



Algorithmen und ihr Diskriminierungsrisiko

Eine erste Bestandsaufnahme



Judith Scheer

Algorithmen und ihr Diskriminierungsrisiko

Eine erste Bestandsaufnahme



Inhalt

Vorwort	6
1. Einleitung	7
2. Überblick: Algorithmische Entscheidungsprozesse	8
2.1. Begriffliche Grundlagen	8
2.2. Folgen algorithmenbasierter Entscheidungsprozesse	10
2.3. Ursachen von und Beispiele für Diskriminierungen durch ADM-Prozesse	12
2.4. Übersicht des rechtlichen Regulierungsrahmens	14
2.5. Übersicht der Forderungen nach Handlungsbedarfen	17
3. Relevanz von ADM-Prozessen für das Land Berlin	19
3.1. Einsatz von ADM-Prozessen in der Berliner Verwaltung	20
3.2. Handlungsmöglichkeiten für das Land Berlin und die LADS	24
4. Initiativen zur Auseinandersetzung mit ADM-Prozessen	26
5. Literatur	28
Anhang A: Normtext EU-DSGVO (Auszug)	33
Anhang B: Algo.Rules – Regeln für die Gestaltung algorithmischer Systeme	36

Vorwort



des Senators für Justiz, Verbraucherschutz und Antidiskriminierung, Dr. Dirk Behrendt

Die Digitalisierung immer weiterer Lebensbereiche birgt große Chancen, die Lebensqualität aller Bürger*innen zu verbessern. Gleichzeitig stellt sie einen neuen Bereich dar, der ein hohes Diskriminierungspotential aufweisen kann. So tauchen beispielsweise Themen wie *Hate Speech* und *Cybermobbing* zunehmend in gesellschaftlichen Debatten auf. Zivilgesellschaft und Wissenschaft beschäftigen sich auch mit einem weiterem, deutlich an Relevanz gewinnenden Thema: Diskriminierungsrisiken durch Algorithmen und algorithmenbasierte Entscheidungsprozesse.

Der Einfluss von Algorithmen auf unser Leben und unseren Zugang zu ökonomischen und sozialen Ressourcen und damit auf gesellschaftliche Teilhabe und Lebenschancen wächst exponentiell – genauso wie ihr Diskriminierungspotential. Die Verarbeitung personenbezogener Daten durch Algorithmen geht mit dem grundsätzlichen Risiko einher, dass Menschen mittelbar oder unmittelbar schlechter gestellt und diskriminiert werden. Die vorliegende Untersuchung nennt hier einige sehr nachdenklich machende Beispiele.

Daher freut es mich, dass Judith Scheer für die Landestelle für Gleichbehandlung – gegen Diskriminierung (LADS) mit dieser ersten Bestandsaufnahme eine verständliche Einführung zur Problematik der algorithmenbasierten Diskriminierung verfasst hat, die allen interessierten Bürger*innen eine Informationsgrundlage und Orientierungshilfe bietet.

Die Sensibilisierung sowohl der breiten Öffentlichkeit als auch der Politik und Verwaltung in diesem neuen und zentralen Handlungsfeld der Antidiskriminierungspolitik ist mir ein besonderes Anliegen.

Ich wünsche Ihnen eine spannende Lektüre.

Dr. Dirk Behrendt

Senator für Justiz, Verbraucherschutz und Antidiskriminierung

1. Einleitung

Unbestritten gewinnt das Internet einen immer größeren Einfluss auf unseren Alltag und unsere Gesellschaft. Die Digitalisierung bietet neue Möglichkeiten der Vernetzung, Kommunikation, der politischen und sozialen Partizipation sowie des Zugangs zu Informationen (Baldauf et al. 2018: 8). Auch in Bezug auf kommerzielle Aspekte ermöglicht die Digitalisierung durch Dienstleistungen und den Online-Handel einen komfortablen Zugang zu Gütern. Im Jahr 2018 hatten zum ersten Mal über 90 Prozent der deutschsprachigen Bevölkerung ab 14 Jahren einen Internetzugang. Die Anzahl der Personen, die das Internet täglich nutzen, lag bei 77 Prozent und auch die tägliche Nutzungszeit stieg in den letzten Jahren kontinuierlich (ARD/ZDF 2019).

Den vielen neuen Chancen, die eine digitale Gesellschaft ermöglicht, stehen jedoch auch neue Herausforderungen gegenüber. Durch die rasante Entwicklung und die stetig steigende Komplexität der Onlinewelt stellen sich neue Fragen zu Grundrechten wie Meinungsfreiheit und Privatsphäre, zu Datenschutz und -sicherheit, Informationsfreiheit, zur öffentlichen Sicherheit und letztendlich zur Demokratie (Baldauf et al. 2018: 8). Phänomene wie *Hate Speech* und *Cybermobbing* in sozialen Medien und die Verbreitung von *fake news* und extremistischen Inhalten sind Beispiele für die „dunkleren Aspekte des Internets“ und werden mittlerweile öffentlich diskutiert (ebd.). Andere Formen der Ungleichbehandlung durch Algorithmen im digitalen Raum waren bis vor kurzem fachspezifischer Natur und kommen erst allmählich im politischen Diskurs an (u.a. dazu Jaume-Palasi/Spielkamp 2017). Wie aktuell das Thema ist, zeigt die am 13.02.19 vom Europarat verabschiedete Aufforderung an die Mitgliedsstaaten, sich den Risiken automatisierter Entscheidungen für demokratische Gesellschaften entgegenzustellen (Committee of

Ministers 2019). Weitere Diskriminierungsrisiken durch die Digitalisierung, wie beispielsweise die Verwehrung von Teilhabe durch einen mangelhaften oder fehlenden Zugang zum Internet, scheinen bislang nur auf der Bildungsagenda zu stehen, wie die Diskussion um den Digitalpakt des Bundes und der Länder zeigt (Spiegel Online 2019). In jedem Fall ergeben sich durch die Digitalisierung neue Handlungsfelder für die Antidiskriminierungspolitik und -arbeit.

