund*innengespräch keine Abkürzungen wie AMAS verwendet werden sollen und welche Antworten auf welche Fragen zu geben sind. Neben sehr konkreten Antworten zum Modell, zur Einstufung und zu den verwendeten persönlichen Merkmalen etc. finden sich auch folgende Erklärungen, warum AMAS eingesetzt wird: Die Integration in den Arbeitsmarkt soll schneller geschaffen werden, mehr Effizienz beim Fördermitteleinsatz soll erreicht werden, es soll mehr Zeit für das „Kerngeschäft“ bleiben, nämlich Kund*innen mit „mittleren“ Chancen zu beraten und Personen mit „niedrigen“ Chancen neuen Betreuung- und Beratungseinrichtungen zu überlassen. In der Erklärung der Arbeitsmarktchance spielen auch die Textbausteine – „Segmentzusatzinformationen“ – eine wichtige Rolle, die Arbeitssuchenden erklären sollen, warum sie in Gruppe hoch/mittel/niedrig gelandet sind (siehe Abbildung 17). Diese beziehen sich auf positive und negative Einflussfaktoren auf die berechnete Chance: *„Ich kann Ihnen gerne die 2 wichtigsten Kriterien sagen, die für Ihre Einschätzung am stärksten gewirkt haben“* [RICHT_2 2019, S. 3]. Auf die Frage nach Diskriminierung soll geantwortet werden, dass es sich nicht um eine Diskriminierung durch das Assistenz-System handelt: *„Nicht das Modell diskriminiert, sondern teilweise der Arbeitsmarkt wie z. B. bei älteren Menschen oder Menschen mit Behinderung“* [RICHT_2 2019, S. 4].

Vorgehen bei offenen Fragen

Schließlich wird noch die Trefferquote erklärt, was passiert, wenn die Chance „niedrig“ ist, und dass die berechneten Arbeitsmarktchancen nicht an Unternehmen weitergegeben werden. Sollten dennoch Fragen offenbleiben, stünden in jeder Landes- und regionalen Geschäftsstelle „eingeschulte und kompetente“ Ansprechpersonen zur Verfügung. Letzte Instanz bei offenen Fragen, fragwürdigen Arbeitsmarktchancen oder strittigen Fällen ist die Bundesgeschäftsstelle (Email Dr. Ernst Haider, 18.3.2020). Aber auch davor können „Korrekturen mittels Umstufungen“ vorgenommen werden, wie oben beschrieben. Zusammenfassend wird das Prozedere des Kund*innengesprächs und die Rolle von AMS-Mitarbeiter*innen mit dem Slogan *„BeraterInnen haben den Klangstab in der Hand“* [SCHU_7 2020, S. 32] versinnbildlicht.

4.2 SPANNUNGEN UND AMBIVALENZEN IN DER AMS PRAXIS

Dieser Abschnitt identifiziert kritische Aspekte der Einführung von AMAS in die Beratungspraxis und reflektiert die Veränderungen, die damit – vor dem Hintergrund der laufenden Umgestaltung der Beratungspraxis zu post-bürokratischer Interaktionsarbeit unter Einsatz von Techniken des New Public Management – einhergehen. Eine zentrale Anforderung an AMAS ist die Steigerung von Effizienz und Effektivität im Beratungsprozess, sowie die Standardisierung der Beratungstätigkeit und Vermeidung von „Willkür“, wie eingangs ausführlich diskutiert (siehe Kapitel 1). Doch können diese Ziele erreicht werden, wenn wir uns die Funktion von AMAS im Beratungsgespräch im Detail vor Augen führen? Diese Frage werden wir nun auf drei Ebenen diskutieren: 1) auf der Ebene der Berater*innen, 2) auf der Ebene der Kund*innen und 3) auf institutioneller Ebene.

kritische Aspekte

Spannungsverhältnisse auf der Ebene der Berater*innen

Aus Sicht der AMS-Berater*innen ist ein Spannungsverhältnis zwischen AMAS als nötige Entlastung und als Einschränkung ihres Handlungsspielraums zu beobachten. In ihrer Studie zur Umstrukturierung des AMS in teilautonome Serviceeinrichtung sehen Penz et al. (2015, S. 27) im Wandel bürokratischer Verwaltungsarbeit in kundenorientierte Interaktionsarbeit einen „Widerspruch zwischen dem Effizienzdruck“ dem AMS-Berater*innen ausgesetzt sind „und der geforderten Personalisierung der Beratung“. Sie sollen auf die Bedürfnisse von Arbeitssuchenden eingehen, die sich oftmals in vulnerablen Lebenssituationen befinden und dieser Interaktion „einen persönlichen Charakter“ geben (Penz et al. 2015, S. 28). Kontinuierlich wachsende Fallzahlen – vor allem in Wien – führen aber dazu, dass „die ArbeitsvermittlerInnen in der Interaktion mit ihren KlientInnen distanziert (bleiben) und darum bemüht (sind), die Vielzahl an Fällen möglichst effizient zu bearbeiten.“ (Penz et al. 2015, S. 28) Vor diesem Hintergrund kann es durchaus entlastend für Berater*innen sein, die mitunter schwierigen Entscheidungen mit einer scheinbar „objektivierten“ computergenerierten Einstufung von Kund*innen zu legitimieren. Bereits vor der Einführung des AMS-Algorithmus hat die Nutzung des Computers während des Kund*innenkontakts zu einer Formalisierung der Beratungssituation und der Distanzierung des*r Berater*in von den Kund*innen geführt (Penz et al. 2017, S. 553).

Formalisierung
der Beratungspraxis

Mit der Fokussierung auf Kennzahlen geht jedoch auch eine Einschränkung des Ermessensspielraums der AMS-Berater*innen einher. Durch die konkreten Vorgaben und Tasks bleibt wenig Spielraum ihre Erfahrungen und Einschätzungen einzubringen; zumal nun viel Zeit und Energie für den Umgang mit der errechneten Einstufung aufgewendet werden muss. AMAS greift in alle vier Phasen des Kernprozesses „Arbeitskräfte unter-

AMAS als praktizierte
„Erstmeinung“

stützen“ ein. Anhand der Segmentierung soll der Beratungsprozess standardisiert werden und einem Schema folgen, das durch AMAS vorgegeben wird. Je nach Segmentzuteilung werden unterschiedliche „Programme“ aktiviert: unterschiedliche Terminintervalle, Maßnahmenpakete und, letztendlich, unterschiedlich lange Verweildauer der Kund*innen am AMS oder in externen Betreuungseinrichtungen (BBEN). Der Fokus des Kund*innenkontakts verschiebt sich vom persönlichen Förderbedarf einer Einzelperson hin zu Arbeitsmarktchancen auf Basis einer Populationsberechnung. Damit ändert sich die Arbeitspraxis von AMS-Berater*innen gravierend. Anstatt ein offenes Gespräch mit Arbeitssuchenden über ihre Biographie und ihren Förderbedarf zu führen, gilt es eine Reihe von „Tasks“ abzuarbeiten. Die Arbeitspraxis wird formalisiert und durch vorgegebene Schritte strukturiert. Die holistische Sichtweise auf die arbeitssuchende Person wird dabei auf ein Abarbeiten einer Einstufung reduziert, auch wenn dieser Prozess gemeinsam – „dialogisch“ – mit den Arbeitssuchenden erfolgen soll. Die Gesprächsführung ist stark durch die Handhabung des Scores definiert, wie der standardisierte Fragenkatalog zu AMAS mit vorgefertigten Antworten an Kund*innen-Fragen zeigt. Während vor AMAS Ermessensspielräume genutzt werden konnten, um eine Person und deren Motivation und Kompetenzen einzuordnen, sind Berater*innen nun angehalten strikt nach Vorgaben der Einstufung vorzugehen oder diese bei Bedarf zu ändern²⁹. Bei Änderung der Einstufung ist eine Erklärung erforderlich, also eine Legitimation, die Zeit und Engagement braucht, die nicht immer vorhanden ist oder sein kann, wie wir unten näher ausführen. AMAS wird in den AMS Dokumenten durchgehend als „zweite Meinung“ für Berater*innen beschrieben, welche aufgefordert sind als „soziales Korrektiv“ zu fungieren. Ein Blick auf die detaillierten Handlungsanweisungen und die technische Umsetzung im System zeigt allerdings, dass die Gefahr groß ist, dass AMAS in der Praxis als „Erstmeinung“ gehandhabt wird und den Beratungsprozess stark prägt (vgl. Eubanks 2017).

Spannungsverhältnis auf Ebene der Arbeitssuchenden

Reduktion der Biographie auf Score

Mit der Verschiebung von persönlichem Förderbedarf einer Einzelperson hin zu Arbeitsmarktchancen auf Basis einer Populationsberechnung ändert sich die Beratungssituation auch für Arbeitssuchende drastisch. Der Anspruch des Systems ist eine bessere, maßgeschneiderte „personalisierte“ Betreuung von Arbeitssuchenden zu ermöglichen. Das „neue Personenmerkmal“ soll Berater*innen die Möglichkeit bieten, Arbeitssuchende und deren Chancen auf dem Arbeitsmarkt besser einschätzen zu können, so das Versprechen. Die Idee der personalisierten Betreuung steht jedoch in einem Spannungsverhältnis zu einer potentiellen Reduktion der Bio-

²⁹ Diese Ermessensspielräume sollten laut Penz et al. (2015, S. 24) aber auch vor der Einführung von AMAS „im Sinne der ökonomischen Vernunft nutzbringend“ ausgestaltet werden.

graphie auf einen Score, den es abzuarbeiten gilt, statt sich an Unterstützungs- und Förderbedarfen von Arbeitssuchenden zu orientieren.

Arbeitsvermittler*innen im (semi-)staatlichen Bereich haben ein Doppelmandat, so Penz et al. (2015, S. 21): „Sie sind dazu angehalten, den Arbeitssuchenden beratend und motivierend zur Seite zu stehen sowie deren individuelle Bemühungen im Rahmen der Stellensuche zu kontrollieren.“ Mit Instrumenten, wie genau festgelegten Kontaktintervallen und vor Ort von den Arbeitssuchenden zu unterschreibenden Betreuungsvereinbarungen, werden „in den Interaktionen symbolische Herrschaftsverhältnisse konstituiert“ (Penz et al. 2015, S. 21). Die Einführung von AMAS bedeutet für Arbeitssuchende, dass nun ein neues Charakteristikum ihrer Person existiert, das sie nur marginal beeinflussen können. Weder Alter, Geschlecht, noch gesundheitliche Einschränkung sind Eigenschaften, die beliebig zu ändern sind. Zudem hat die Einstufung einer Person nur eingeschränkt mit der eigenen Person zu tun, da sich der Wert aus einer Populationsberechnung speist, die auf Daten aus der Vergangenheit zurückgreift. Für Arbeitssuchende bedeutet das, dass ihre Biographie und Fähigkeiten auf einen scheinbar „objektiven“ Wert reduziert werden, der Auskunft über ihre Chancen am Arbeitsmarkt geben soll. Während Arbeit einen zentralen identitätsstiftenden Faktor für Menschen darstellt, beruht eine Identitätszuweisung als „Service-, Beratungs- oder Betreuungskund*innen“ auf deren Arbeitslosigkeit und numerischen kurz- und langfristigen IC-Werten. Daran geknüpft sind die Vorgabe der Kontaktintervalle und eine Spezifizierung des Angebots, das im Rahmen einer Betreuungsvereinbarung von Seiten der Kund*innen verhandelt werden kann. Der Score gibt ein Schema vor, dem zu folgen ist oder das abzuändern ist – mittels BAM und Begründung. Statt von der Biographie und dem Förderbedarf der Einzelperson zu starten, müssen sich Arbeitssuchende ebenfalls mit dem Score auseinandersetzen. Statt ein segmentiertes Angebot anzubieten und an selbstbestimmte Kund*innen zu kommunizieren, segmentiert AMAS die Personen und knüpft standardisierte Maßnahmenpakete daran.

statistischer Wert statt individueller Förderbedarf

Nachdem das System in der Praxis bestimmte Anreize setzt ihm zu folgen, besteht für Arbeitssuchende die Gefahr auf einen numerischen Wert reduziert zu werden, der die eigenen Perspektiven und Wünsche in den Hintergrund verschiebt. Besonders motivierte und engagierte Arbeitssuchende können in diesem System an ihre Grenzen stoßen, wenn sie ihre Ideen und Berufswünsche einbringen wollen. Zudem zieht eine schlechte Einstufung häufig Demotivation und ein Gefühl der Machtlosigkeit auf Seiten der Arbeitssuchenden nach sich (vgl. Gangadharan & Niklas 2019). Deutsche Ungleichheitssoziolog*innen haben etwa festgestellt, dass Schüler*innen, die eine Empfehlung fürs Gymnasium bekommen, motiviert sind und ihre Leistung weiter steigern, während Schüler*innen, die keine Empfehlung bekommen, ins Phlegma verfallen und noch weiter zurück fallen (Billmann-Mahecha & Tiedemann 2006; vgl. auch Staunæs & Brøgger 2020). Dies zeigt, dass die Zuordnung zu Gruppen die eigenen Erwartungshaltungen und Möglichkeiten beeinflussen und damit identitätsstiftend sein

Demotivation durch niedrige Einstufung

kann. Ähnliche Dynamiken könnten mit der Einführung von AMAS hervorgerufen werden (vgl. Pultz 2016).

Ambivalenzen auf institutioneller Ebene

mehr Zeit und
Engagement nötig

Die Erwartungshaltung des AMS, mittels AMAS eine effizientere und effektivere Vergabe von Fördermitteln zu erzielen, steht in einem Spannungsverhältnis zu Anforderung an das System als „Zweitmeinung“ zu fungieren und an Berater*innen als „soziales Korrektiv“ zu agieren. Die Ziele, die an AMAS herangetragen werden, umfassen effizientere Beratungsgespräche durch das zusätzliche „Personenmerkmal“ und mehr Beratungszeit im „mittleren“ Segment, da künftig weniger Beratungszeit für Arbeitssuchende aus den Gruppen „hoch“ und „niedrig“ aufgewendet werden soll. Der Fokus liegt damit auf dem „Kerngeschäft“ der AMS, nämlich der Betreuung von „Beratungskund*innen“. Schließlich wird durch die evidenzbasierte Entscheidung und daran geknüpfte Standardisierung der Abläufe willkürlicher Fördermittelvergabe vorgebeugt und ein effektiverer Einsatz von Förderungen erhofft. Das „mittlere“ Segment könne die Förderungen am besten in eine kürzere Vermittlungsdauer umsetzen, so die Erwartungshaltung. Diese Hoffnung wird jedoch durch die Praxis von AMAS als Zweitmeinung konterkariert. Entsprechend den Handlungsanleitungen für AMS-Berater*innen sollen diese die jeweilige Einstufung mit den Arbeitssuchenden dialogisch diskutieren. Es wird von ihnen erwartet, die Einstufung genau zu erklären, diese mit der eigenen Einschätzung der Person abzugleichen und gemeinsam mit der arbeitssuchenden Person zu einer Letztentscheidung zu gelangen, welche in der Betreuungsvereinbarung festgehalten wird. Dieses Vorgehen ist nicht trivial, wenn es ernst genommen wird. Eine ausführliche Unterhaltung über die Einstufung und eine allfällige Umstufung benötigt mehr Zeit und Engagement, als den CAM Wert einfach zu übernehmen. Wenn ein BAM angelegt wird, muss eine Begründung gefunden und festgehalten werden, die die Umstufung – und einen damit verbundenen Mehraufwand an Ressourcen, z. B. wenn von N auf M umgestuft wird – darlegt und legitimiert. Angesichts der kurzen Zeit, die AMS-Berater*innen pro Beratung zur Verfügung steht (etwa zwischen dreißig und in den östlichen Bundesländern nur zehn Minuten [DSB 2020, S. 5]), gibt es einen gewissen Anreiz den CAM zu übernehmen und die Unterhaltung mit Kund*innen darüber knapp zu gestalten. Im Bescheid der Datenschutzbehörde [DSB 2020] wurde dieses Vorgehen als „routinemäßige Übernahme“ beschrieben (siehe Kapitel 2).

neue Kompetenzen
sind erforderlich

Das heißt, dass zwar die Letztentscheidung und Verantwortung bei den AMS-Berater*innen liegt, dass das System und seine Einbettung in die AMS Praxis, jedoch deutliche Anreize setzt eher dem CAM zu folgen, als diesen zu „korrigieren“. Zudem ist die Erklärung des Scores nicht einfach, wenn sie über die beiden „Segmentzusatzinformationen“ hinaus gehen soll. Die Einstufung bezieht sich auf einen Wahrscheinlichkeitswert auf Basis einer ausgewählten Population, die Betreuungsvereinbarung muss jedoch mit einer konkreten Einzelperson geschlossen werden. Dieses Vorgehen erfordert neue Kompetenzen mit Korrelationen und Kausalitäten

umzugehen und diese, in einem zweiten Schritt, den Kund*innen konzipie zu erklären. Es braucht daher ganz neue Fähigkeiten, wenn Berater*innen den komplexen Anforderungen des Systems und den Vorgaben des AMS gerecht werden sollen. Diese erforderlichen Fähigkeiten – vom Umgang mit numerischen Personenmerkmalen, dem Nachvollziehen von Wahrscheinlichkeitsberechnungen bis hin zu Anti-Diskriminierungsmaßnahmen – müssen in der Ausbildung von AMS-Berater*innen berücksichtigt und gefördert werden. De facto ändert sich die Berufsbeschreibung von AMS-Berater*innen durch die Einführung von AMAS drastisch.