Aussagen wie „Lieber Kongress, es ist nicht länger okay, nicht zu wissen, wie das Internet funktioniert“¹ (Baldauf/Guhl 2018: 68), oder der Wunsch aus der Zivilgesellschaft, „dass Politiker endlich begreifen, wie das Netz funktioniert, damit sie sinnvolle Gesetze (...) auf den Weg bringen können“² (Böckem/Ley 2019), verdeutlichen, wie wichtig die Auseinandersetzung von Politik und Verwaltung mit den Funktionsweisen und Auswirkungen der Digitalisierung ist. Insbesondere im Bereich der Antidiskriminierungsarbeit fehlt es jedoch – außerhalb der Beleuchtung von *Hate Speech* – an Forschung und Erkenntnissen über die gegenseitige Einflussnahme von Diskriminierung und Digitalisierung sowie die Auswirkungen derselben auf Individuen und Gesellschaft. So konstatiert Jaume-Palasi (2017): „Im Zweifel müssen wir erstmal verstehen, was online passiert.“

Der Fokus dieser Bestandsaufnahme liegt auf einer ersten Analyse von Diskriminierungsrisiken durch Algorithmen bzw. algorithmische Entscheidungsprozesse (*Algorithmic Decision Making*, im Folgenden ADM-Prozesse), da – wie eben erläutert – für dieses Handlungsfeld bisher nur wenig Wissen und Sensibilität vorhanden ist.³

- Die zunehmende Digitalisierung bietet Chancen und Risiken für unsere Gesellschaft.
- Dadurch ergeben sich neue Handlungsfelder in der Antidiskriminierungspolitik und -arbeit.
- Der Fokus dieser Bestandsaufnahme liegt auf Diskriminierungsrisiken durch algorithmische Entscheidungsprozesse.

1 Übersetzung durch die Autorin. Die Aussage stammt aus dem Jahr 2011 von der Journalistin Janus Kopfstein. Sie zielte darauf ab, die fehlende technische Kompetenz der Abgeordneten bei einer Anhörung zu einem Gesetzesentwurf zur Verhinderung der Online-Verbreitung von urheberrechtlich geschützten Inhalten im US-Senat zu kritisieren.

2 Die Aussage stammt von Hannes Ley, dem Gründer der Facebook-Gruppe #ichbinhier, die sich gegen Hass im Internet und für digitale Zivilcourage einsetzt.

3 So wissen laut einer aktuellen Studie 48 Prozent der Europäer*innen nicht, was Algorithmen sind und dass sie in fast jedem Lebensbereich an Einfluss gewinnen (Grzymek/Puntschuh 2019: 7).

2. Überblick: Algorithmische Entscheidungsprozesse

In diesem übergeordneten Kapitel werden zunächst die wichtigsten (technischen) Begriffe definiert. Daraufhin wird auf die Folgen algorithmischer Entscheidungsprozesse eingegangen. Durch verschiedene Beispiele vor allem aus den USA werden anschließend die Ursachen von Diskriminierung durch ADM-Prozesse skizziert. Es folgt ein Überblick über die einschlägigen rechtlichen Rahmenbedingungen zu ADM-Prozessen, in dem neben antidiskriminierungsrechtlichen Gesetzen auch die im Mai 2018 in Kraft getretene europäische Datenschutzgrundverordnung (EU-DSGVO) relevant ist. Es folgt, unter Berücksichtigung verschiedener Politikfelder, eine Übersicht zu Forderungen und Handlungsbedarfen über eine Regulierung von ADM-Prozessen.

2.1. Begriffliche Grundlagen

Algorithmus

Ein Algorithmus ist ein Werkzeug zur automatischen Lösung eines mathematischen Problems. Mit den nötigen Informationen – dem Input – liefert er durch seine Berechnungen ein Resultat – den Output. Ein Beispiel für eine Anwendung ist die Berechnung einer Wegstrecke. Die Straßenkarte und die Informationen zum Standpunkt A und zum Ziel B stellen den Input dar. Der Algorithmus berechnet die kürzeste Fahrtzeit, den Output. Das Verhältnis zwischen Input und Output ist das mathematische Problem. Diese recht einfachen Algorithmen und Problemstellungen sind weit verbreitet (bspw. in der Suchfunktion von Datenbanken). Laut Zweig (2018) stehen diese „Algorithmen (...) bisher aber nicht in Verdacht, die gesellschaftliche Teilhabe von Menschen zu erhöhen oder zu senken“ (Zweig 2018: 11).

Lernende Algorithmen

Unter lernenden Algorithmen wird eine spezielle Form von Algorithmen verstanden, „die das aktuelle Verhalten von Menschen bewerten und/oder Vorhersagen über das zukünftige Verhalten von Personen machen“ (ebd.). Im Bereich des maschinellen Lernens bzw. der Künstlichen Intelligenz bedeutet ‚lernen‘, dass der Algorithmus aus einer sehr großen Datenmenge (*Big Data*) bisher unbekannte Muster identifiziert, die in Korrelation mit dem menschlichen Verhalten stehen (Zweig 2018: 11f.). Diese besondere Form des maschinellen Lernens wird auch *deep learning*⁴ genannt. Einige Algorithmen ‚lernen‘ während ihres Einsatzes weiter, da sie stetig mit neuen Daten versorgt werden (Zweig 2018: 27). Aufgrund künstlicher neuronaler Netze sind keine vorher programmierten Modelle mehr nötig. Hinter diesen

Netzen stehen komplexe mathematisch-statistische Funktionen. Die theoretische Nachvollziehbarkeit der Ergebnisse dieser vielschichtigen Entscheidungsfindungen ist nicht mehr möglich, so dass Entwickler*innen oft nicht wissen, wie bestimmte Ergebnisse zustande kommen (Schallbruch 2018: 48f.).

Profiling

Unter dem Begriff *Profiling*, werden anhand der durch den Algorithmus durchgeführten Analysen Vorhersagen zum Verhalten angefertigt, indem vorherige Verhaltensweisen automatisiert bewertet werden. Die Vorhersagen zielen vor allem auf private Aspekte wie Kaufkraft, Ortswechsel, Vorlieben, Gesundheit usw. ab. Online-Shops nutzen die Daten ihrer Kund*innen, um weitere personalisierte Produkte vorzuschlagen, die mit einer hohen Wahrscheinlichkeit gekauft werden, da sie auch von anderen Kund*innen mit ähnlichem Kaufverhalten erworben wurden. Je nach Funktion erstellen die lernenden Algorithmen anhand der Daten eine Priorisierung (z.B. die Suchfunktion bei Google), eine Klassifizierung bzw. Scoring (z.B. *Creditscore* der Schufa), eine Zuordnung (z.B. das Empfehlungssystem von Amazon) oder eine Filterung (z.B. Spamfilter) (Busch 2018). Neben den genannten Werbeinstrumenten wird *Profiling* auch vermehrt von staatlichen Stellen, beispielsweise der Polizei, eingesetzt (Berliner Beauftragte für Datenschutz und Informationsfreiheit 2018a: 207).

Decision Support System und Automated Decision Making

Lernende Algorithmen werden auch eingesetzt, um menschliche Entscheidungen zu unterstützen (*decision support system*). Diese sind in Deutschland weit verbreitet und werden zum Beispiel bei der Errechnung von Kreditwürdigkeit eingesetzt. Am Ende entscheidet ein Mensch, wie die Informationen genutzt werden. Ein weiterer Einsatz sind die Systeme, die eigenständige bzw. automatische Entscheidungen treffen (*automated decision-making system*). Beispielweise würde ein erster Algorithmus die finanzielle Situation einer Person einschätzen, ein weiterer Algorithmus würde daraufhin die Auszahlung von Arbeitslosengeld einstellen oder ausführen (Zweig 2018: 12). Dabei wird der von Menschen implementierte Ermessensspielraum für Einzelfälle eingeschränkt oder ist gar nicht mehr möglich (Vieth/Wagner 2017: 11). Daher ist der Einsatz solcher Prozesse für das deutsche Sozialrecht nur zulässig, wenn kein Ermessensspielraum besteht. Der vollständig automatisch erstellte Steuerbescheid ist ein bekanntes Beispiel für den Einsatz eines *automated decision-making*

⁴ Zum besseren Verständnis wird in dieser Bestandsaufnahme der Begriff der lernenden Algorithmen verwendet.

systems in der Verwaltung (Busch 2018: 16f.). Wie hoch der Anteil dieser Systeme in Deutschland ist, ist nicht bekannt (Zweig 2018: 12).

Der Vollständigkeit halber sei an dieser Stelle erwähnt, dass es algorithmenbasierte Entscheidungssysteme gibt, die keine direkte (diskriminierende) Auswirkung auf Personen haben, wie beispielsweise der Einsatz von Kameras in der Produktion, um fehlerhafte Produkte auszusortieren. Zweig unterscheidet zwischen diesen Algorithmen und „teilhaberelevanten Algorithmen“ (Zweig 2018: 13). Grundsätzlich sind lernende Algorithmen aufgrund ihrer Vielfältigkeit und ihrer verschiedenen Einsatzgebiete ein nicht einfach zu umreißender Forschungsgegenstand (Busch 2018: 10).

Algorithmischer Entscheidungsprozess (ADM-Prozess)

Bei der Betrachtung von algorithmenbasierter Diskriminierung wäre es zu kurz gegriffen, sich nur mit (lernenden) Algorithmen zu beschäftigen, da diese ausschließlich rechnen und ein Ergeb-

nis präsentieren. Um der Komplexität der Thematik gerecht zu werden, ist es notwendig, von einem algorithmischen Entscheidungsprozess (ADM – *Algorithmic Decision-Making-Prozess*) zu sprechen (Zweig 2018). Vieth und Wagner bezeichnen einen ADM-Prozess als „den Gesamtprozess von der Datenerfassung über die Datenanalyse bis hin zur Deutung und Interpretation der Ergebnisse und der Ableitung einer Entscheidung oder einer Entscheidungsempfehlung aus den Ergebnissen“ (Vieth/Wagner 2017: 10). Dieser Prozess umfasst eine Vielzahl an (technischen) Schritten, Unterprozessen und verantwortlichen Akteur*innen sowie eine soziale Einbettung des Systems. Entsprechend potenziert sich die Anfälligkeit für Fehler und Diskriminierungen⁵. Insbesondere der Einbettungsprozess in den gesellschaftlichen Kontext mit seinen verschiedenen Fehlerquellen ist bisher kaum erforscht (Zweig 2018: 17).