Die datenbasierte Entscheidungsunterstützung stellt im Rahmen der Einführung von New Public Management einen weiteren Schritt in Richtung „Modernisierung“ der Institution AMS dar und ist als Dienstleistungsrationalisierung zu charakterisieren. AMAS kann dabei als weiteres Tool verstanden werden, das „sowohl die Effizienz der staatlichen Organisation als auch die Marktfähigkeit der KundInnen befördert“ (Penz et al. 2015, S. 24).


5 SZENARIEN ZUR ILLUSTRATION DER ERGEBNISSE

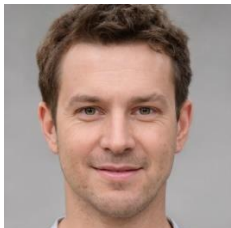

Im Folgenden stellen wir vier konkrete Szenarien vor, die die Ergebnisse unserer Analyse insbesondere in Hinblick auf unterschiedliche Folgen von Bias in der Beratungspraxis verdeutlichen. Nachdem wir im Rahmen dieser Studie weder Zugriff auf die Daten des AMS, sowie auf AMAS selbst hatten, wie eingangs beschrieben, sind die Szenarien als plausible Annäherungen auf Basis der uns zur Verfügung stehenden Dokumente und Materialien zu verstehen. Sie können keine tatsächliche IC-Werte oder Fehlerraten berechnen, aber sie können die mit AMAS verbundenen Probleme im Zusammenspiel von technischem Bias und sozialer Praxis plausibel illustrieren. Um ein algorithmisches System tatsächlich evaluieren zu können, bräuchte es neben Zugang zu Daten und System zudem eine evidenz-basierte Gegenüberstellung von Vorher-Nachher Bestandsaufnahmen, die in Bezug auf AMAS nicht vorliegen, soweit uns bekannt ist. Nur wenn Vergleichsdaten zur Situation vor der Einführung des Systems vorliegen, kann eine Verbesserung oder Verschlechterung von Effizienz, Effektivität und Fairness von Fördervergaben konkret evaluiert werden.


Jedes der vier Szenarien hat das Ziel ein konkretes Problem, das sich in der praktischen Handhabung von AMAS ergibt, zu veranschaulichen. Anhand von fiktiven Personen³⁰ und Biographien gehen wir dabei von möglichst realen Situationen aus, die wir auf Basis unserer Ergebnisse gebildet haben und in denen wir die Schritte der Prozessbeschreibung der KP1 Richtlinie [RICHT_2 2019] nachvollziehen. Das erste Szenario verdeutlicht das Problem der Fehlerraten von AMAS, indem es eine Person trotz vollständiger Datensätze in eine falsche Gruppe einstuft und aufzeigt welche Auswirkungen sich daraus auf individueller und institutioneller Ebene ergeben. Das zweite Szenario beschäftigt sich mit einem technischen Bias und welche Konsequenzen sich daraus für die AMS Beratungspraxis ergeben. Das dritte Szenario problematisiert einen gegebenen und emergenten Bias, der sich über die Zeit am Arbeitsmarkt herausgebildet hat und über AMAS weiter verfestigt wird, und illustriert welche Implikationen sich daraus für individuelle Arbeitssuchende ergeben. Das vierte und letzte Szenario verdeutlicht, welche Auswirkungen die Handhabung von AMAS in der Beratungspraxis hat, indem es insbesondere die Standardisierung von Abläufen durch den IC-Wert in den Blick nimmt.

³⁰ Die für die Szenarien verwendeten Fotos sind computergeneriert und stammen von der Website <https://generated.photos>

 <p>Martin B. Herr Martin B. ist 27 Jahre alt und war einige Jahre als Reinigungskraft in einem familiär geführten Betrieb tätig. Leider hat er recht plötzlich seine Anstellung verloren und musste sich daher erstmals arbeitslos melden.</p>	<p>Variablenausprägung Geschlecht: Männlich Alter: 27 Ausbildung: Pflichtschulabschluss Staatenzugehörigkeit: Österreich RGS-Typ: 1 (Eferding) Sektor: Dienstleistungen Population: Basispopulation – vollständige Daten (fiktiver) IC-Wert kurzfristig: 52 % (fiktiver) IC-Wert langfristig: 63 % Segment: M (Mittel)</p>
<p>Beratungssituation: Berater*in</p> <ul style="list-style-type: none"> • Berater*in sieht sich Personenstammdaten von Martin B. an und denkt, dass die Einstufung als M trotz niedriger Ausbildung plausibel ist, da der Kunde noch jung ist und in einem vorteilhaften Bezirk lebt. Die Einstufung wird beibehalten. • Ermutigender Erklärungstext „<i>Sie waren über einen sehr langen Zeitraum in den letzten Jahren beschäftigt.</i>“ Aber „<i>Sie haben keine über das Pflichtschulniveau hinausgehende Berufsausbildung absolviert</i>“ als besondere Herausforderung • Aufgrund der Zuteilung zu Segment M und weil noch Kursplätze frei sind, wird eine zweimonatige Maßnahmenteilnahme vereinbart. 	<p>Beratungssituation: Kund*in</p> <ul style="list-style-type: none"> • Der Kunde ist sich bewusst, dass er keine Ausbildung hat, aber bisher war er erfolgreich. • Martin B. hätte es vorgezogen, erst einmal eine neue Stelle zu suchen, widerspricht aber auch nicht der Maßnahmenteilnahme.
<p>Resultat der operativen Umsetzung</p> <ul style="list-style-type: none"> • Beibehaltung der Einstufung als M • Martin B. wird in drei Wochen für die Dauer von zwei Monaten eine Maßnahme besuchen. 	
<p>Relevanter Bias</p> <p>Die Trefferquoten für Segment M sind nicht dokumentiert, die Trefferquoten bewegen sich für die anderen Segmente im Bereich 75-90%. Bei insgesamt ca. 600.000 Geschäftsfällen können daher ca. 60.000-150.000 Personen falsch eingestuft werden. Das ist auch bei Martin B. der Fall. Eigentlich sind seine Chancen hoch: In Eferding werden gerade viele Arbeitskräfte gesucht, darunter auch Reinigungskräfte.</p>	
<p>Zielsetzungen des AMS</p> <p>AMAS hat zum Ziel, die Effizienz der Fördermittelvergabe zu steigern und den Beratungsprozess zu unterstützen. Da Martin B. tatsächlich „hohe“ (H) Integrationschancen hat, läuft dieses Szenario den Zielsetzungen Effektivität (Besuch einer nicht notwendigen Maßnahme) und Effizienz (Ressourcen für die Maßnahme) zuwider. Die Effektivität könnte weiters verschlechtert werden, wenn aufgrund der Maßnahmen-Teilnahme weniger Energie in die Stellensuche gesteckt werden kann – und sich somit der Wiedereintritt in die Beschäftigung verzögert. Für Herrn B. bedeutet das darüber hinaus, dass er länger mit den finanziellen Einbußen des Arbeitslosengeldes konfrontiert ist. Die Fehleinstufung führt auch zu längeren Arbeitslosengeld-Zahlungen als notwendig.</p>	

 <p>Schifteh A. Frau Schifteh A. ist iranische Staatsbürgerin, die vor 3 Jahren nach Wien migriert ist, um an der TU Wien ihren Master in Informatik zu absolvieren. In dieser Zeit war sie Studienassistentin an der TU Wien. Sie ist nun, nach Studienabschluss und Ende ihrer Anstellung arbeitslos gemeldet und auf der Suche nach einer Arbeit in der IT Branche. Ihre tatsächlichen Aussichten sind exzellent (Segment H); sie hat Deutschkenntnisse auf B1-1 Niveau, ihr Englisch und Farsi sind exzellent, und in den meisten größeren IT-Firmen ist englische Kommunikation Standard. Sie ist hochmotiviert, erwartet keine Probleme, eine Arbeit zu finden, und hat bereits potentielle Stellen auf Ihrer Bewerbungsliste.</p>	<p>Variablenausprägung Gender: Weiblich Ausbildung: Matura + Staatszugehörigkeit: Drittstaaten RGS-Typ: 5 (1100 Wien) Sektor: Dienstleistungen Population: Partiiell valide (Migrationshintergrund) (fiktiver) IC-Wert: 22 % Segment: N (Niedrig)</p>
<p>Beratungssituation: Berater*in</p> <ul style="list-style-type: none"> • Beraterin vergleicht IC-Wert mit Biographie • Es werden keine Erklärungstexte angezeigt • Basierend auf Ausbildung und Sparte stuft sie die Kundin auf (inkl. Begründung des neuen BAMM); der ursprüngliche CAMN legt der Beraterin trotz sehr guter Jobaussichten lediglich eine Aufstufung auf M nahe • Nach Aufstufung liegen verschiedene Maßnahmen nahe (z. B. Deutsch-Kurs) 	<p>Beratungssituation: Kund*in</p> <ul style="list-style-type: none"> • Kundin erwartet gute Chancen und plant, selbstständig Jobangebote zu suchen und auf Inserate zu antworten • Nach Niedrig-Einstufung und Aufstufung hat sie Zweifel an ihrer schnellen Vermittelbarkeit • Kundin nimmt das Maßnahmenangebot an (da sie ohnehin vorhatte, ihr Deutsch zu verbessern)
<p>Resultat der operativen Umsetzung</p> <ul style="list-style-type: none"> • Aufstufung von N auf M • Demotivation der Kundin und Erwartungshaltung, dass Maßnahmen zur Jobaufnahme nötig sind • Verzögerung der Jobaufnahme durch erhöhte Betreuungs-Intervalle und ev. Maßnahmen-Teilnahme 	
<p>Relevanter Bias</p> <p>Das Informationsdefizit von AMAS in Bezug auf detaillierte Informationen zur Ausbildung (Matura+ inkludiert auch höhere Ausbildungsgrade) führt in Kombination mit den anderen Merkmalen zu einer signifikanten Unterschätzung der Integrationschance. Durch die Stratifikation in gleiche Konstellationen kommt es in diesem Szenario zu Vergleichen mit völlig anderen Personenhintergründen, die aufgrund anderer Voraussetzungen (Bildungsgrad, genauer Arbeitsmarkt) tendenziell schlechtere Chancen haben. Da die Kundin in der Population ‚partiell valide schätzbar mit Migrationshintergrund‘ geführt wird, gibt es auch keine Erklärungstexte für ihre niedrige Bewertung. Dies konstituiert einen technischen Bias.</p>	
<p>Zielsetzungen des AMS</p> <p>Gemäß der Zielsetzung ‚Effizienz‘ und ‚Effektivität‘ für das System sind in diesem Fall eher Verschlechterungen als Verbesserungen zu erwarten. Die Beratungssituation wird durch widersprüchliche Einschätzungen zwischen Menschen und System verkompliziert; die fehlenden Erklärungstexte/Segmentzusatzinformationen machen die Evaluierung des Fehlerausmaßes schwierig. Die ungenügende Aufstufung in das mittlere Segment führt zu erhöhtem Zeit- und Ressourcenaufwand.</p>	

	<p>Michael S. Herr Michael S. ist 34 Jahre alt und war nach einer Lehre als Maschinenschlosser mit kurzen Unterbrechungen in einem metallverarbeitenden Betrieb beschäftigt.</p> <p>Aufgrund seiner Qualifikation und Beschäftigungshistorie hat er grundsätzlich sehr gute Aussichten einen Arbeitsplatz zu finden. Aufgrund von Betreuungspflichten sucht Herr S. eine Teilzeitarbeit, was in seiner Branche eher unüblich ist.</p>	<p>Variablenausprägung Geschlecht: männlich Alter: 34 Ausbildung: Lehre Sektor: Produktion Population: Basispopulation – vollständige Daten (fiktiver) IC-Wert kurzfristig: 66 % (fiktiver) IC-Wert langfristig: 72 % Segment: H (Hoch)</p>		<p>Sabine K. Frau Sabine K. ist 34 Jahre alt und war nach einer Lehre als Bürokauffrau mit kurzen Unterbrechungen in einem metallverarbeitenden Betrieb beschäftigt.</p> <p>Aufgrund ihrer Qualifikation und Beschäftigungshistorie hat sie sehr gute Aussichten einen Arbeitsplatz zu finden. Sie hat Betreuungspflichten, die sie mit ihrem Partner flexibel teilt und sucht eine Voll- oder Teilzeitstelle.</p>	<p>Variablenausprägung Geschlecht: weiblich Alter: 34 Ausbildung: Lehre Betreuungspflichten: ja Sektor: Produktion Population: Basispopulation – vollständige Daten (fiktiver) IC-Wert kurzfristig: 59 % (fiktiver) IC-Wert langfristig: 65 % Segment: M (Mittel)</p>
<p>Beratungssituation</p> <ul style="list-style-type: none">• Erklärungstexte: „Sie waren über einen sehr langen Zeitraum in den letzten Jahren beschäftigt.“ „Sie verfügen über eine Berufsausbildung mit formalem Abschluss zumindest auf Facharbeiterniveau.“• Kunde bezweifelt seine Einstufung aufgrund seiner Präferenz für eine branchenunübliche Teilzeitstelle und führt seine Betreuungspflichten als Grund an.• Es wird eine Abstufung auf M vereinbart (inkl. Begründung des BAMM).			<p>Beratungssituation</p> <ul style="list-style-type: none">• Erklärungstexte: „Sie waren über einen sehr langen Zeitraum in den letzten Jahren beschäftigt.“ „Sie haben Betreuungspflichten in ihrem familiären Umfeld.“• Kundin wird aufgefordert, über Betreuungspflichten Auskunft zu geben.• Eine Beibehaltung des CAMM oder eine mögliche Aufstufung in H (inkl. Begründung des BAMH) wird abgewogen.		
<p>Resultat der operativen Umsetzung</p> <ul style="list-style-type: none">• Abstufung von H auf M• Möglichkeit qualifizierende Maßnahmen zu vereinbaren			<p>Resultat der operativen Umsetzung</p> <ul style="list-style-type: none">• Reproduktion des Stereotyps der Zuständigkeit von Frauen für Betreuungspflichten• Ungleichbehandlung in der Einstufung trotz Irrelevanz der Betreuungspflichten der Kundin		
<p>Gegebener, technischer und emergenter Bias</p> <p>Entsprechend des gesellschaftlich gegebenen Stereotyps, dass Frauen mit Kindern durch Betreuungspflichten in ihrer Erwerbstätigkeit eingeschränkt seien, kommt die Variable ‚Betreuungspflichten ja/nein‘ nur bei weiblichen Arbeitssuchenden zur Anwendung. Die Erwerbstätigkeit von Männern sei nicht durch Betreuungspflichten eingeschränkt und daher kommt diese Variable bei männlichen Arbeitssuchenden nicht zur Anwendung. In der Regressionsanalyse bildet sich diese Diskriminierung in den Daten über Arbeitssuchende ab und Frauen mit Kind haben demnach, unabhängig davon ob sie durch Betreuungspflichten tatsächlich eingeschränkt sind, einen Nachteil am Arbeitsmarkt. Ausgehend von dieser Beobachtung wurde die Variable ‚Betreuungspflichten ja/nein‘ in die Berechnung des IC-Wertes von Frauen aufgenommen, nicht aber in die Berechnung des IC-Wertes von Männern. Bei ansonsten gleichen Variablenausprägungen wird Herr S. aufgrund der technischen Einschränkung, die durch die Wahl von diskreten Schwellenwerten ohne Graubereiche entsteht, in Gruppe H eingestuft, während Frau K. in Gruppe M eingestuft wird. Der gesellschaftliche Bias wird hier in einen technischen Bias übersetzt. Dieser führt den Bias in der Beratungssituation fort und schreibt damit gesellschaftliche Werte, die im Wandel begriffen sind, fest.</p>					
<p>Zielsetzungen des AMS</p> <p>Neben den Zielen ‚Effizienz‘ und ‚Effektivität‘ ist das AMS auch dem Grundsatz der Gleichbehandlung sowie der Frauenförderung verpflichtet. Der CAMM eröffnet Frau K. grundsätzlich die Möglichkeit beschäftigungsfördernde Maßnahmen und Frauenfördermaßnahmen in Anspruch zu nehmen. Sie erfährt allerdings keine Gleichbehandlung, weil sie schlechter eingestuft wird und damit ihre Betreuungsintervalle höher sind.</p>					