- Ein Algorithmus beschreibt eine mathematische Handlungsanweisung.
- Ein lernender Algorithmus kann durch *Big Data* trainiert werden, um Menschen zu bewerten und ihr Verhalten im Internet vorauszusehen (*Profiling*).
- Dies wird auch *deep learning* durch künstliche neuronale Netze genannt.
- Ein algorithmischer Entscheidungsprozess oder *Algorithmic Decision-Making-Prozess* (ADM-Prozess) bezieht sich auf den Gesamtablauf der Anwendung von (lernenden) Algorithmen inkl. seiner Einbettung in den gesellschaftlichen Kontext.

⁵ Für eine schematische Darstellung von Akteur*innen, Phasen der Entwicklung sowie möglichen Fehlerquellen siehe Zweig 2018: 21.

2.2. Folgen algorithmenbasierter Entscheidungsprozesse

Der Einfluss von algorithmischen Entscheidungsprozessen ist fast in jedem Lebensbereich allgegenwärtig. Mit den Argumenten potenzieller Kostenreduzierung, Gewinnmöglichkeiten und Produktivitätssteigerung werden ADM-Prozesse für wissenschaftliche, administrative, militärische und kommerzielle Interessen genutzt (Shadowen 2017: 2). Nationale und internationale Beispiele für die Verwendung von ADM-Prozessen reichen vom Errechnen von Rückfallquoten vor Gericht (USA) und Kriminalitätsprognosen (europaweit) über die Bewertung der Kreditwürdigkeit, Online-Bewerbungen, und die Vergabe von Studienplätzen (Frankreich) bis zum autonomen Fahren (Lischka/Klingel 2017). Insbesondere im Online-Handeln werden ADM-Prozesse für personalisierte Werbung, Produktempfehlungen und sogar automatische Preisanpassungen genutzt (Busch 2018: 13). Das Ergebnis der Google-Suche, der Spamfilter des E-Mail-Programms und der vorgeschlagene neue Facebook-Freund*innen sind nur einige Beispiele dafür, wie sich ADM-Prozesse individuell auswirken (Shadowen 2017: 2).

Diese Entwicklung steckt noch in den Kinderschuhen und so prognostizieren Jaume-Palasi und Spielkamp, dass in „wenigen Jahren (...) alle Bürgerinnen und Bürger täglich auf die eine oder andere Art von Entscheidungen dieser Systeme betroffen sein (werden). Automatisierung wird alle politischen und rechtlichen Felder erfassen“ (Jaume-Palasi/Spielkamp 2017: 2). Der Einsatz von ADM-Prozessen hat Einfluss auf die eigene Lebensgestaltung, Teilhabechancen, Konsumententscheidungen und Autonomie (vzbv 2017: 3). Das Vertrauen in ADM-Prozesse beruht auf der Annahme, dass diese hocheffizient und konsistent eine große – von Menschen nicht leistbare – Fallzahl bearbeiten und die Entscheidungen bzw. Entscheidungsvorschläge der Algorithmen somit besser seien. Im Gegensatz zu Menschen seien ADM-Prozesse objektiv und neutral, da sie keine Vorurteile oder persönliche Vorlieben in ihre Entscheidungen einfließen lassen könnten (kritisch hierzu f/21 2017; Ulbricht 2017; Wolfangel 2018). So ermöglichen ADM-Prozesse konsistente Entscheidungen, da unterschiedliche Fälle immer mit den gleichen Kriterien bearbeitet würden. Dadurch bestehe die Möglichkeit einer Reduzierung menschlicher Fehler (vzbv 2017: 7).

Wenn jedoch eine Gesichtserkennungssoftware Fotos Schwarzer Menschen automatisch mit ‚Gorilla‘ beschriftet (Wolfangel 2018) oder als weiblich identifizierten Personen bei der Jobsuche im Internet schlechter bezahlte Stellen vorgeschlagen werden als männlichen Personen (Ulbricht 2017), dann wird deutlich, „dass Ungerechtigkeit und Diskriminierung keineswegs aus automatisierten Entscheidungsverfahren verbannt sind“ (f/21 2017). Diese Phänomene sind auch als *Algorithmic Bias* bekannt (Moorstedt 2017). Die Gründe dafür sind vielfältig: Software-Entwickler*innen sind nicht frei von Vorurteilen und Stereotypen, Datensätze können Verzerrungen enthalten und maschinelles Lernen nimmt menschliche Vorurteile auf und reproduziert sie (f/21 2017; Ulbricht 2017). Der Nachweis von Diskriminierungen in ADM-Prozessen ist schwierig. Beispiele und Studien sind vor allem aus den USA bekannt. In Deutschland sind die Auswirkungen und Gründe bisher wenig erforscht. Zudem mangelt es an unabhängigen Studien (Ulbricht 2017).

Wenngleich auch von der Fachwelt kritisiert (Jaume-Palasi 2017), zeigen Forderungen nach einem „digitalen Antidiskriminierungsgesetz“ und „vorurteilsfreiem Programmieren“ vom damaligen Bundesjustizminister Heiko Maas, dass die politische Debatte in Deutschland bereits begonnen hat⁶ (Maas 2017). Aktuell widmet sich die Verbraucherschutzministerkonferenz der Verhinderung von Diskriminierung durch die Regulierung von Algorithmen (Verbraucherschutzministerkonferenz 2018) und in der Zivilgesellschaft werden außerdem Kriterienkataloge für ADM-Prozesse entwickelt (iRights.Lab/Bertelsmann Stiftung 2019). Eine breite Debatte um den verantwortungsvollen Umgang mit ADM-Prozessen sowie die Sensibilisierung der Öffentlichkeit steht noch aus (f/21 2017). Schließlich wirft der Einsatz von ADM-Prozessen grundsätzlich neue ethische, rechtliche und gesellschaftliche Fragen auf (vzbv 2017: 3).

Wie problematisch der Einsatz von ADM-Prozessen auch für die Demokratie sein kann, zeigte sich im US-amerikanischen Wahlkampf 2016. Das Wahlkampfteam von Donald Trump hatte sich Unterstützung des Unternehmens Cambridge Analytica eingekauft, um gezielte Wahlwerbung durch Micro-Targeting zu ermöglichen. Anhand einer großen Datenmenge von beispielsweise *Social Media*-Nutzungen werden gezielte Persönlichkeitsprofile erstellt und die Ansprache

⁶ Der Sachverständigenrat für Verbraucherfragen (SVRV) veröffentlichte 2017 ein Gutachten zu digitaler Souveränität sowie entsprechende Handlungsempfehlungen. Des Weiteren publizierte 2018 das Kompetenzzentrum Öffentliche IT die vom Bundesinnenministerium geförderte Schrift ‚(Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft‘.

durch Wahlkampfhelfer*innen an der Tür entsprechend angepasst. Die Folgen des Einsatzes dieser Methode auf den Wahlkampf und das Wahlergebnis sind nicht eindeutig geklärt. Dabei wird aber die Möglichkeit einer Durchführung von intensiven Persönlichkeitsanalysen deutlich (Schallbruch 2018: 44f.). Auch der Einsatz von sogenannten *Social Bots* für die öffentliche Meinungsbildung im Wahlkampf wurde u.a. in den USA diskutiert. *Social Bots* sind algorithmengestützte Ro-

boter in sozialen Netzwerken, die zur Propaganda eingesetzt werden und u.a. automatisierte *fake news* verbreiten. Sie geben vor, ein Mensch zu sein, und können durch die Interaktion mit Verbraucher*innen die Reichweite von Themen und Meinungen verstärken. Dadurch kann eine Minderheitenmeinung als gesellschaftlicher Konsens dargestellt werden. In Deutschland wird diese Gefahr im Vergleich zu den USA jedoch als eher gering eingeschätzt (Neudert 2017).

- Der Einsatz von ADM-Prozessen nimmt in fast jedem Lebens- und Arbeitsbereich zu.
- Das Vertrauen in die Technik beruht auf der Annahme, diese sei neutral.
- Beispiele zeigen jedoch, dass Vorurteile und Stereotype reproduziert werden.
- Die politische Diskussion in Deutschland hat begonnen, eine breite gesellschaftliche Debatte steht noch aus.

2.3. Ursachen von und Beispiele für Diskriminierungen durch ADM-Prozesse

Die Ursachen von Diskriminierungen durch ADM-Prozesse können sehr komplex und vielfältig sein.

Schon in der Phase des Designs des ADM-Systems können durch handwerkliche Fehler zum Beispiel fehlerhafte Codes entstehen. Zusätzlich kann die Auswahl von Algorithmen, Methoden oder Operanten für die jeweilige Problemstellung unpassend sein und somit Fehler erzeugen bzw. einen ungeeigneten Output generieren. Insbesondere die Modellierung von gesellschaftlich bedeutsamen Fragen, wie zum Beispiel ‚Welche Nachrichten sind relevant?‘, ist anfällig für Vorurteile von Programmierer*innen (Zweig 2018: 21ff.; Busch 2018: 21).

Der Input, also die Daten anhand derer der Algorithmus lernt, kann Verzerrungen aufweisen. Dieser *bias* wird vom Algorithmus reproduziert und verfälscht entsprechend den Output. Dies kann an unvollständigen oder veralteten Datensätzen liegen. Die mangelhafte Qualität von Daten kann aber auch auf die Art und Weise ihrer Entstehung zurückgeführt werden. Für das sogenannte *data mining* gibt es zum einen keine Standards, zum anderen werden Datensätze oft eingekauft, ohne dass die Entwickler*innen wissen, wie diese generiert wurden (Moorstedt 2017; Shadowen 2017: 10; Zweig 2018: 23). Gleichwohl kann ein *bias* auch aufgrund der fehlenden Sensibilität der Entwickler*innen für die Notwendigkeit diverser Datensätze entstehen. So wurde vermutlich bei der Entwicklung des oben genannten Beispiels der Gesichtserkennungssoftware der lernende Algorithmus nur mit *weißen* Personen gespeist. Entsprechend war es für das Programm nicht möglich, Schwarze Personen korrekt zu identifizieren⁷ (Busch 2018 S. 21f.) Wolfangel (2018) pointiert dies mit der Aussage: „Stecken wir Vorurteile rein, kommen Vorurteile raus.“ Oft wird erst am Ende der Entwicklung deutlich, ob ein Algorithmus diskriminierend ist, und das auch nur, wenn Entwickler*innen bereit sind, ihre Ergebnisse zu hinterfragen.

Weitere Probleme mit Datensätzen können durch einen historischen *bias* entstehen (Hacker 2018a). So bewertet eine vom österreichischen Arbeitsmarktservice eingesetzte Software die Chancen von Frauen auf dem Arbeitsmarkt grundsätzlich negativ. Der Algorithmus lernte anhand von vergangenen Arbeitsmarktstatistiken, dass es Frauen auf dem Arbeitsmarkt schwerer haben, und projizierte dies auf die Zukunft. Die Ursachen für Diskriminierungen auf dem Arbeitsmarkt werden dadurch nicht aufgedeckt, sondern eher verstärkt, da Frauen durch diese negativen Kategorisierungen grundsätzlich weni-

ger Förderungen erhalten können (Fröhlich/Spiecker 2018). In einem weiteren Beispiel berichtet Moorstedt (2017) von einer Übersetzungssoftware, die akademische Berufe in die männliche Form und weniger qualifizierte Berufsbezeichnungen in die weibliche Form übersetzt. Der ADM-Prozess nutze vermutlich Datensätze aus der analogen Welt und reproduziere die bestehenden Verhältnisse bzw. die bestehende ungleiche Verteilung von Berufen auf Geschlechter.