 <p>Harald G. Herr Harald G. ist 49 Jahre alt und hat sein Leben lang als Kellner in der Tourismusbranche gearbeitet. Aufgrund einer langwierigen Knie-Operation, die ihm längeres Stehen und Gehen unmöglich machte, hatte er zuletzt längere arbeitslose Perioden. Zudem musste er seine Mutter über längere Zeit pflegen und konnte daher Graz nicht verlassen. Kurz vor der Corona-Krise hat er eine Stelle in einem Grazer Stadthotel angenommen, die mit seiner gesundheitlichen Einschränkung zu vereinbaren war. Aufgrund der Krise wurde er nun wieder gekündigt. Er kommt hochmotiviert zum Erstgespräch und möchte sich gerne umschulen lassen, da er mit seinen Einschränkungen langfristig nicht mehr als Kellner arbeiten kann. Seine Betreuungspflichten sind beigelegt und er möchte beruflich neu durchstarten jetzt wo er zeitlich wieder flexibel ist.</p>	<p>Variablenausprägung Gender: M Alter: 49 Ausbildung: Lehre Sektor: Dienstleistungen Gesundheitliche Einschränkung: Ja Population: Basispopulation – vollständige Daten (fiktiver) IC-Wert kurzfristig: 23 % (fiktiver) IC-Wert langfristig: 18 % Segment: N (Niedrig)</p>
<p>Beratungssituation: Berater*in</p> <ul style="list-style-type: none"> Beraterin sieht sich seine Biographie und die Einstufung N an. Sie ist erst seit kurzem beim AMS und folgt aufgrund von geringer Erfahrung und Zeitdruck der Einstufung. Sie erklärt mithilfe der Erklärungstexte warum er in Segment N gelandet ist: „<i>Sie waren nur über kurze Zeiträume in den letzten Jahren beschäftigt.</i>“ und „<i>Sie weisen gesundheitliche Einschränkungen auf, die die Arbeitssuche erschweren.</i>“ Beraterin bietet dem Kunden die Teilnahme an dem BBEN Betreuungsformat an, da er noch keinen Anspruch auf Fördermöglichkeiten für Menschen 50+ hat. 	<p>Beratungssituation: Kund*in</p> <ul style="list-style-type: none"> Herr G. fühlt sich missverstanden und abgestempelt und hinterfragt die Einstufung. Er möchte keine Stabilisierungsmaßnahme, sondern eine Qualifizierungsmaßnahme. Vor seinem Erstgespräch hat er recherchiert und möchte sich zum Restaurantmanager umschulen, da er Erfahrung in der Branche hat und diese Tätigkeit mit seinen Knieproblemen vereinbar ist. Er fordert aktiv die Weiterbildung „Management und Marketing in Hotellerie, Gastgewerbe und Tourismus“ ein.
<p>Resultat der operativen Umsetzung</p> <ul style="list-style-type: none"> Algorithmus stuft Herrn G. in N ein, da hohe Motivation und Eigeninitiative nicht in die Berechnung einfließen. CAMN wird ‚routinemäßig‘ übernommen, da AMS-Berater*in wenig Zeit und Erfahrung hat. Herrn G.s Alter ist sehr nahe am Grenzwert von 50 Jahren, was weder durch den Score noch von der Beraterin berücksichtigt wird. 	
<p>Relevanter Bias</p> <p>Herr G. hat u. a. aufgrund seiner gesundheitlichen Beeinträchtigung und seines Alters bereits Schwierigkeiten, eine Stelle zu finden. Beides sind auch entscheidende Einflussfaktoren für seine Einstufung in das Segment „niedrig“ (N). Dadurch wiederum wird es ihm erschwert, Zugang zu hilfreichen Maßnahmen zu bekommen, was zu sogenannter kumulativer Benachteiligung führt. Außerdem bildet AMAS seine komplexen Umstände nicht adäquat ab.</p>	
<p>Zielsetzungen des AMS</p> <p>Das Beratungsgespräch gestaltet sich langwierig und für beide Seiten höchst frustrierend. Herr G. beginnt seine erwünschte Förderung einzufordern, die Beraterin erklärt, dass seine Einstufung zu schlecht sei. Statt über seine Potentiale zu sprechen, reden sie primär über die Einstufung und ob/wie diese zu ändern sei. Die Beraterin kann in der eingeschränkten Zeit nicht beurteilen, ob eine teure Qualifizierungsmaßnahme zielführend ist. Sie hat schon einige andere Kund*innen umgestuft und ist in diesem Fall unsicher, ob eine Umstufung gerechtfertigt ist. Dass für Herrn G. eine BBEN Maßnahme effektiv ist, ist zu bezweifeln, da sich seine gesundheitliche Einschränkung nicht ändert und er ansonsten fit für den Arbeitsmarkt ist. Eine Qualifizierungsmaßnahme scheint in seinem Fall die einzige Möglichkeit zu sein, der Langzeitarbeitslosigkeit zu entkommen, da ihm 50+ Maßnahmen noch nicht zustehen (und zudem darin enthaltene finanzielle Anreizsysteme für Unternehmen in seinem Fall wenig sinnvoll erscheinen, da in seiner derzeitigen Branche wenige Stellen mit seiner gesundheitlichen Einschränkung vereinbar sind).</p>	

6 VERGLEICHBARE SYSTEME IN ANDEREN LÄNDERN

Systeme zur algorithmisch unterstützten Entscheidungsfindung sind im Bereich des Arbeitsmarktservices in der Form des AMS-Algorithmus in Österreich neu. Im internationalen Vergleich zeigt sich, dass in anderen Ländern vergleichbare Systeme bereits seit längerem im Einsatz sind bzw. waren. Erste Systeme zum statistischen Profiling von Erwerbsarbeitslosen wurden etwa bereits 1993 in den USA und 1994 in Australien eingeführt (Loxha & Morgandi 2014). Aber auch nicht-statistische Formen des Profiling von Erwerbsarbeitslosen sind in vielen Ländern im Einsatz.

Dieses Kapitel bietet einen Einblick in Systeme zum Profiling von Erwerbsarbeitslosen in anderen Ländern. Auch wenn sich die Arbeitslosenversicherungssysteme in diesen Ländern vom Österreichischen System unterscheiden, bietet der internationale Vergleich algorithmischer Entscheidungsunterstützungssysteme wichtige Erkenntnisse für die Diskussion in Österreich. Einem Überblick über die Bandbreite an verwendeten Daten und Methoden folgt ein Blick auf die Verknüpfung von Profiling und dem Zugang zu Ressourcen. Schließlich werden ausgewählte Länder ausführlicher betrachtet.

6.1 DATEN & METHODEN

Profiling-Systeme in anderen Ländern zeigen eine große Bandbreite an verwendeten Daten, eingesetzten Methoden und deren Einbettung in Prozesse in den jeweiligen Arbeitsmarktservice-Einrichtungen.

Bei vielen Systemen kommen administrative Daten zum Einsatz, das heißt Daten, die bereits für andere Zwecke bei den Arbeitsmarktservice-Einrichtungen vorhanden sind. Bei manchen Systemen bilden sie die einzige Datenbasis (z. B. Italien, Schweden), bei anderen werden sie mit anderen Daten kombiniert (z. B. USA, Irland, Lettland) (Desiere, Langenbacher & Struyven, 2019). Ein weiterer häufiger Datentyp sind Interviews (z. B. Australien, Lettland) und Fragebögen, wobei diese überwiegend online durchgeführt werden (z. B. Dänemark, Niederlande, USA). Teilweise finden diese Erhebungen während der Meldung als erwerbsarbeitslos statt (Irland). Neben den genannten Datentypen gibt es selten auch ausgefallenerere Daten, wie etwa das Klickverhalten auf den Webseiten des Flämischen Arbeitsmarktservice (Desiere et al. 2019).

Bei den Profiling-Methoden können grob drei Ansätze identifiziert werden (vgl. Barnes, Wright, Irving & Deganis 2015). Der erste Ansatz setzt

auf die *Berater*innen* zum Erstellen von Profilen von Erwerbsarbeitslosen. Die persönliche Erfahrung und Einschätzung der Berater*in, möglicherweise unter Zuhilfenahme von administrativen Daten bilden die Grundlage des Profilings. Der zweite ist ein *regelbasierter Ansatz*. Hier variiert z. B. die Betreuung mit der Dauer der Erwerbsarbeitslosigkeit, oder je nach Gruppenzugehörigkeit gibt es unterschiedliche Angebote. *Statistisches Profiling* ist der dritte Ansatz. Mit statistischen Modellen werden Schätzungen bzw. Vorhersagen getroffen. Oft wird Langzeitarbeitslosigkeit vorhergesagt (wobei die Definitionen von Langzeitarbeitslosigkeit variieren), in den USA hingegen ist es eine Vorhersage über die Wahrscheinlichkeit, die Arbeitslosigkeitsansprüche auszuschöpfen. Unter den Methoden des statistischen Profiling dominiert logistische Regression (Australien, Italien, Niederlande, Schweden und USA). Weitere Methoden sind Probit Regression (Irland), Random forest Modelle (Neuseeland, Flandern/Belgien), Gradient boosting (Neuseeland) und Faktorenanalyse (Lettland). Komplexere Machine Learning Methoden – etwa (tiefe) Neuronale Netze – scheinen derzeit keine Verwendung zu finden (Desiere et al. 2019).

Häufig kommt es zu Mischformen dieser Ansätze, z. B. indem vorgegebene Regeln das statistische Profiling aufheben können (z. B. Niederlande) oder statistische Vorhersagen in Kombination mit Berater*innen (z. B. Dänemark) (Desiere et al. 2019).

6.2 PROFILING UND RESSOURCENZUGANG

Der AMS-Algorithmus wird eingesetzt um Erwerbsarbeitslose in drei Gruppen einzuteilen. Diese Segmentierung basiert auf der Wahrscheinlichkeit der Reintegration in den Arbeitsmarkt und soll dem Ziel dienen, Effizienz und Effektivität bei der Ressourcenzuteilung zu erhöhen sowie den Aufwand der Berater*innen zu reduzieren. Diese Motive sind international verbreitet, jedoch fehlen eindeutige wissenschaftliche Studien und Beweise um die Effektivität und Effizienz in der Praxis belegen zu können (Johnson, Scoppetta, & Buckenleib 2018).

Beim österreichischen AMS-Algorithmus ist das Ziel, der Gruppe mit prognostizierter „mittlerer“ Chance auf Arbeitsmarktintegration die meisten Ressourcen zur Verfügung zu stellen, jenen mit schlechter Prognose hingegen weniger Ressourcen [BER_15, 2017]. Konträr dazu wird die überwiegende Mehrheit der Profiling-Systeme im internationalen Vergleich eingesetzt, um Personen mit erhöhtem Risiko von Langzeitarbeitslosigkeit früher und/oder stärker zu fördern (Kureková 2014; Desiere et al. 2019; Johnson et al. 2018). Eine Ressourcenverteilung wie in Österreich ist die Ausnahme. In der EU folgte einzig Polen ebenfalls dieser Logik (Desiere et al. 2019).

6.3 AUSGEWÄHLTE LÄNDERVIGNETTEN

6.3.1 NIEDERLANDE

Die Niederlande sind das EU-Land mit der längsten Geschichte von statistischem Profiling von Erwerbsarbeitslosen. Bereits 1999 wurde der Kansmeter (Chancenmesser) eingeführt, ein statistisches Tool, das die Ferne der Erwerbsarbeitslosen vom Arbeitsmarkt berechnen sollte (Konle-Seidl 2011). Aufgrund einer negativen Evaluierung durch das niederländische Sozial- und Arbeitsministerium wurde es 2007 durch eine Segmentierung in zwei Gruppen ersetzt. Auch dieses System wurde 2009 eingestellt. Stattdessen gab es seit 2009 ein regelbasiertes System, die sogenannte WERKformule. Unter WERKformule ist die Dauer der Erwerbsarbeitslosigkeit ausschlaggebend für die angebotenen Ressourcen.

Bereits seit 2006 wurde in den Niederlanden der Work Profiler entwickelt, aufgrund von umfangreichen Vorstudien und Umstrukturierungen (vor allem die Digitalisierung des niederländischen Arbeitsmarktservices) allerdings erst ab 2013 schrittweise eingeführt (Wijnhoven & Havinga 2014). Work Profiler soll helfen, Faktoren zu erkennen, die die Wahrscheinlichkeit vorhersagen können, binnen 12 Monaten neue Erwerbsarbeit zu finden. In den Vorstudien wurde eine Reihe an „harten“ und „soften“ Faktoren identifiziert. Harte Faktoren sind etwa das Alter und Dauer an der letzten Arbeitsstelle. Softe Faktoren sind z. B. Einschätzungen, wie lange die Jobsuche dauern wird, wie oft es im letzten Monat Kontakte zu möglichen Arbeitgebern gab und welche externen Faktoren Erwerbsarbeitslose als entscheidend bei der Jobsuche einschätzen (Glück, Gelegenheit oder persönliche Netzwerke). Die Daten für Work Profiler werden freiwillig von Erwerbsarbeitslosen über einen online-Fragebogen eingegeben.

Wie viele andere Systeme versucht auch Work Profiler, aus den Daten die Wahrscheinlichkeit, nach 12 Monaten noch arbeitslos zu sein, vorherzusagen (Wijnhoven & Havinga 2014). Und auch beim Work Profiler soll darauf basierend nur einem Teil der Erwerbsarbeitslosen direkter Kontakt zu Berater*innen ermöglicht werden. Darüber hinaus soll der Fragebogen auch dazu dienen, „Schwächen“ zu identifizieren.

Work Profiler steht teilweise im Kontrast zu dem System des österreichischen AMS: Freiwilligkeit, die Wichtigkeit „softer“ Information und die Identifikation von „Schwächen“, die unter Umständen adressiert werden können, sind offensichtliche Unterschiede. Dennoch ähnelt Work Profiler in vielen Punkten dem AMS-Algorithmus, etwa im Einsatz von statistischem Profiling und dem Fokus auf Erwerbsarbeitslose.

6.3.2 POLEN

In Polen wurde statistisches Profiling von Erwerbsarbeitslosen im Jahr 2014 zum ersten Mal eingesetzt (Niklas et al. 2015). Erst kürzlich, 2019, wurde

das System durch einen richterlichen Beschluss wieder abgeschafft. Grund waren Formalfehler bei der Einführung des Systems (Laub 2020). Dieses System ist aus österreichischer Sicht besonders interessant, da seine Grundlagen jenen des AMS-Algorithmus in vielerlei Hinsicht sehr ähneln.

Grund für die Einführung war das begrenzte Budget des polnischen Arbeitsmarktservices. Das neue System wurde als ein modernes, wissenschaftliches und technisches Werkzeug für mehr Effizienz und Effektivität sowie gezieltere Förderung präsentiert. Die Motivation und die Legitimierung sind damit sehr ähnlich wie in Österreich. Die Daten werden im Unterschied zum AMS-Algorithmus aber durch Befragungen erhoben, wobei viele dabei gemessene Variablen denen im System des AMS ähneln: Alter, Geschlecht, Bildung, gesundheitliche Beeinträchtigung, Dauer der Arbeitslosigkeit und Hauptwohnsitz werden abgefragt. Weiters wird auch noch die Motivation der Erwerbsarbeitslosen erhoben. Welche statistischen Modelle zum Einsatz kamen, war nicht öffentlich bekannt.

Das System teilte Erwerbsarbeitslose je nach Chancen-Prognose in drei Gruppen ein. Den Gruppen mit den besten und schlechtesten Chancen wurden reduzierte Ressourcen zugeteilt. Eine Analyse des Handbuchs für Berater*innen stellte stigmatisierende Sprache in der Beschreibung der Gruppe mit den schlechtesten Chancen fest (Niklas, Sztandar-Sztanderska & Szymielewicz 2015). So wurde die Gruppe im Dokument etwa zusammengefasst als eine mit „schwerwiegende Lebensproblemen“ („serious life problems“), die „passiv“ sei, keine oder kaum Bildung/Erfahrung habe, durch Sozialhilfe unterstützt werde und gesundheitliche Probleme habe. Weiters wurde Arbeitslosigkeit als eine Entscheidung der Betroffenen dargestellt und sei mit mentalem Charakter zu begründen. In diese Gruppe fielen zum Beispiel alleinerziehende Mütter, Menschen mit Behinderung oder Menschen aus kleineren Ortschaften, für die es schwer ist zu pendeln (Niklas, Sztandar-Sztanderska & Szymielewicz 2015).

In der Praxis hat sich gezeigt, dass Berater*innen Nachfragen zur Kategorisierung durch Erwerbsarbeitslose als verdächtig beurteilten (Niklas, Sztandar-Sztanderska & Szymielewicz 2015). Im Handbuch wurde weiters ausschließlich dargestellt, wie eine Abstufung innerhalb des Systems durchgeführt werden kann. Im Gegensatz dazu gab es kein Beispiel zur Aufstufung einer Person. Darüber hinaus wurde die Abstufung in dem Beispiel mit dem Verdacht auf falsche Angaben bei der Befragung begründet. Beim polnischen System hat sich gezeigt, dass bloß 0,58 % der Zuteilungen des statistischen Profiling-Systems durch Berater*innen geändert wurden. Das System wurde von verschiedenen Organisationen wegen Diskriminierung, Verletzung der Privatsphäre, fehlender Transparenz und der Gefahr von Stigmatisierung kritisiert.

6.3.3 SCHWEDEN

Das Schwedische Arbeitsmarktservice Arbetsförmedlingen hat 2012 das Assessment Support Tool (AST) eingeführt (Assadi & Lundin 2018). Wenn

sich Erwerbsarbeitslose beim Arbeitsförmedlingen melden, werden im Zuge eines Interviews Daten erhoben, die ein erhöhtes Risiko von Langzeitarbeitslosigkeit indizieren sollen. Mit Hilfe dieser Daten sollen Erwerbsarbeitslose identifiziert werden, die Gefahr laufen, länger als sechs Monate arbeitslos zu bleiben. Dazu wird logistische Regression eingesetzt (Desiere et al. 2019) und das Risiko auf einer vierteiligen Skala angezeigt (Assadi & Lundin 2018).

Die Berater*innen müssen der Einschätzung des Systems nicht folgen, sie werden sogar ermutigt, ihre professionelle Erfahrung in die Entscheidungsfindung einfließen zu lassen (Assadi & Lundin, 2018). Trotz dieser Freiwilligkeit müssen Berater*innen begründen, wenn sie von der algorithmischen Einschätzung abweichen. Assadi & Lundin (2018) erhoben in einer Studie, dass es 82 % der lokalen Arbeitsförmedlingen-Manager*innen wichtig war, dass Berater*innen den Empfehlungen folgen. Weiters fanden sie heraus, dass Berater*innen mit zunehmender Erfahrung häufiger von der algorithmischen Einschätzung abweichen. Das trifft für Männer noch deutlicher zu als für Frauen. Unerfahrene Berater*innen orientierten sich hingegen stärker am algorithmischen System. Im Durchschnitt, so das Arbeitsförmedlingen, ist das Assessment Support Tool genauer als Berater*innen – wobei Assadi & Lundin (2018) betonen, dass unklar ist, ob das etwa auch für erfahrene Berater*innen der Fall ist.

Im Gegensatz zum System des österreichischen AMS werden Erwerbsarbeitslose mit hohem Risiko erweiterte Ressourcen zur Verfügung gestellt. Außerdem werden ihnen diese Ressourcen früher zugänglich gemacht als Erwerbsarbeitslosen mit geringem Risiko. Zuvor wurde in Schweden ein regelbasiertes System eingesetzt: Mit steigender Dauer der Erwerbsarbeitslosigkeit stieg auch der Zugang zu Ressourcen in Form von Intensivberatung und Kursen. Im Gegensatz dazu hat das neue System das erklärte Ziel, ab dem ersten Tag Personen mit erhöhtem Risiko von Langzeitarbeitslosigkeit intensive Ressourcen zur Verfügung zu stellen (Assadi & Lundin 2018). Ähnlich wie in Österreich war die Motivation auch in Schweden, begrenzte Ressourcen effektiver zu verteilen, allerdings indem Langzeitarbeitslosigkeit möglichst früh bekämpft wird.

6.3.4 SCHWEIZ

In einigen Ländern scheiterten statistische Systeme am Widerstand bzw. an fehlender Akzeptanz durch Mitarbeiter*innen. Ein gut dokumentierter Fall ist das Schweizer Pilotprojekt Statistisch assistierte Programmselektion (SAPS). Neben SAPS in der Schweiz scheiterten ähnliche Systeme auch in Südkorea und Finnland an fehlender Akzeptanz (Johnson et al. 2018). Bei dem Schweizer System handelt es sich jedoch nicht um ein Profiling-System wie AMAS, sondern um ein statistisches Targeting-System. Während Profiling in der Regel die Wahrscheinlichkeit von Langzeitarbeitslosigkeit schätzt und davon abhängig die Intensität von arbeitsmarktpolitischen Maßnahmen bestimmen soll, setzt Targeting darauf, Maßnahmen zu identifizieren, die für eine Person die beste Wirksamkeit verspricht (Behncke,

Frölich & Lechner 2007). Je nach Person können Maßnahmen unterschiedlich effektiv (oder sogar kontraproduktiv) sein. Pauschale Aussagen, dass Maßnahmen sinnvoll oder sinnlos sind, können daher nicht getätigt werden. Profiling errechnet jedoch nur einen pauschalen Chancenwert pro Person, ohne die unterschiedliche Wirksamkeiten zu berücksichtigen. Das führt, so die Studienautor*innen, zu einer geringen Aussagekraft von Profiling.

SAPS wurde im Jahr 2005 in einem randomisierten Feldexperiment in 21 regionalen Arbeitsvermittlungszentren erprobt. Als Webinterface sollte es die arbeitsmarktliche Maßnahmenverteilung unterstützen. Die abschließende Entscheidung verblieb bei den Berater*innen. Diese wurden vor dem Einsatz entsprechend geschult. Im Anschluss an die Probephase wurden die Beschäftigungsverläufe der Erwerbsarbeitslosen ein Jahr lang beobachtet. Dabei konnte kein Einfluss auf die Maßnahmenauswahl der Berater*innen oder eine Steigerung der Arbeitsmarktchancen von Teilnehmer*innen ermittelt werden.

In einigen Interviews konnten die Studienautor*innen mehrere Gründe für die fehlende Akzeptanz feststellen. Erstens fürchteten Berater*innen, dass ihre Freiheiten eingeschränkt werden. Weiteres wurde der Zeitaufwand bei einer Teilnahme als zu hoch eingestuft. Und drittens war die Präzision enttäuschend. Berater*innen nannten einige Beispiele in denen das System in ihrem Verständnis komplett unsinnige Empfehlungen ausgab. Trotz der Einbeziehung vieler Charakteristika lieferte SAPS nur durchschnittliche Einschätzungen für Untergruppen und entsprechend waren Ungereimtheiten zu erwarten. Das wurde Berater*innen zu Projektbeginn kommuniziert und sie wurden gebeten bei eindeutig falscher Klassifikation SAPS zu widersprechen.

In einigen Rückmeldungen wurde auch die Qualität der zugrundeliegenden Daten angezweifelt. Besorgnis wurde etwa über die Zuverlässigkeit aufgrund fehlender Details durch zeitliche Engpässe, Aktualität der Datenbank (z. B. Fehlen kurzfristiger Maßnahmen/Ereignisse) und die Subjektivität von Variablen wie etwa der Einschätzungen der Sprachkenntnisse geäußert. Die Prognosen von SAPS basieren ähnlich wie beim AMS-Algorithmus auf einer großen Bandbreite an Variablen, wie Alter, Geschlecht, Muttersprache, Nationalität, regionale Faktoren und Erwerbsgeschichte.

Abschließend ist noch zu erwähnen, dass die Dokumentation des Pilotprojekts im Abschlussbericht mit 180 Seiten sehr umfassend ist. Der Detailgrad dieser Dokumentation ist im Vergleich zur Dokumentation durch die Synthesis Forschung beachtlich. Die letzten uns bekannten Versuche zu statistischem Profiling sind in der Schweiz 2015 durchgeführt worden, haben aber nicht zu einer landesweiten Umsetzung geführt (Desiere et al. 2019).

6.3.5 DEUTSCHLAND

Ein kürzlich veröffentlichter Bericht der NGO Algorithm Watch (Matz et al. 2019) listet vier Systeme der Bundesagentur für Arbeit bzw. der Job-

center die automatisierte Entscheidungshilfen oder Entscheidungssysteme darstellen: (1) DELTA-NT ist eine computergestützte psychologische Begutachtung im Rahmen der beruflichen Orientierung. (2) VERBIS ist ein zentrales Informationssystem der Arbeitsagentur und erlaubt zum Beispiel das Abgleichen von Arbeitssuchenden mit offenen Stellen. (3) Unter dem Namen „3A1“ firmiert die „Automatisierte Antragsbearbeitung Arbeitslosengeld“. Dieses System soll eine automatisierte „Entscheidungsvorbereitung“ bieten und ist aktuell noch in Entwicklung, ein Prototyp wurde bereits getestet (Bundesministerium für Arbeit und Soziales, 2018). (4) Am ähnlichsten zum österreichischen System ist jedoch die Berechnungshilfe Arbeitsmarkt Chancen der „PP-Tools“. Ein paar Details bietet eine Beantwortung einer Anfrage der Bundestagsfraktion Die Linke aus dem Jahr 2018. Zum Zeitpunkt der Anfrage war dieses Tool bereits produktiv im Einsatz. Laut der Anfragebeantwortung werden in diesem System Wahrscheinlichkeiten der „Vermittlungsfähigkeiten des Kunden pro Zielregion“ errechnet. Dazu werden pro Dienststelle (und damit pro Zielregion) Arbeitshilfen, Stellenangebote und Stellengesuche der letzten 24 Monate herangezogen. Welche Daten im Detail erhoben werden ist dabei unklar, besonders, ob auch noch weitere Informationen als die Berufskennzahl (BKZ) über die Arbeitssuchenden in die Berechnung einfließen. Interessant ist eine darauf aufbauende Funktion: Die sogenannte „Umkreiswahrscheinlichkeit“ berechnet, „wie weit der Kunde umziehen/pendeln müsste, um mit einer bestimmten Wahrscheinlichkeit vermittelt werden zu können.“ (Bundesministerium für Arbeit und Soziales 2018, S. 46)

Abgesehen von diesen konkreten algorithmischen Verfahren kommt in Deutschland das sogenannte 4-Phasen-Modell zur Anwendung. Im Rahmen dieses Phasenmodells wurde lange Zeit Profiling der Erwerbsarbeitslosen durchgeführt, auf dessen Basis sie einer von sechs „Profillagen“ zugeteilt wurden: Die Profillagen bilden einerseits die prognostizierte Dauer der Arbeitsmarktintegration ab, andererseits aber auch, welcher Handlungsbedarf besteht. Eine Profillage war z. B. das Aktivierungsprofil, wo besonders im Bereich „Motivation“ Handlungsbedarf festgestellt wurde, oder das „Stabilisierungsprofil“, das ein Defizit bei der „Leistungsfähigkeit“ impliziert (Obermeier & Oschmiansky 2014). Als diese Profillagen verwendet wurden, kam es auch zu einem kleineren Skandal. 2013 wurde die deutsche Bundesagentur für Arbeit für Zielsetzungen kritisiert, die in der Praxis zu sogenanntem „Creaming“ geführt hatten (Dahlkamp, Dettmer & Tietz 2013). Da jede Vermittlung von Arbeitssuchenden gleich viel ‚zählte‘ wurde überwiegend auf Arbeitssuchende fokussiert, die leicht vermittelbar waren. Der deutsche Rechnungshof hatte diese Praxis scharf kritisiert: „Der Bundesrechnungshof hält es nicht für sachgerecht, dass die Bundesagentur ihre Ressourcen überwiegend für ‚gute Risiken‘ einsetzt.“ (zit. nach Dahlkamp, Dettmer & Tietz 2013, S. 33) Der Ansatz des österreichischen AMS verunmöglicht ein „Creaming“, da die aussichtsreichsten Erwerbsarbeitslosen keine Unterstützung bekommen sollen, die Verknappung der Ressourcen für jene mit schlechten Chancen ist aber ähnlich.

Im Jahr 2016 wurden diese Profillagen von der Bundesagentur für Arbeit abgeschafft (Bundesagentur für Arbeit 2016). Mit dem Ziel, die Komplexität des Beratungs- und Vermittlungsprozesses zu reduzieren und Freiräume für „komplexe Kundenanliegen“ zu schaffen wurden die Profillagen durch eine „Integrationsprognose“ ersetzt. Diese basiert auf der Einschätzung der Sachbearbeiter*innen und unterscheidet nur zwischen „marktnah“ (wahrscheinlich binnen 6 Monaten in einer neuen Beschäftigung) und „nicht marktnah“ (wenn eine Beschäftigung binnen 6 Monaten unwahrscheinlich erscheint). Personen, bei denen auch eine Arbeitsmarktintegration in 12 Monaten unwahrscheinlich erscheint, können Zugang zur internen ganzheitlichen Integrationsberatung („Inga“) bekommen. Dieses Programm zeichnet sich durch engere Kontaktmöglichkeiten und breite individuelle Leistungen und Unterstützungen aus (Bundesagentur für Arbeit 2020).