Neben den Problematiken bezüglich des Inputs kann auch der Output eine Ursache für Diskriminierungen darstellen. Lernende Algorithmen sind darauf ausgelegt, Korrelationen zu erkennen. Ein statistischer Zusammenhang zwischen Variablen bedeutet jedoch nicht unbedingt eine Kausalität (Busch 2018: 20). Des Weiteren kann eine Fehlinterpretation des Outputs durch die Anwender*innen stattfinden, wenn nicht klar ist, dass es sich bei den Ergebnissen um eine „gruppenbasierte Wahrscheinlichkeit für ein bestimmtes Verhalten“ und nicht um Fakten handelt (Zweig 2018: 25).

Mittelbare Diskriminierungen können auch durch Korrelationen von Variablen entstehen, die nicht als geschützter Diskriminierungsgrund eingestuft werden. So stellte ein in den USA eingesetzter ADM-Prozess eine Kausalität zwischen einer hohen Fluktuation von Mitarbeiter*innen und deren Wohnortsentfernung zum Arbeitsplatz fest. Das Programm riet der Firma, Bewerber*innen nicht einzustellen, die in den Außenbezirken lebten. Diese waren aber vornehmlich Schwarze Personen. Obgleich der Datensatz aufgrund von anonymisierten Bewerbungsverfahren keine Angaben zur ethnischen Herkunft enthielt, wurde durch die Korrelation eine indirekte Diskriminierung bzw. Benachteiligung hergestellt (Hacker 2018b; Wolfangel 2018). Daher beschreibt Wolfangel (2018) lernende Systeme als „dermaßen konsequente und perfekte Korrelationsaufpürmaschinen, dass sie Zusammenhänge erkennen, die uns Menschen nicht bewusst sind.“

Die Algorithmen, die während ihres Einsatzes stetig neu dazulernen und sich mit einem Echtzeitfeedback weiterentwickeln, können eine weitere Schwierigkeit aufweisen. Erst in der Interaktion mit Verbraucher*innen und Anwender*innen kann ein gesellschaftlich relevanter Aspekt des ADM-Prozesses zum Vorschein kommen. So kann ein technisch einwandfreies ADM-System, das mit repräsentativen Daten gelernt hat, nicht-intendierte Auswirkungen haben. Beispielsweise wurde der *Chatbot Tay* programmiert, um sich in Forum-Diskussionen einzubringen. Der Einsatz wurde jedoch nach 24

⁷ Ausführlich siehe hierzu Köver 2019.

Stunden wieder eingestellt, da der *Chatbot* rechtsradikale Äußerungen von Menschen im Forum lernte und diese reproduzierte (Zweig 2018: 26).

Ein weiteres Problem kann durch selbstverstärkende Echtzeitfeedbacks entstehen. Beim *Predictive Policing* (vorrusschauende Polizeiarbeit) in den USA werden u.a. Straßenzüge anhand ihrer Kriminalitätsstatistiken kategorisiert. In besonders gefährdeten Gegenden wird dann vermehrt Streife gefahren. Indem dadurch auch mehr Straftaten entdeckt werden, lernt das System, dass die Kriminalität in dieser Gegend steigt und noch mehr Streifen eingesetzt werden sollen. Die tatsächlich verübten Verbrechen werden dabei ungleich verfolgt (Zweig 2018: 27).

Ein weiteres Beispiel von Diskriminierung durch staatliche Institutionen wurde durch die investigative Journalist*innen-Plattform ProPublica aufgedeckt. Die in den USA eingesetzte Software COMPAS wird vor Gericht verwendet, um die Rückfallquote von Straftäter*innen zu prognostizieren. Die Recherche von ProPublica ergab, dass die Software eine Korrelation zwischen einer dunklen Hautfarbe und einer erhöhten Rückfallwahrscheinlichkeit herstellte. Schwarzen Häftlingen wird dadurch fälschlicherweise eine höhere Rückfallquote als *weißen* Häftlingen attestiert (Busch 2018; Shadowen 2017: 20f.; Wolfangel 2018; Zweig 2018: 30f.). Dieser Fall ist besonders besorgniserregend, da die Software seit mehreren Jahren landesweit eingesetzt wurde und sowohl mehrere interne als auch externe Studien die Genauigkeit der Vorhersagen bescheinigten. Zudem enthielten die Testdaten, anhand derer der Algorithmus lernte, keine Angaben zur ethnischen Herkunft (Shadowen 2017: 11ff.).

Auch in Deutschland verdeutlicht ein Beispiel im Kontext von Asylverfahren, wie Behörden auf Softwarelösungen setzen und diese trotz mehrerer Fehlerquellen über Menschenleben entscheiden lassen. Mit dem Argument der Beschleunigung von Verfahren setzt das Bundesamt für Migration und Flüchtlinge (BAMF) ein preisgekröntes Programm zur Analyse von Dialekten ein, um im Zweifel die Herkunft von Asylsuchenden zu klären (Biselli 2018b). Nach einer zweitminütigen Sprachaufnahme

wird die Wahrscheinlichkeit der möglichen Herkunft der Person berechnet (Biselli 2017). Die Fehlerwahrscheinlichkeit lag jedoch Ende 2017 bei 20 Prozent und Anfang 2018 bei 15 Prozent. Durch eine falsche Prognostizierung können Menschen in Gefahr geraten, wenn der Asylantrag aufgrund eines Computers abgelehnt wird. So zeigen Recherchen, dass Entscheider*innen sich auf die Aussagen der Software verlassen, da sie nicht im richtigen Umgang mit derselben geschult sind (Biselli 2018a). Problematisch ist zudem, dass sich Dialekte von Menschen verändern und in formellen Gesprächen die eigene Sprache oft angepasst wird. Die Intransparenz der Analyse und des ADM-Prozesses ist ein weiteres Problem. Weder für die Öffentlichkeit noch für die Betroffenen besteht die Möglichkeit, den Einsatz und die Software transparent und nachvollziehbar zu verstehen (Biselli 2017).

Das letzte Beispiel zeigt noch eine weitere Komplexität auf. So wird Frauen bei der Google-Suche weniger häufig Werbung für Führungskräfte-seminare angeboten als Männern. Dies liegt jedoch nicht an verzerrten Daten oder vorurteilsbelasteten Codierungen, sondern an der Art und Weise, wie Google den Verkauf von Werbeanzeigen strukturiert. So ist die Schaltung von Werbung für Frauen im Alter von 18 bis 35 Jahren am teuersten, da diese die interessanteste Zielgruppe darstellt. Werbung, die ohne geschlechtsspezifische Präferenz geschaltet wird, ist daher nicht nur günstiger, sondern wird auch mehr Männern als Frauen angezeigt, da Frauen ausschließlich die speziell auf sie ausgerichtete Werbung erhalten (Busch 2018: 21).

Diese Vielzahl verschiedener Fehlerquellen und Diskriminierungsvorfälle verdeutlicht die Komplexität der Problematik einer Diskriminierung durch ADM-Prozesse. Zwischen den verschiedenen Fehlerquellen zu unterscheiden, ist für Betroffene daher sehr schwierig. Selbst die Entwickler*innen können die Ergebnisse zunehmend nicht mehr nachvollziehen, da sie aufgrund der Komplexität und Menge der Daten oft nicht wissen, wie das jeweilige Resultat zustande gekommen ist (Jaume-Palasi/Spielkamp 2017: 5) Dies ist jedoch für den Nachweis einer Ungleichbehandlung unabdingbar (Hacker 2018b).

Die Ursachen von Diskriminierungen durch ADM-Prozesse sind komplex und der Nachweis für Betroffene kaum möglich. Ursachen können u.a. sein:

- Handwerkliche Fehler in der Modellierung der Algorithmen
- Input-Verzerrung/*data bias*: fehlende Diversität oder historischer *bias*
- Output: Herstellung von nicht bestehenden Kausalitäten bei Korrelationen
- Fehlinterpretation des Outputs durch Anwender*innen; fehlende Schulungen
- Marktmechanismus der Werbeschaltung

2.4. Übersicht des rechtlichen Regulierungsrahmens⁸

Wie verdeutlicht wurde, birgt der Einsatz von ADM-Prozessen ein hohes Risiko für Diskriminierungen. Neben den technischen und ethischen Aspekten sind die rechtlichen Regulierungen und Rahmenbedingungen relevant, um sich mit der Problematik umfassend auseinanderzusetzen. Für diese Bestandsaufnahme sind insbesondere das Antidiskriminierungsrecht und das Datenschutzrecht ausschlaggebend⁹. Hacker empfiehlt sogar „eine integrierte Vision des Antidiskriminierungs- und Datenschutzrechts, um Fairness im digitalen Zeitalter durchzusetzen“¹⁰ (Hacker 2018a: 1). Dazu werden die beiden Rechtsbereiche zunächst in ihrer Anwendbarkeit auf Diskriminierung durch ADM-Prozesse beleuchtet sowie die Problematik der Rechtsdurchsetzung skizziert.

Allgemeines Gleichbehandlungsgesetz (AGG)

Das Allgemeine Gleichbehandlungsgesetz (AGG) ist bei einer Diskriminierung zunächst der wichtigste gesetzliche Bezugsrahmen zur Durchsetzung von Gleichbehandlung. In Bezug auf ADM-Prozesse gibt es eine Diskussion, bei der auf der einen Seite (verkürzt dargestellt) ein digitales AGG gefordert wird (Maas 2017). So geht Hacker (2018a) auch davon aus, dass verzerrte Input-Daten zu den häufigsten Gründe für algorithmenbasierte Diskriminierung gehören und diese durch die derzeitigen europäischen Antidiskriminierungsrichtlinien nicht abgedeckt seien, da in diesen Fällen die Wahrscheinlichkeit einer nicht-intendierten Diskriminierung bestehe (Hacker 2018a: 10). Auf der anderen Seite wird der Bedarf an dieser rechtlichen Anpassung eher kritisch gesehen. So behauptet Jaume-Palásí (2017), dass durch den „technischen Fortschritt (...) keine rechtsfreien Räume entstanden [sind]“. Auch Busch verweist darauf, dass die Vorschriften des AGGs „technologieneutral konzipiert [sind] und (...) daher grundsätzlich auch ADM-Prozesse [erfassen]“ (Busch 2018: 47).