6.3.6 DÄNEMARK

In Dänemark wurde statistisches Profiling anfangs unter dem Namen Job Barometer eingesetzt, um die Chance einer Anstellung in den nächsten 6 Monaten zu errechnen (Loxha et al. 2014). Dabei wurden Betroffene in eine von fünf Kategorien je nach Chancenwert eingeteilt. Dieses System wurde jedoch eingestellt, unter anderem wegen dem hohen Dokumentationsaufwand, Verunsicherung und Mehraufwand aufgrund von notwendigen Mehrfachregistrierungen in IT-Systemen, dem Risiko von Diskriminierung auf Basis von Wohnorten und dem mangelhaften Überblick über den Arbeitsmarkt, welches das System bot. Die Abkehr passierte zu einer Zeit als Dienststellen mehr Autonomie bekamen (Loxha et al. 2014).

Das statistische Profiling wurde anschließend durch ein von Berater*innen durchgeführtes Profiling ersetzt (Loxha et al. 2014). Hierbei entscheiden Berater*innen auf Basis der Förderungsgeschichte sowie zwei Fragen an den*die Kunden*in: (1) ob ein Job angenommen werden kann und (2) ob eine Kursteilnahme möglich ist. Anhand dieser Information werden Erwerbsarbeitslose in drei Kategorien (National Labour Market Authority 2011) eingeteilt, die denen in AMAS ähneln:

- *Gruppe 1:* Bereit einen Job anzunehmen in 3 Monaten
- *Gruppe 2:* Bereit für Unterstützungsmaßnahmen und wird in 3 Monaten voraussichtlich keine Anstellung finden
- *Gruppe 3:* Temporäre passive Unterstützung, erwerbsarbeitslose Person steht vor so großen Problemen bzw. Herausforderungen, dass eine Vermittlung oder Teilnahme an Angeboten nicht möglich ist

Uns ist nicht bekannt, ob dieses Berater*innen-basierte Profiling in der beschriebenen Form aktuell noch landesweite Praxis ist. Seit der genannten Umstellung wurde auch ein statistisches Profiling für die Priorisierung von jungen Erwerbsarbeitslosen mit einem hohen Risiko auf Langzeitarbeitslosigkeit in der Beratung eingesetzt (Pultz 2016). Laut einer kürzlich veröffentlichten OECD Studie (Desiere et al. 2019) ist aktuell in Dänemark nur ein statistisches Profiling-System im Einsatz. Dieses wird verwendet, um die Wahrscheinlichkeit einer Langzeitarbeitslosigkeit (> 26 Wochen) zu bestimmen, um eine Priorisierung bei der Beratung zu ermöglichen. Die Verwendung dieses Systems ist freiwillig sowohl für Berater*innen also auch Erwerbsarbeitslose und damit auch nicht direkt mit Konsequenzen für den Zugang zu Unterstützungsmaßnahmen verbunden. Die errechneten Ergebnisse sind für Erwerbsarbeitslose einsehbar. Limitationen des Systems sind etwa, dass lokale Arbeitsmarktfaktoren nicht berücksichtigt werden und dass die Trefferquote nur bei etwas über 60 % liegt.

6.4 TRENDS UND PROBLEME

Die Menge an unterschiedlichen Methoden und Daten, welche für Profiling-Systeme verwendet werden, machen einen direkten Vergleich schwierig. In Anbetracht der berücksichtigten Variablen und auch der Prognose von Langzeitarbeitslosigkeit gibt es einige vergleichbare Systeme. Die gewählte Methodik von AMAS anhand von Konstellationen Integrationschancenwerte zu berechnen, stellt eine Ausnahme dar. Ebenso die Art der Verknüpfung des AMS-Algorithmus mit der Ressourcenverteilung unter Erwerbsarbeitslosen, eine Fokussierung auf die Gruppe mit „mittleren Chancen“ am Arbeitsmarkt, bildet, bis auf Polen, eine Ausnahme. In Polen gab es breite öffentliche Kritik, zur Abschaffung des Systems führte jedoch ein Formalfehler. In Deutschland war die Bevorzugung von aussichtsreichen Erwerbsarbeitslosen ein Nebeneffekt der Zielvorgaben für Jobcenter, was ebenso zu Kritik und letztendlich zu Änderungen führte.

Digitalisierung, auch in der Verwaltung, ist derzeit in aller Munde. Dennoch ist hervorzuheben, dass es keinen klaren Trend zu statistischem Profiling gibt: In manchen Ländern wird es schon länger eingesetzt, in anderen Ländern wurde es aber auch wieder abgeschafft (Johnson et al. 2018). Häufige Gründe dafür sind zu geringe Genauigkeit des Profiling (z. B. Kanada) und Widerstand bzw. fehlende Akzeptanz von Mitarbeiter*innen (z. B. Schweiz, Finnland, Südkorea). In Polen zeigte sich, dass Berater*innen selten von algorithmischen Einschätzungen abweichen bzw. diesen widersprechen. Ein differenziertes Bild diesbezüglich gibt es aus Schweden, wo erfahrene Berater*innen häufiger abweichen als unerfahrene (Asadi & Lundin 2018).

Alternative Ansätze im Profiling zeigen die illustrierten Systeme in Deutschland, Schweden und den Niederlanden, welche auf Freiwilligkeit setzen und verstärkt besonders Benachteiligte unterstützen. Es gibt jedoch keine stichhaltigen Studien, welche die Effektivität von statistischem Profiling bestätigen. Eine EU-Expert*innen-Gruppe (Johnson et al. 2018) empfahl etwa in Personal statt statistisches Profiling zu investieren, da potenzielle Steigerung von Effektivität und Effizienz nicht die Nachteile und Risiken aufwägen.

7 SCHLUSSFOLGERUNGEN UND EMPFEHLUNGEN

7.1 ZIELHETEROGENITÄT UND UNGLEICHHEITSVERHÄLTNISSE

Zahlreiche AMS-Dokumente kommunizieren die Vorstellung, dass nach Einführung von AMAS in den Kernprozess „Arbeitskräfte unterstützen“ der Integrationschancenwert (IC-Wert) lediglich eine Zusatzfunktion darstellt, die den Berater*innen eine evidenzbasierte Entscheidungshilfe zur Verfügung stellt. Die vorliegende soziotechnische Analyse zeigt allerdings eine Bandbreite an Konsequenzen, die sowohl die operative Praxis der AMS-Berater*innen, das organisationale Selbstverständnis des AMS als semi-staatliches Dienstleistungsunternehmen, arbeitsmarktpolitische Weichenstellungen³¹ sowie rechtspolitische Fragen des Einsatzes algorithmischer Systeme im öffentlichen Sektor betreffen. Betrachten wir die einleitend dargestellten heterogenen Zielsetzungen des Projekts AMAS, so bilden die Ergebnisse der Studie folgende Dynamiken ab:

1. Eine tatsächliche Steigerung der Effizienz des Beratungsprozesses geht möglicherweise nur mit einer vorwiegend routinemäßigen Übernahme der computergenerierten Einstufung einher. Der computerunterstützte Prozess rückt, einer Effizienzlogik folgend, nicht das differenzierte Angebot und den Bedarf der Kund*innen in den Fokus, sondern bringt eine Bewertung von Integrationschancen anhand von Personenmerkmalen mit der Vergabe von Fördermitteln und differenzierten Kontrollmechanismen in Verbindung. Dies mag im Sinne einer Dienstleistungs-rationalisierung nachvollziehbar sein, im Sinne der Serviceorientierung einer semi-staatlichen Institution, die ihrem öffentlichen Auftrag zur Erhöhung der Chancen von Arbeitssuchenden nachkommt, aber am Ziel vorbei gehen.
2. Die Erhöhung der „Schulungseffektivität“ durch Konzentration der Fördermittel im „mittleren“ Segment zielt nicht auf die Zielgenauigkeit und Qualität der Maßnahme für eine*n AMS-Kunden*in ab, sondern verbindet das Ziel des „effektiven“ Mitteleinsatzes mit einer groben – und eben nicht zielgenauen – Einteilung in drei Gruppen. Als Indikatoren werden aktuell der Erfolg der Arbeitsmarktförderung und die Reduzierung der Geschäftsfalldauer im „hohen“ und „mittleren“ Segment definiert. Allerdings wird der Erfolg der Arbeitsmarktförderung gemessen an der Reduzierung der Geschäftsfalldauer wieder als Teil der Personenmerkmale und nicht als qualitatives Merkmal des Maßnahmen-

Konsequenzen

Serviceorientierung
leidet

Maßnahmenevaluation
ist ausständig

³¹ Siehe dazu Allhutter et al. (2020).

<p>Anti-Diskriminierungsstrategien sind mangelhaft</p>	<p>angebots in das System rückgeführt. Ob eine Maßnahme gegriffen hat und Mittel daher effektiv eingesetzt wurden, oder ob andere Gründe für einen schnellen Wiedereinstieg ins Arbeitsleben ausschlaggebend waren, bleibt daher offen; klare Kausalzusammenhänge lassen sich aus den IC-Werten keinesfalls ablesen.</p> <p>3. Ein weniger zentrales, aber doch immer wieder genanntes Ziel war es, einer etwaigen Willkür oder Stereotypisierung durch Berater*innen durch Standardisierung der Beratungsleistung beizukommen. Wie die vorliegende Analyse gezeigt hat, wurden in der Entwicklung von AMAS kaum Verfahren eingesetzt, um Bias im System zu vermeiden; auch in der Anwendung bietet das System den Berater*innen keine Anhaltspunkte um möglichen strukturellen Ungleichbehandlungen vorzubeugen. Während das AMS im Kernprozess „Arbeitskräfte unterstützen“ immer wieder auf Gleichstellungsaspekte – etwa im Jobmatching oder in der Betreuungsvereinbarung – eingeht, kann eine objektivierte, numerische (und nicht hinsichtlich Bias optimierte) Einstufung der Integrationschancen eine Diskrepanz zum gesetzlichen Auftrag der Gleichbehandlung darstellen.</p>
<p>Repräsentation von Arbeitssuchenden und segmentiertes Angebot</p>	<p>Die verfolgten Ziele können auch nochmals hinsichtlich der unterschiedlichen Wirkweisen des Systems beleuchtet werden. Das System AMAS selbst bildet mittels Profiling eine <i>Repräsentation</i> von Arbeitssuchenden als „Service-, Beratungs- und Betreuungskund*innen“, die mit einem numerischen IC-Wert und einer Klassifizierung als „hohes“, „mittleres“ und „niedriges“ Segment versehen sind. Unabhängig von den Maßnahmen, die an diese Segmentierung geknüpft sind, ist diese Kategorisierung nicht wertneutral, sondern stellt eine Hierarchisierung der zu Vermittelnden dar, die auf Individuen und Gesellschaft wirkt. Wenn sich im „niedrigen“ Segment zu Beginn des Geschäftsfalls schließlich Frauen mit „Migrationshintergrund“ zu 16 % im Vergleich zu Männern mit vollständiger Erwerbsgeschichte zu 3 % wiederfinden, überschneidet sich die AMAS-Kategorisierung außerdem mit gesellschaftlichen Ungleichheitslagen. Damit verbunden sind die <i>Entscheidungen</i>, die an die Zuordnung zu einem Segment gebunden sind. Ein segmentiertes Angebot, das unterschiedliche Bedarfe abdeckt, kann eine angemessene Differenzierung von Betreuungsformaten und Kontakthäufigkeit mit Betreuer*innen darstellen. Wenn allerdings bestimmte Personengruppen gehäuft einem Segment zugeordnet werden, weil die Zugehörigkeit zu einer Konstellation stark auf Personencharakteristiken abstellt und andere Einflussfaktoren in geringerem Ausmaß in die Integrationschance einfließen, handelt es sich um eine systematische Ungleichbehandlung.³²</p>
<p>systematische Ungleichbehandlung</p>	<p>Eine Tendenz in diese Richtung zeigt sich in der Rekonstruktion des technischen Systems durch kumulative Effekte und Feedback-Loops. Eine weitere systematische Ungleichbehandlung liegt in der Diskrepanz der Fehler-</p>

³² In der Fachliteratur werden diese beiden Wirkweisen als „representational harm“ und „allocative harm“ bezeichnet (Whittaker et al. 2018, S. 25).

raten für unterschiedliche Personengruppen, so dass eine Falscheinstufung bestimmte Personengruppen in größerem Ausmaß betrifft als andere. Auf der operativen Ebene ist es klar diskriminierend, dass Populationen mit unvollständigen Daten keine Erklärungstexte für ihre Einstufung bekommen und auch in den vorhandenen Erklärungstexten finden sich Schieflagen in der Bewertung von Qualifikationen.

Diese Identifikation von Indizien für in das System eingeschriebenen Bias und an die Kategorisierung gebundene ungleiche Entscheidungen, können allerdings nicht eine Evaluierung von AMAS auf Basis der konkreten Personendaten ersetzen. Eine Körperschaft öffentlichen Rechts unterliegt aus grundrechtlicher, demokratischer und rechtsstaatlicher Sicht bestimmten Transparenz-, Berichts- und Evaluierungspflichten, die hier unbedingt eingefordert werden müssen.

7.2 EMPFEHLUNGEN

Die vorliegende Studie fokussiert auf AMAS, jedoch sind viele der Kritikpunkte und Herausforderungen, die dieses konkrete Projekt betreffen, auch typisch für andere (semi-)automatische Systeme in (teil-)öffentlichen Institutionen. Ein Handlungsbedarf ergibt sich daher auf mehreren Ebenen und wird im folgenden Abschnitt in Form von Empfehlungen für unterschiedliche Domänen formuliert: 1) Der erste Block an Empfehlungen bezieht sich konkret auf AMAS und seine Einbettung in die soziale Praxis. 2) Darauf aufbauend bietet der zweite Block Empfehlungen zur Einführung algorithmischer Systeme im öffentlichen Sektor. 3) Schließlich diskutieren wir, welche rechtlichen und politischen Rahmenbedingungen und Governance-Strukturen notwendig sind, um (semi-)automatisierte Entscheidungssysteme in unterschiedlichen Sektoren gerecht zu gestalten. Dabei beziehen wir uns auf nationale und internationale Initiativen, die sich bereits ausführlich mit ethischen Aspekten in KI-Anwendungen³³ und automatisierten Systemen beschäftigt haben. Die folgenden Empfehlungen betten sich also in einen breiteren Diskurs über Ethik in KI-Systemen und (semi-)automatische Systemen ein, den es in Zukunft weiter zu verfolgen gilt. Die vorliegende Analyse des „AMS-Algorithmus“ kann dabei als empirisches Fallbeispiel für abstrahierte Ethik-Diskussionen dienen, das größere Problemfelder und gesellschaftliche Implikationen detailliert veranschaulicht und greifbar macht.

AMAS als Beispiel
für Ethik-Debatten

³³ Anwendungen unter Einbeziehung künstlicher Intelligenz

7.2.1 EMPFEHLUNGEN ZU AMAS

Kritik und Unverständnis	<p>Die schrittweise Einführung von AMAS war von öffentlicher Kritik begleitet, die sich auf das System selbst, so wie es in den Medien kommuniziert wurde, als auch auf die fehlende Nachvollziehbarkeit und Transparenz bezog. In der Phase des Testbetriebs wurden Teile der Variablen und Modelle veröffentlicht und in den Medien diskutiert. Diese bruchstückhafte Dokumentation hat jedoch mehr Verwirrung gestiftet als Aufklärung geschaffen, wie sich an der nachfolgenden Debatte um die logistische Regression gezeigt hat, die de facto keine wesentliche Rolle in der Berechnung des IC-Werts spielt, wie sich im Rahmen unserer Datenanfrage herauskristallisierte. Auch in der Serie an bereits davor gestellten und vom AMS beantworteten Datenanfragen (vom Rechnungshof, der Volksanwaltschaft, der Anwaltschaft für Gleichbehandlungsfragen, sowie der NGO „Epicenter.Works“), wurde dieses wesentliche Detail nicht aufgeklärt. Eine vollständige Information über die sozio-technischen Funktions- und Wirkweisen von AMAS sind bis dato immer noch nicht vorhanden, sodass eine umfassende Einschätzung des Systems unmöglich ist und öffentliche Debatten nach wie vor von Unsicherheiten und Unverständnis geprägt sind (insbesondere da viele der uns zur Verfügung gestellten Materialien nicht öffentlich zugänglich gemacht wurden).</p>
Transparenz	<ul style="list-style-type: none"> • Systemtransparenz: Das Gewährleisten von Systemtransparenz während des laufenden Entwicklungsprozesses, und insbesondere vor dem Letztentscheid zum Einsatz des Systems, soll solchen grundrechtlich problematischen Vorgehensweisen in Zukunft vorbeugen. Gerade bei komplexen Systemen dieser Art bietet es sich an, nicht nur auf schriftliche Dokumentation zu setzen, sondern durch klar strukturierte Illustrationen oder Beispielberechnungen, sowie durch anonymisierte Open Data-Schnittstellen Systemtransparenz zu schaffen. Diese Transparenz sollte für ein semi-staatliches Dienstleistungsunternehmen nicht als optionales Qualitätsmerkmal eines solchen Systems, sondern als Pflicht verstanden und rechtlich verankert werden. • Datentransparenz: Als notwendiger Gegenpol zur Systemtransparenz ist eine umfassende Datentransparenz unumgänglich, um eine belastbare Evaluation des Systems in Bezug auf Bias und Diskriminierung zur ermöglichen. Auch wenn im Sinne der DSGVO jene Daten, die real im System zum Einsatz kommen, als schützenswert zu behandeln sind, ist eine anonymisierte Publikation der Datensätze möglich. Ohne Zugang zu diesen Daten sind auch viele der seitens des AMS vorgebrachten Argumente für die Unbedenklichkeit des Systems nicht verifizierbar und keinesfalls als Erfüllung des vorgebrachten Bekenntnisses zu umfassender Transparenz zu sehen.
nachvollziehbare Einstufung	<ul style="list-style-type: none"> • Nachvollziehbarkeit: Im Bereich der Rechenschaftspflicht – „<i>Accountability</i>“ – muss sowohl aus rechtlicher Sicht (wie von der Datenschutzbehörde bemängelt [DSB 2020]), als auch aus Sicht des operativen Einsatzes eine klare Nachvollziehbarkeit gegeben sein. Hierbei müssen ausnahmslos alle Betroffenen abgedeckt werden; Erklärungsmodelle

wie die Segmentzusatzinformationen, die nur für Teile der Betroffenen verfügbar sind, sind hier keinesfalls ausreichend.

- **Bedarfserhebung und Umsetzung:** Zentrale Zielvorgaben von AMAS sind, wie oben diskutiert, Effizienz, Effektivität und Vermeidung von Willkür. Alle drei Ziele sind aus Sicht der Gesamtorganisation formuliert und es fehlen Indizien, dass andere Stakeholder dieselben Interessen teilen würden oder ausschlaggebend für das Systemdesign waren. Im Sinne eines *human-centered design* empfiehlt sich eine Bedarfserhebung durchzuführen, welche User*innen und andere Stakeholder der Organisation in den Entwicklungsprozess miteinbezieht und welche deren Interessen in der Umsetzung berücksichtigt. Soll das System darüber hinaus auch die Zielvorgabe der besseren, „personalisierteren“ Betreuung von Kund*innen verfolgen, muss die Bedarfserhebung auch auf externe Gruppen, insbesondere Arbeitssuchende oder Arbeitslosen-Vereinigungen, ausgeweitet werden, um auch deren Interessen und Bedürfnisse einbeziehen zu können. Nur so kann ein Kund*innen-zentriertes System entwickelt werden. In öffentlichen Bereichen, wie der Arbeitsvermittlung, bedürfen diese Fragen einer gesellschaftlichen Mitsprache und Diskussion und sollten nicht hinter Stellvertreterdiskussionen über Algorithmen und deren Objektivität verschwinden.
- **Einsicht und Einspruchsrecht:** Eng im Zusammenhang mit Transparenz und demokratischer Teilhabe und Kontrolle steht die Notwendigkeit, Betroffenen Einsicht in die Berechnungen/Einflussfaktoren ihrer eigenen Einstufung und gegebenenfalls ein Einspruchsrecht bei Falschklassifikation zu gewähren (auch dieser Punkt wurde von der Datenschutzbehörde [DSB 2020] bemängelt). Dazu sollte eine möglichst niederschwellige Ombudsstelle eingerichtet werden, die Betroffene in strittigen Fällen aufsuchen können, um Hilfe und Rechtsbeistand zu erlangen.³⁴ Unsere Analyse der Umgestaltung der sozialen Praxis durch AMAS zeigt sehr deutlich, dass Betroffene mitunter Unterstützung brauchen sich gegen ihre (falsche) Einstufung auszusprechen. Unter den gegebenen Umständen der aktuellen Beratungspraxis, insbesondere bei Kund*innen in schwierigen Lebenssituationen oder mit mangelnden Deutschkenntnissen, muss sichergestellt sein, dass diese Ombudsstelle auch stichprobenartig in Eigeninitiative tätig wird, da eben jenen Kund*innen nicht zumutbar ist, selbst aktiv gegen etwaige (falsche) Einstufungen vorzugehen.
- **Neue Kompetenzen:** Algorithmen können nur eine aggregierte Sichtweise auf Gruppen errechnen, die nicht einfach auf individuelle Personen umgelegt werden kann. Der tägliche Umgang mit Korrelationen auf Basis von ausgewählten Populationen und deren Bedeutung für den Einzelfall einer arbeitssuchenden Person erfordert neue Kompetenzen

human-centered
design

rechtlicher Beistand

kritische
Wissensvermittlung

³⁴ Zwar existiert bereits eine Ombudsstelle, die aber keine auf AMAS spezialisierte Rechtshilfe leistet oder als niederschwellig betrachtet werden kann (siehe: https://www.oesterreich.gv.at/themen/leben_in_oesterreich/ombudsstellen_und_anwaltschaften/Seite.3240001.html).

und Fähigkeiten auf Seite der AMS-Berater*innen, die über Handlungsanleitungen zur Gesprächsführung hinausgehen müssen. Die Ausbildung von AMS-Mitarbeiter*innen sollte ein kritisches Wissen über die Quantifizierung der arbeitssuchenden Bevölkerung und deren Chancenberechnungen vermitteln, um damit einhergehende Bias und Fehlklassifikationen adäquat verstehen und vermitteln zu können. Die vom AMS geplanten, kurzen Schulungen sind dafür keineswegs als ausreichend zu bewerten. Gleichzeitig betrifft diese Problematik Kund*innen in ebensolcher Weise. Soll der IC-Wert, wie derzeit beschrieben, mit Kund*innen besprochen werden und auch für diese als Orientierungshilfe nützlich sein, müssen Kund*innen ebenso neue Kompetenzen entwickeln. Ob und wie dies im konkreten Fall möglich sein könnte ist unklar; um sicherzustellen, dass die angezeigten IC-Werte nicht zu Fehlschlüssen seitens der Kund*innen über ihre tatsächliche Situation führen, wäre dies jedoch wünschenswert.

nicht-technische
Lösungen

- **Besserer Betreuungsschlüssel:** Schließlich veranschaulicht die vorliegende Studie einen weiteren Kritikpunkt der Datenschutzbehörde [DSB 2020], nämlich jenen der „routinemäßigen Übernahme“, sehr deutlich. Sie zeigt, dass AMAS – und seine Einbettung in die AMS Praxis – Anreize setzt, der computergenerierten Einstufung eher zu folgen, als diese zu „korrigieren“. Fehlende Zeit und Legitimation sind zwei zentrale Gründe dafür, weswegen nicht-technische Lösungen – insbesondere mehr Budget für einen besseren Betreuungsschlüssel – nicht aus den Augen verloren werden dürfen. Die Anweisung an Betreuer*innen als menschliches „Korrektiv“ zu fungieren ist angesichts der knappen Betreuungszeiten am AMS nicht ausreichend, um diese Dynamiken oder Automatismen abzufangen.

7.2.2 EMPFEHLUNGEN ZUM EINSATZ ALGORITHMISCHER SYSTEME IM ÖFFENTLICHEN SEKTOR

Fairness und
Gerechtigkeit

- **Anti-Diskriminierungs-Strategien:** Internationale Studien (Dobbe et al. 2018) zeigen, dass ein biasfreies System unmöglich zu realisieren ist, da jede Modellierung eine Reduktion der komplexen Realität erfordert. Dennoch müssen Schritte entsprechend dem aktuellen wissenschaftlichen Stand zu algorithmischer Fairness und Anti-Diskriminierung durchgeführt und offengelegt werden. Einen Schritt in diese Richtung macht z. B. die Stadtverwaltung von Amsterdam. Dort wird aktuell ein Algorithmen-Register aufgebaut, wo neben detaillierter Auflistung der verwendeten Daten, Algorithmen und Systemarchitektur auch explizit angeführt wird, wie diskriminierende Effekte des Algorithmus verhindert werden, wie menschliche Aufsicht umgesetzt und wie potentiellen Risiken begegnet wird. Ein solches Vorgehen empfiehlt sich für alle algorithmischen Systeme im öffentlichen Bereich, die gesellschaftliche Interessen berühren.

gesellschaftliche
Teilhabe

- **Öffentliche Konsultationen:** Ein transparenter Entwicklungsprozess, sowie das Offenlegen von Anti-Diskriminierungsmaßnahmen ermög-

licht eine öffentliche Debatte und gesellschaftliche Teilhabe an der Einführung von algorithmischen Systemen im öffentlichen Dienst. Als probates Mittel dafür bieten sich öffentliche Konsultationen an, um Bedenken, Kritikpunkte und offene Fragen zu klären.

- **Informationspflicht:** In sensiblen Bereichen wie dem Einsatz von algorithmischen Systemen im öffentlichen Sektor fordert das Gutachten der deutschen Datenethikkommission (DEK) das „Recht einer betroffenen Person auf aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen eines algorithmischen Systems (vgl. DSGVO)“ (Datenethikkommission 2019, S. 27)³⁵. In Bereichen des öffentlichen Interesses sollte darüber hinaus nicht nur unmittelbar betroffenen Personen ein Recht auf Information eingeräumt werden, sondern auch einem erweiterten Personenkreis bestehend aus Forscher*innen, Journalist*innen oder (neu geschaffenen) Aufsichtsorganen. Darüber hinaus sollten diese Rechte nicht nur für automatisierte Entscheidungssysteme gelten, sondern auch für semi-automatisierte Systeme, die mit Profilbildungen arbeiten, unabhängig davon, ob Menschen die Letztentscheidung obliegt oder nicht. Angesichts der vorliegenden Studie – und insbesondere der darin beschriebenen Gefahr der „routinemäßigen“ Übernahme des IC-Werts in der Beratungspraxis – empfiehlt es sich, diese Forderungen nach Informationspflicht möglichst rasch umzusetzen, um einen sicheren und nachhaltigen Umgang mit (semi-)automatisierten Systemen in Bereichen des öffentlichen Dienstes zu gewährleisten.

Recht auf
aussagekräftige
Informationen

7.2.3 INSTITUTIONELLE EMPFEHLUNGEN

Um die oben formulierten Empfehlungen zu implementieren, sind bestehende Grundrechte umzusetzen und neue/adaptierte gesetzliche Grundlagen, Gremien und Aufsichtsorgane, sowie neue Auditing-Verfahren einzurichten, wo notwendig. Transparenz, demokratische Kontrolle und Anti-Diskriminierung sind Eckpfeiler der Rechtsstaatlichkeit. Die fortschreitende Digitalisierung kann neue Prozesse und Institutionen erfordern, die standardisierte Verfahren dafür entwickeln. Sowohl nationale, als auch internationale Initiativen im Bereich Ethik und KI, als auch Ethik in automatisierten Entscheidungssystemen geben diesbezüglich bereits wichtige Anhaltspunkte. Hier sind insbesondere österreichische KI Strategien wie die „Artificial Intelligence Mission Austria 2030“³⁶ oder das „White Paper des Österreichischen Rats für Robotik und Künstliche Intelligenz“³⁷ auf nationaler Ebene, sowie das „White Paper on Artificial Intelligence. A Eu-

bestehende
Grundrechte und neue
Verfahren

³⁵ <https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html>

³⁶ <https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/aimat.html>

³⁷ https://www.acrai.at/wp-content/uploads/2019/08/ACRAI_White_Paper_DE.pdf

ropean Approach to excellence and trust“³⁸ auf EU-Ebene zu nennen. Diese Initiativen beziehen sich primär auf automatisierte und autonome KI-Anwendungen, sind im Hinblick auf Transparenz, Stakeholder-Einbindung und Zertifizierung algorithmischer Systeme aber auch auf teilautomatisierte Systeme zu übertragen. Insbesondere möchte diese Studie aber Erkenntnisse der deutschen Datenethikkommission (DEK) und der österreichischen Datenschutzbehörde [DSB 2020] unterstützen, die sich nicht nur zu autonomen Systemen, sondern zur Governance von semi-automatisierten Entscheidungssystemen äußern – Zweitere explizit zu AMAS. Insbesondere folgende konkrete Empfehlungen können durch die detaillierte Analyse von AMAS empirisch untermauert werden: Die Ausweitung von gesetzlichen Rahmenbedingungen für automatisierte Entscheidungssysteme auf semi-automatisierte Systeme, die in der Praxis oft als automatisierte Entscheidungssysteme fungieren (können), ist eine zentrale Empfehlung der deutschen Datenethikkommission, die die vorliegende empirische Analyse nahelegt. Die klare Nachvollziehbarkeit des Systems für Betroffene, sowie für einen erweiterten Personenkreis werden sowohl von der Datenethikkommission, als auch von der Datenschutzbehörde gefordert. Im Sinne der Rechenschaftspflicht („*Accountability*“) unterstützen die Ergebnisse der vorliegenden Studie dieses Anliegen sowohl aus rechtlicher Sicht, als auch aus Sicht des operativen Einsatzes von AMAS. Die Schaffung einer rechtlichen Grundlage für Informationspflicht und Einspruchsrecht für Betroffene, geschützte Gruppen und ausgewählte Stakeholder-Gruppen ist daher eine weitere zentrale Forderung der österreichischen Datenschutzbehörde [DSB 2020], der sich die vorliegende Studie anschließt.

neue Gremien,
Aufsichtsbehörden und
Auditverfahren

Über diese konkreten gesetzlichen Rahmenbedingungen hinaus braucht es neue Gremien und Aufsichtsbehörden, die eine unabhängige Begleitung, Auditierung und Evaluierung algorithmischer Systeme und des sozio-technischen Kontexts ihres Einsatzes durchführen können (sowohl national, als auch auf EU-Ebene). Dabei sollte auf sozialwissenschaftliche, technische und rechtliche Expertise sowie auf die fachliche Expertise von zivilgesellschaftlichen Akteur*innen zurückgegriffen werden. Wissenschaft und Zivilgesellschaft sollten in den Aufbau von institutionalisierten Auditverfahren, -prozessen und -institutionen einbezogen werden. Oftmals rasch benötigte Technikbewertungen in sensiblen Bereichen wie Arbeit oder Gesundheit, wie zum Beispiel im Falle der Stopp-Corona-App, die ehrenamtlich von Grundrechts-NGOs und Wissenschaftler*innen rund um Epicenter.Works durchgeführt wurde³⁹, sollte nicht allein dem Idealismus engagierter Akteur*innen überlassen werden, sondern einem professionellen Prozedere folgen, das ausreichend mit finanziellen, technischen und rechtlichen Ressourcen ausgestattet ist. Neben punktuellen Audits, könnten durch solche professionalisierten Verfahren kontinuierliche Evaluierungen (semi-)automatisierter Entscheidungssysteme und deren Auswirkun-

³⁸ https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

³⁹ <https://epicenter.works/document/2497>

gen über einen längeren Zeitraum bewerkstelligt werden. Ebenso ist die Exploration alternativer Ansätze zur Problemlösung unabdingbar, um die grundlegende Notwendigkeit einer rein technischen Lösung zu hinterfragen und die Möglichkeit zu schaffen, jenseits einer technikdeterministischen Zieldefinition potentiell weniger risikoreiche Lösungen zu finden.

Auf Basis fundierter Analysen und Audits durch Expert*innen-Gremien könnte sich eine gesamtgesellschaftliche Debatte entfalten, die weniger von Unsicherheiten und Ängsten, als von technisch-wissenschaftlicher Expertise, Rationalität und einem Fokus auf Gerechtigkeit und Fairness geprägt ist. Das würde den „europäischen Weg“ fortsetzen (vgl. Mager 2017), „wonach sich europäische Technologien durch konsequente Ausrichtung an europäischen Werten und Grundrechten, wie sie insbesondere auch in der Charta der Grundrechte der Europäischen Union und in der Konvention zum Schutz der Menschenrechte und Grundfreiheiten des Europarats zum Ausdruck kommen, auszeichnen sollten“ (Datenethikkommission 2019, S.32)⁴⁰, wie es die Datenethikkommission abschließend auf den Punkt bringt.

**fundierte
gesellschaftliche
Debatte**

⁴⁰ <https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html>

LITERATUR

- AI Now Institute (Hg.), 2018, Litigating algorithms: Challenging government use of algorithmic decision systems, New York, <https://ainowinstitute.org/litigatingalgorithms.html>.
- Allhutter, D., Cech, F., Fischer, F., Grill, G. und Mager, A., 2020, Algorithmic profiling of job seekers in Austria: How austerity politics are made effective, *Frontiers In Big Data, Special Issue Critical Data and Algorithm Studies*, 17, <https://doi.org/10.3389/fdata.2020.00005>.
- Assadi, A. und Lundin, M., 2018, Street-level bureaucrats, rule-following and tenure: How assessment tools are used at the front line of the public sector, *Public Administration*, 96(1), 154–170.
- Barnes, S., Wright, S., Irving, P. und Deganis, I., 2015, Identification of latest trends and current developments in methods to profile jobseekers in European Public Employment Services [Final report], EUROPEAN COMMISSION Directorate-General for Employment, Social Affairs and Inclusion Directorate C – Europe 2020: Employment Policies Unit C.3 – Mobility and Employment Services European Commission B-1049 Brussels, <http://ec.europa.eu/social/BlobServlet?docId=14173&langId=en>.
- Barocas, S. und Selbst, A. D., 2016, Big data's disparate impact, *Cal. L. Rev.*, 104, 671.
- Behncke, S., Frölich, M. und Lechner, M., 2007, Abschlussbericht zum Pilotprojekt Statistisch Assistierte Programmselektion (SAPS), Schweizerisches Institut Für Aussenwirtschaft Und Angewandte Wirtschaftsforschung (SIAW), St Gallen.
- Bickel, P. J., Hammel, E. A. und O'Connell, J. W., 1975, Sex bias in graduate admissions: Data from Berkeley, *Science*, 187(4175), 398–404.
- Billmann-Mahecha, E. und Tiedemann, J., 2006, Übergangsempfehlung als kritisches Lebensereignis: Migration, Übergangsempfehlung und Fähigkeitsselbstkonzept, In: *Risikofaktoren kindlicher Entwicklung*, VS Verlag für Sozialwissenschaften, S. 193-207.
- Bowen, G. A., 2009, Document analysis as a qualitative research method, *Qualitative research journal* 9, 27–40.
- Bowker, G. C., 2005, *Memory practices in the sciences*, MIT Press Cambridge, MA.
- Bowker, G. C. und Star, S. L., 1999, *Sorting things out: Classification and its consequences*, MIT Press.
- Boyd, D. und Crawford, K., 2012, Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon, *Information, Communication & Society*, 15(5), 662–679.
- Bundesagentur für Arbeit, 2016, Weisung 201611035 vom 16.11.2016 – Fokussierung auf Handlungsstrategien – Wegfall von Profillagen, https://www.arbeitsagentur.de/datei/weisung201611035_ba014507.pdf.
- Bundesagentur für Arbeit, 2020, Auf neuen Wegen zum Erfolg. Interne ganzheitliche Integrationsberatung (Inga) – Karlsruhe-Rastatt, <https://www.arbeitsagentur.de/vor-ort/karlsruhe-rastatt/inga>.
- Bundesministerium für Arbeit und Soziales, 2018, Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Jessica Tatti, Susanne Ferschl, Doris Achelwilm, weiterer Abgeordneter und der Fraktion DIE LINKE.
- Böheim R., Eppel R., Mahringer H., Gatterbauer M., Wilk M., Fuchs S., Lorenz C. und Böhs, G., 2017, Die Auswirkungen einer verbesserten Betreuungsrelation für Arbeitslose in der Arbeitsvermittlung des AMS—Ergebnisse eines kontrollierten Experiments des AMS Österreich in der Beratungszone der RGS Esteplatz in Wien, Arbeitsmarktservice Österreich Bundesgeschäftsstelle Marlies Gatterbauer, Marius Wilk A-1200 Wien, Treustraße 35-43 Tel: (+43 1) 331 78-0, http://www.forschungsnetzwerk.at/downloadpub/2017_ams-personalpiilot_Wien_esteplatz-2017.pdf.

- Costanza-Chock, S., 2020, Design Justice: Community-Led Practices to Build the Worlds We Need. MIT Press.
- Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., McElroy, E., Sánchez, A., Raji, D., Rankin, J., Richardson, R., Schultz, J., West, S. und Whittaker, M., 2019, AI Now 2019 Report. AI Now Institute.
- Dahlkamp, J., Dettmer, M., und Tietz, J., 2013, Mit allen Mitteln, Der Spiegel, <https://magazin.spiegel.de/EpubDelivery/spiegel/pdf/99311751>.
- Daston, L. und Galison, P., 2007, Objectivity, Zone Books, <https://press.princeton.edu/books/paperback/9781890951795/objectivity>.
- Datenethikkommission, 2019, Gutachten der Datenethikkommission, 1-240, Datenethikkommission der Bundesregierung Bundesministerium des Innern, für Bau und Heimat, <https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html>.
- Desiere, S., Langenbucher, K. und Struyven, L., 2019, Statistical profiling in public employment services: An international comparison (OECD Social, Employment and Migration Working Papers No. 224). <https://doi.org/10.1787/b5e5f16e-en>.
- Desiere, S. und Struyven, L., 2020, Using Artificial Intelligence to classify Jobseekers: The Accuracy-Equity Trade-off. Journal of Social Policy, 1–19, <https://doi.org/10.1017/S0047279420000203>.
- Dobbe, R., Dean, S., Gilbert, T., Kohli, N., 2018, A broader view on bias in automated decision-making: Reflecting on epistemology and dynamics.
- Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C. und Venkatasubramanian, S., 2017, Runaway feedback loops in predictive policing. ArXiv:1706.09847 [Cs, Stat]. <http://arxiv.org/abs/1706.09847>.
- Eubanks, V. 2018, Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.
- Friedler, S. A., Scheidegger, C. und Venkatasubramanian, S., 2016, On the (im) possibility of fairness. ArXiv Preprint ArXiv:1609.07236.
- Friedman, B. und Nissenbaum, H., 1996, Bias in computer systems. ACM Transactions on Information Systems (TOIS), 14(3), 330–347.
- Fröhlich, W. und Spiecker genannt Döhmann, I., 2019, Können Algorithmen diskriminieren?, Verfassungsblog, <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/>.
- Gaigg V. und Simoner M., 2018, Verfassungsgerichtshof bestätigt Recht auf drittes Geschlecht, <https://www.derstandard.at/story/2000082511550/verfassungsgerichtshof-bestaetigt-recht-auf-drittes-geschlecht>.
- Gajane, P. und Pechenizkiy, M., 2017, On formalizing fairness in prediction with machine learning. ArXiv Preprint ArXiv:1710.03184.
- Gandy, O. H., 2016, Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage. Routledge.
- Gangadharan, S. und Niklas, J., 2019, Decentering technology in discourse on discrimination. Information, Communication & Society 22, 882–899.
- Hainmueller, J., Hofmann, B., Krug, G. und Wolf, K., 2016, Do lower caseloads improve the performance of public employment services? New evidence from German employment offices. The Scandinavian Journal of Economics, 118(4), 941–974.

- Hodosi, G., Johansson, D. und Rusu, L., 2017, Does it matter the loss of tacit knowledge in IT outsourcing? A study in a Swedish Governmental Agency. *Procedia Computer Science*, 121, 491–502.
<https://doi.org/10.1016/j.procs.2017.11.066>.
- Johnson, T., Scoppetta, A. und Buckenleib, A., 2018, Tackling long-term unemployment through risk profiling and outreach: A discussion paper from the employment thematic network. (Technical Dossier No. 6). Retrieved from ESF Transnational Platform website: <https://publications.europa.eu/en/publication-detail/-/publication/ee87b6da-6d21-11e8-9483-01aa75ed71a1/language-en>.
- Kitchin, R., 2014, *The data revolution: Big data, open data, data infrastructures and their consequences*, Sage.
- Konle-Seidl, R., 2011, Profiling systems for effective labour market integration, Use of profiling for resource allocation, action planning and matching, PES to PES Dialogue, European Commission Mutual Learning Programme for Public Employment Services.
- Kureková, L. M., 2014, Review of profiling systems, categorization of jobseekers and calculation of unit service costs in employment services-implications and applications for Slovakia, Central European Labour Studies Institute (CELSI).
- Laub, I., 2020, Warum der polnische „AMS“-Algorithmus gescheitert ist, <https://epicenter.works/content/warum-der-polnische-ams-algorithmus-gescheitert-ist>.
- Lopez, P., 2019, Reinforcing Intersectional Inequality via the AMS Algorithm in Austria, *Proceedings of the 18th Annual IAS-STS Conference on Critical Issues in Science and Technology Studies*, i. E.
- Loxha, A. und Morgandi, M., 2014, Profiling the Unemployed: A Review of OECD Experiences and Implications for Emerging Economies.
- Mager, A., 2017, Search Engine Imaginary. Visions and Values in the Co-Production of Search Technology and Europe. *Social Studies of Science*, 47(2), 240–262.
- Matzkat, L., Zielinski, L., Cocco, M., Penner, K., Spielkamp, M., Gießler, S., Lang, S. und Thiel, V., 2019, *Atlas der Automatisierung, Automatisierte Entscheidungen und Teilhabe in Deutschland*, Berlin: AW AlgorithmWatch GmbH.
- McQuillan, D., 2018, Data Science as Machinic Neoplatonism. *Philosophy & Technology*, 31(2), 253–272, <https://doi.org/10.1007/s13347-017-0273-3>.
- National Labour Market Authority, 2011, Employability Profiling System – The Danish Experience, <https://ec.europa.eu/social/BlobServlet?docId=7584&langId=en>.
- Niklas, J., Sztandar-Sztanderska, K. und Szymielewicz, K., 2015, Profiling the unemployed in Poland: Social and political implications of algorithmic decision making, Fundacja Panoptykon, Warsaw Google Scholar.
- Obermeier, T., und Oschmiansky, F., 2014, Die öffentliche Arbeitsvermittlung, <https://www.bpb.de/politik/innenpolitik/arbeitsmarktpolitik/55132/oeffentliche-arbeitsvermittlung?p=all>.
- O’Neil, C., 2016, *Weapons of math destruction: How big data increases inequality and threatens democracy*, Broadway Books.
- Penz, O., Glinsner, B., Gaitsch, M., Hofbauer, J. und Sauer, B., 2015, Affektive Interaktionsarbeit in der öffentlichen Arbeitsvermittlung in Österreich, Deutschland und der Schweiz. *AIS-Studien*, 8(1), 21-36.
- Pultz, S., 2016, Governing homo economicus: Risk management among young unemployed people in the Danish welfare state, *Health, Risk & Society*, 18(3/4), 168-187, <https://doi.org/10.1080/13698575.2016.1190003>.

- Rieder, G. und Simon, J., 2016, Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data, *Big Data & Society*, 3(1), 2053951716649398, <https://doi.org/10.1177/2053951716649398>.
- Schiprowski, A., 2017, The role of caseworkers in unemployment insurance: Evidence from Unexpected Absences, 50.
- Sierminska, E., 2015, Does it pay to be beautiful?, *IZA World of Labor*.
- Silverman, D., 2015, *Interpreting qualitative data: Methods for analyzing talk, text and interaction*, London: SAGE, 5th edn.
- Staunæs, D. und Brøgger, K., 2020, In the Mood of Data and Measurements. Experiments as Affirmative Critique or How to Curate Academic Value with Care.Feminist Theory, in print.
- Treasury Board of Canada, 2019, Directive on automated decision-making.
- Veale, M., Van Kleek, M. und Binns, R., 2018, Fairness and accountability design needs for algorithmic support in High-Stakes Public Sector Decision-Making, *Proceedings of CHI 2018*, ACM, 440.
- Wattenberg M., Viégas F. und Hardt M., 2016, Attack discrimination with smarter machine learning. <https://research.google.com/bigpicture/attacking-discrimination-in-ml/>.
- Weichselbaumer, D., 2016, Discrimination against female migrants wearing headscarves, SSRN Scholarly Paper ID 2842960, Social Science Research Network, <https://papers.ssrn.com/abstract=2842960>.
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., Myers West, S. Richardson, R., Schultz, J. und Schwartz, O., 2018, AI Now Report 2018. AI Now Institute, https://ainowinstitute.org/AI_Now_2018_Report.html.
- Wijnhoven, M. A. und Havinga, H., 2014, The Work Profiler: A digital instrument for selection and diagnosis of the unemployed, *Local Economy*, 29(6–7), 740–749, <https://doi.org/10.1177/0269094214545045>.
- Will, A.-K., 2019, The German statistical category „migration background”: Historical roots, revisions and shortcomings, *Ethnicities*, 19(3), 535–55.

DOKUMENTE-VERZEICHNIS

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikationsdatum
AGB	ALLGEMEINE VERTRAGSBEDINGUNGEN ARBEITSMARKTSERVICE	AMS	01.09.2015
AUSSCHR_1	AUSSCHREIBUNGSUNTERLAGE VERGABEVERFAHREN „Entwicklung eines Integrationschance-Prognosemodell zur KundInnensegmentierung im AMS“	AMS	09.11.2015
AUSSCHR_2	FORMVORLAGEN UND FORMBLÄTTER	AMS	09.11.2015
BEGL_1	Beschäftigungsintegration von AMS- Kundinnen/ -Kunden Beobachtung, Analyse, Prognose	Jürgen Holl, Günter Kernbeiß Michael Wagner-Pinter (Synthesis Forschung)	17.04.2019
BER_1	Evaluierung des Betreuungsformates für Personen mit multiplen Vermittlungshindernissen (BBEN) – Management Summary	Eva Auer, Petra Tamler (AMS), Friederike Weber (Prospect)	01.01.2019
BER_2	Profiling-systeme für eine effektive Arbeitsmarktintegration: Neue Ansätze für berufliches Profiling und holistisches Assessment	Jenny Bimrose, Sally Anne Barnes (Warwick Institute of Employment Research, European Commission (EC))	01.05.2011
BER_3	Bericht vom PES – Profiling Workshop: „Profiling for customers at risk from long-term unemployment“	Trude Hausegger (Prospect Research and Solution)	01.04.2015
BER_4	Evaluierung des Betreuungsformates für Personen mit multiplen Vermittlungshindernissen (BBEN) Endbericht	Eva Auer, Petra Tamler (AMS), Friederike Weber (Prospect Research & Solution)	01.01.2019
BER_5	Evaluierung der PPC-Pilotierung: Perspektivencheck Begleitforschung zur Pilotierung. im Auftrag des AMS Österreich	Claudia Liebeswar, Mario Taschwer und Andrea Egger-Subotitsch (abif, AMS)	01.02.2019
BER_6	Arbeitsmarktchancen-Assistenzsystem – AMS-Interne Befragung	Michael Auer, Tobias Krüse (AMS, Prospect Research & Solution)	18.11.2019
BER_7	AMS–Chance 2020 Das AMS–Arbeitsmarktchancen–Modell	Ernst Haider, Judit Marte- Huaing, Marius Wilk (AMS) Jutta Gamper, Günter Kernbeiß, Michael Wagner- Pinter (Synthesis Forschung)	01.12.2019
BER_8	Projektkonzept Neue Segmentierungsstrategie – „Integrationschancendeskriptor“ (ICD)	Ernst Haider (AMS)	22.06.2015
BER_9	Interne Kommunikation, strategische Diskussion und Statusberichte zur Umsetzung einer neuen Strategie des AMS in der Arbeitsmarktpolitik	AMS	22.06.2015
BER_10	Arbeitsmarktchancen als Merkmal zur Bildung von KundInnengruppen im AMS	Ernst Haider (AMS)	20.04.2018
BER_11	Profiling 2008/09 Schlussfolgerungen aus den internationalen Erfahrungen mit Profiling und Targeting und der Pilotversuch 2008/09	Synthesis Forschung	01.05.2009
BER_12	Vom Arbeitsmarkt vorgegeben: Wie die Verteilung von individuellen Integrationschance das Handlungsfeld des AMS strukturiert	Petra Gegeritsch, Stefanie Gude, Paul Timar, Michael Wagner-Pinter (Synthesis Forschung)	01.11.2011

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikationsdatum
BER_13	Arbeitsmarktpolitische Zielvorgaben	BM Beate Hartinger-Klein (Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz)	01.02.2019
BER_14	Bericht der Volksanwaltschaft an den Nationalrat und an den Bundesrat: Kontrolle der öffentlichen Verwaltung	Volksanwaltschaft	01.03.2019
BER_15	Bericht des Rechnungshofes: Arbeitsmarktservice (AMS)	Rechnungshof	22.12.2017
BRIEF_1	AMS-Arbeitsmarktchancenmodell; Algorithmus zur Chancenbewertung	Herbert Buchinger (AMS)	15.03.2015
BRIEF_2	Arbeitsmarktchancendeskriptor; Algorithmus zur Chancenbewertung am Arbeitsmarkt	Herbert Buchinger (AMS)	10.12.2018
BRIEF_3	Beantwortung der Datenanfrage des Instituts für Technikfolgen-Abschätzung (ITA) der Österreichischen Akademie der Wissenschaften (ÖAW) zum Projekt „Soziotechnische Analyse des AMS-Algorithmus«	Ernst Haider (AMS)	04.02.2020
DAT_1	Schnittstellenbeschreibung	Synthesis Forschung	11.11.2019
DAT_2	RGS Typen/Modell inkl. Einteilung	Judit Marte-Huainigg (AMS)	17.12.2019
DAT_3	Trefferquoten: Datentabelle	AMS	16.01.2020
DOK_1	Das AMS-Arbeitsmarktchancen-Modell: Dokumentation zur Methode	Jürgen Holl, Günter Kernbeiß, Michael Wagner-Pinter (Synthesis Forschung)	01.10.2018
DOK_2	Das Assistenzsystem AMAS. Zweck, Grundlagen, Anwendung	Jutta Gamper, Günter Kernbeiß, Michael Wagner-Pinter (Synthesis Forschung)	01.02.2020
DOK_3	Die Informationsbasis des Assistenzsystems AMAS. Statistisch methodische Dokumentation zur Version „Logistische Regression«	Synthesis Forschung	01.02.2020
DSB	Bescheid der Datenschutzbehörde	DSB	19.08.2020
HAND_1	Handbuch AMS-REQ-000930 AMAS Arbeitsmarktchancen – Assistenz-System	AMS, IBM	20.11.2019
INFO_1	Kundinnen-Information zur Datenschutz-Grundverordnung – DSGVO	AMS	
NOTES_1	Die Veränderung der Integrationschancen im Laufe eines Geschäftsfalles: Anomalie oder Informationsquelle?	Jürgen Holl, Günter Kernbeiß, Michael Wagner-Pinter (Synthesis Forschung)	01.01.2017
NOTES_2	Klausur des Vorstandes mit den LandesgeschäftsführerInnen 25. – 27.2.2015: Zusammenfassung und Vorschläge	AMS	19.01.2015
NOTES_3	Unterlage zu den Fragestellungen vom 16.3.2020	Synthesis Forschung	30.03.2020
NOTES_4	»AMS-Algorithmus« am Prüfstand	Michael Wagner-Pinter (Synthesis Forschung)	30.03.2020
NOTES_5	Rückfragen an Synthesis/Antworten	Günter Kernbeiß (Synthesis Forschung)	08.09.2020
NOTES_6	Personenbezogene Wahrscheinlichkeitsaussagen (»Algorithmen«): Stichworte zur Sozialverträglichkeit	Jürgen Holl, Günter Kernbeiß, Michael Wagner-Pinter (Synthesis Forschung)	09.05.2019
NOTES_7	Anfragebeantwortung AMS/Epicenter.Works	Marius Wilk (AMS)	16.08.2019

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikationsdatum
NOTES_8	Datenanfrage des Instituts für Technikfolgen-Abschätzung der Österreichischen Akademie der Wissenschaften an das AMS	Doris Allhutter, Astrid Mager (ITA, ÖAW)	23.10.2019
NOTES_9	Verifizierung letzter Unklarheiten	Fabian Fischer (TU Wien)	31.08.2020
NOTES_10	Anfragebeantwortung AMS/GBA	GBA/Herbert Buchinger (AMS)	24.04.2019
PARL_1	Parlamentarische Anfragen (Heinisch-Hosek): „Personalisierte Arbeitsmarktbetreuung“ durch das AMS und Algorithmus zur Segmentierung von beim AMS vorgemerkten Arbeitsuchenden	Abg. Heinisch-Hosek (Parlament)	22.11.2018
PARL_2	Parlamentarische Anfrage: Beantwortung	BM Beate Hartinger-Klein (Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz)	17.01.2019
PFLI_1	Pflichtenheft KundInnensegmentierung	Ernst Haider, Marius Wilk (AMS, IBM)	15.06.2016
PFLI_2	Pflichtenheft KundInnensegmentierung	Ernst Haider, Marius Wilk (AMS, IBM)	30.07.2018
POLICY_1	Profiling tools for early identification of jobseekers who need extra support	Kristine Langenbacher, Theodora Xenogiani (OECD)	01.12.2018
PPC_CHECK_1	Checkliste Berufswahl	AMS	
PPC_FORM_1	Information zu Zielsetzung und diagnostische Verfahren	AMS	
PPC_FRAGE_1	Fragebogen für den Einstieg in den „Perspektivencheck“	AMS	
PPC_FRAGE_2	Fragebogen Arbeitsbewältigung	AMS	
PPC_FRAGE_3	Arbeitsbewältigungsindex Verrechnungsbogen (Anleitung f. SB zum Ausfüllen des Berechnungstools)	AMS	
PPC_FRAGE_4	Ergebnisbericht: Mit neuen Perspektiven zu besseren Arbeitsmarktchancen	AMS	
PPC_FRAGE_5	Ergebnisbericht: Mit neuen Perspektiven zu besseren Arbeitsmarktchancen: Ausfüllhilfe	AMS	
PPC_FRAGE_6	Dokumentationsblatt Handlungsempfehlungen	AMS	
PPC_HAND_1	Handanweisung Perspektivencheck	AMS	
PPC_LEIT_1	Leitfadengestütztes Interview: Termin 1	AMS	
PPC_LEIT_2	Leitfadengestütztes Interview: Termin 2	AMS	
PPC_RICHT_2	Unterlage zur Begehrensstellung zur Förderung des Projektes Perspektivencheck zur Abklärung der Arbeitsmarktchancen	AMS	01.10.2019
PPC_RICHT_3	„Perspektivencheck“ (BBEP) Umsetzungsinfo für RGS und LGS	AMS	01.10.2019
PPC_TOOL_1	Berechnungstool Arbeitsbewältigung	AMS	
PRÄS_1	AMS: strategische Herausforderungen – Überleben für weitere 20 Jahre	Herbert Buchinger (AMS)	01.02.2015
PRÄS_2	Personalisierte Arbeitsmarktbetreuung – KundInnengruppen im AMS	Herbert Buchinger (AMS)	01.06.2018
PRÄS_3	Arbeitsmarktchancen als Merkmal zur Bildung von KundInnengruppen im Arbeitsmarktservice	Ernst Haider (AMS (Bundesgeschäftsstelle/VOR))	01.06.2018
PRÄS_4	KundInnengruppen nach Arbeitsmarktchancen: kumulative Erweiterung/Einschränkung	„AMS7594“ (AMS)	17.12.2019

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikationsdatum
PRÄS_5	AMS-Profilung im Pilotbetrieb	BGS/SfA: Elisabeth Oehry, BGS/ABI: Sabine Putz, Synthesis: Daniel Kamleitner (AMS/Synthesis Forschung)	01.01.2008
PRÄS_6	Profilingtool: Umsetzung im AMS Wien	Thomas Haider/LGS Wien (AMS)	27.08.2012
PRÄS_7	KundInnensegmentierung OÖ	AMS, Synthesis Forschung	09.07.2015
PROT_12_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	17.05.2016
PROT_13_JF	Jour Fixe Protokoll AMS IT: KundInnensegmentierung Projektstatus Meeting	AMS	06.06.2016
PROT_13_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	06.06.2016
PROT_14_JF	Jour Fixe Protokoll AMS IT: KundInnensegmentierung Projektstatus Meeting	AMS	20.06.2016
PROT_14_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	20.06.2016
PROT_15_JF	Jour Fixe Protokoll AMS IT: KundInnensegmentierung Projektstatus Meeting	AMS	27.06.2016
PROT_15_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	27.06.2016
PROT_16_JF	Jour Fixe Protokoll AMS IT: KundInnensegmentierung Projektstatus Meeting	AMS	17.08.2016
PROT_16_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	17.08.2016
PROT_17_JF	Jour Fixe Protokoll AMS IT: KundInnensegmentierung Projektstatus Meeting	AMS	24.08.2016
PROT_17_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	24.08.2016
PROT_18_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	05.09.2016
PROT_19_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	25.01.2017
PROT_20_TO	Tagesordnung Jour Fixe – Synthesis	Synthesis Forschung	20.02.2017
PROT_21_JF	Jour Fixe Protokoll AMS IT KundInnen-SegmentierungsUpdate 2018-Besprechung	AMS	14.06.2017
PROT_21_TO	Tagesordnung Jour Fixe – Synthesis „AMS Chance Update“	Synthesis Forschung	14.06.2017
PROT_22_JF	Jour Fixe Protokoll AMS IT KundInnen-SegmentierungsUpdate 2018-Besprechung	AMS	07.09.2017
PROT_22_TO	Tagesordnung Jour Fixe – Synthesis „AMS Chance Update“	Synthesis Forschung	07.09.2017
PROT_ORG_1	KundInnensegmentierung und neue Betreuungsstrategie im AMS: Stand der Arbeitsgruppen	Ernst Haider (AMS)	03.11.2015
PROT_ORG_2	Strategieausschuss: AM-Changes als Merkmal zur Bildung KundInnengruppen	AMS	03.07.2018
PROT_ORG_3	Protokoll Verwaltungsratssitzung: TOP 15.1. Personalisierte Arbeitsmarktbetreuung Assistenz-System	AMS	04.12.2018
PROT_ORG_4	Protokoll Verwaltungsratssitzung: TOP 9 Grundsatzbeschluss zur Umsetzung der PAMAS-Strategie	AMS	25.06.2019
PROT_ORG_5	Protokoll Verwaltungsratssitzung: TOP 5 Grundsatzbeschluss zur Umsetzung der PAMAS-Strategie	AMS	17.09.2019

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikationsdatum
PROT_ORG_6	Protokoll Verwaltungsratssitzung: TOP 4 Entwicklung eines Prognosemodells zur KundInnensegmentierung (Synthesis Forschung; Univ.-Prof. Dr. Wagner-Pinter)	AMS	08.03.2016
PROT_ORG_7	Protokoll Verwaltungsratssitzung: TOP 14 BBE Neu für Personen mit multiplen Vermittlungshindernissen; Konzeption und Erfolgsmessung	AMS	08.11.2016
PROT_ORG_8	Protokoll Verwaltungsratssitzung: TOP 9 Strategie im AMS – BBENeu – Modellvergleiche aus den Bundesländern	AMS	14.12.2016
PROT_ORG_9	Protokoll Verwaltungsratssitzung: TOP 6 Umsetzung der Pilotprojekte für Personen mit multiplen Vermittlungshindernissen	AMS	28.03.2017
PROT_ORG_10	Protokoll Verwaltungsratssitzung: TOP 10 Personalisierte Betreuung von Arbeitslosen aufgrund von Arbeitsmarktchancen als Merkmal zur Bildung von KundInnengruppen im AMS	AMS	08.05.2018
PROT_ORG_11	Strategieausschuss: 5. Funktionsperiode, 21. Sitzung, Protokoll, Teil 1	AMS	06.11.2019
PROT_ORG_12	Tagesordnung Präsidium 3.1.2015 – Screenshot	AMS	03.11.2015
PROT_ORG_13	Präsidium des Verwaltungsrates – 4. Funktionsperiode – Protokoll 14. Sitzung	AMS	03.11.2015
RICHT_1	Bundesrichtlinie Kernprozess Arbeitskräfte unterstützen	Herbert Buchinger, Johannes Kopf (AMS)	20.12.2019
RICHT_2	Arbeitsmarktchancen – Assistenz-System (AMAS): KundInnenfragen beantworten	AMS	16.12.2019
SCHU_1	Programm AMAS TTT Veranstaltung (Train the Trainers)	AMS	16.03.2020
SCHU_2	AMAS ganz kurz	Bettina Huber (AMS)	09.12.2019
SCHU_3	HANDOUT – Sfa – AMAS – Arbeitsmarktchancen – Assistenz-System	Alexander Peinhaupt (AMS)	27.01.2020
SCHU_5	Eine kompakte Information über das Arbeitsmarktchancen Assistenz-System (AMAS) für externe Trainer und Trainerinnen des AMS	Bettina Huber, Judit Marte-Huainigg (AMS)	09.03.2020
SCHU_6	Übersicht über die Kontaktintervalle	Gerlinde Wieser-Böhm (AMS)	01.01.2020
SCHU_7	Arbeitsmarktchancen Assistenz-System – Informationsworkshop	Karin Ostermann, Bettina Huber, Ernst Haider, Karin König, Judit Marte-Huainigg (AMS)	14.02.2020
SCR_1	Screenshots AMS interne Community – Info f. PAMAS	AMS	
SCR_2	Screenshots AMS eLearning Plattform – PAMAS	AMS	
SPEZ_1	Variablen für das Basismodell	Günter Kernbeiß (AMS, IBM)	03.06.2019
SPEZ_2	Textbausteine zur Erläuterung der Arbeitsmarktchancen Modell 2020	„AMS7594“ (AMS)	17.12.2019
SYN_1	Assessment des informationstechnischen Systems PAMAS des Arbeitsmarktservice Österreich – Worum geht es?	Synthesis Forschung	
TO_1	Tagesordnung Wissensvermittlung Kommunikation von AMAS Kommunikation – Grundlage 6 Phasen Modell	AMS	

Kürzel	Dokument-Titel/Bezeichnung	AutorInnen	Publikations datum
WP_1	Statistical Profiling in Public Employment Services: An international comparison	Sam Desiere, Kristine Langenbucher, Ludo Struyven (OECD)	13.02.2019



ÖAW

ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN

www.oeaw.ac.at/ita



INSTITUT FÜR
TECHNIKFOLGEN
ABSCHÄTZUNG

ISSN: 1819-1320 | ISSN-Online: 1818-6556