Jedoch wird auch von letztgenannten Personen die Problematik der Rechtsdurchsetzung thematisiert. Dafür gibt es vielfältige Gründe. Für Einzelpersonen ist eine Diskriminierung mitunter nicht erkennbar. So ist ihnen weder klar, dass ein ADM-Prozess im Einsatz ist, noch dass eine Diskriminierung vorliegt. Auch

wenn dies der Fall wäre, bedeutet es nicht, dass eine vorliegende Diskriminierung durch das AGG abgedeckt wird (Busch 2018: 48; Hacker 2018b). Der Nachweis über eine individuelle Diskriminierung ist ohne Zugang zur Datenbasis und dem Algorithmus nicht möglich. Dieser Zugang ist für Betroffene jedoch nicht gegeben (Hacker 2018b). Jaume-Palásí (2017) gibt zudem zu bedenken, dass selbst eine Transparenz von Algorithmen und den eingespeisten Daten nicht zu mehr Gerechtigkeit führen würde, da durch die hohe Komplexität nur Expert*innen die entsprechenden Prozesse nachvollziehen könnten. Zudem ist es für Entwickler*innen und Anwender*innen von ADM-Prozessen einfach, eine Ungleichbehandlung mit den statistischen Daten und der akkuraten Wahrscheinlichkeit der Ergebnisse zu rechtfertigen. Für Betroffene sei es also unmöglich, Fehler oder *bias* im System zu beweisen (Hacker 2018a: 11f.).

Um sich dieser Problematik anzunehmen, empfiehlt Hacker (2018a; 2018b), das europäische Antidiskriminierungsrecht in Verbindung mit der EU-Datenschutzgrundverordnung (EU-DSGVO) anzuwenden. So könnten die Instrumente der DSGVO die schwierige Rechtsdurchsetzung des Antidiskriminierungsrechts kompensieren (Hacker 2018a: 30).

EU-Datenschutzgrundverordnung (EU-DSGVO)

Grundsätzlich untersagt Artikel 9 (1) DSGVO die *Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person (...)*.

Wenngleich Artikel 9 (2) Ausnahmen zu dieser Vorgabe beinhaltet, wie beispielsweise die ausdrückliche Einwilligung der betroffenen Person¹¹ in Artikel 9 (2a), so werden auch diesen Ausnahmen Rahmenbedingungen gesetzt. Bei der Verarbeitung personenbezogener Daten müssen sich Entwickler*innen und Anwender*innen von ADM-Prozessen an Grundsätze des Artikels 5 DSGVO halten. Für Algorithmen

8 Im Juni 2019 erschien die Monographie ‚Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz‘ von Mario Martini im Springer Verlag. In diesem werden ADM-Prozesse aus einer verfassungsrechtlichen Perspektive analysiert, Vergleiche zu Regulierungen in z.B. der Gendiagnostik gezogen und Reformvorschläge zur Regulierung entworfen.

9 Busch (2018) nennt zusätzlich noch das Wettbewerbsrecht (Kartellrecht und Lauterkeitsrecht) sowie sektorenspezifische Regelungen im medizinischen Bereich und im Hochfrequenzhandel.

10 Übersetzung durch die Autorin.

11 Vermutlich tun wir Verbraucher*innen dies jedoch täglich, wenn ungelesene AGBs als gelesen markiert oder Hinweise auf *Cookies* auf Webseiten weggeklickt werden.

basierte Diskriminierung argumentiert Hacker, dass ein Output, der durch verzerrte Datensätze entstehe, dem Grundsatz der Richtigkeit nach Art. 5 (d) DSGVO widerspreche (Hacker 2018a: 26). Zudem verletzt eine unberechtigte Diskriminierung durch Algorithmen den Grundsatz der rechtmäßigen bzw. fairen Datenverarbeitung nach Art. 5 (a) DSGVO. Im Erwägungsgrund 71 DSGVO wird explizit darauf hingewiesen, dass unter einer fairen Verarbeitung u.a. verstanden wird, zu „verhindern“, dass es gegenüber Personen aufgrund der in Artikel 9 (1) DSGVO genannten Merkmale „zu diskriminierenden Wirkungen oder zu einer Verarbeitung kommt, die eine solche Wirkung hat“ (DSGVO, Erwägungsgrund 71).

So ist auch der beschriebene automatisierte Entscheidungsprozess nach Artikel 22 (1) DSGVO rechtswidrig¹². Jede Person hat bei einer sie betreffenden Entscheidung mit rechtlicher oder ähnlich wichtiger Wirkung das Recht auf eine nicht ausschließlich automatisierte Verarbeitung ihrer Daten. Dies betrifft auch das *Profiling* zum Beispiel bei Bonitätsprüfungen für Kreditvergaben. Dieses Verbot bezieht sich jedoch nur auf vollständig automatisierte Entscheidungen. Die Systeme zur Unterstützung von menschlichen Entscheidungen sind davon ausgenommen. Wenn Sachbearbeiter*innen die Resultate nochmals überprüfen, bevor sie eine Entscheidung treffen, bzw. ein System nicht automatisch die Kreditvergabe in die Wege leitet, ist diese Art der Errechnung von Bonitäten legitim (Busch 2018: 30f.). Die Gefahr von Scheinprüfungen besteht jedoch, je nachdem wie unabhängig von dem Ergebnis die Anwender*innen entscheiden dürfen. Artikel 22 (4) DSGVO wiederholt die unter Artikel 9 DSGVO genannten Einschränkungen für sensible Daten in automatisierten Entscheidungsprozessen. Zudem bezieht sich der oben geschilderte Erwägungsgrund 71 DSGVO auch auf das *Profiling* unter Artikel 22 DSGVO. Folglich dürfen keine sensiblen personenbezogenen Daten bei der Entwicklung und Verwendung eines ADM-Prozesses verarbeitet werden (Hacker 2018a: 33).

Für von Diskriminierung betroffene Personen im Rahmen von ADM-Prozessen sind des Weiteren die Informationspflichten und das Auskunftsrecht der Artikel 12–15 DSGVO relevant. Artikel 12 (1) DSGVO verpflichtet die Verantwortlichen, die Informationen „in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache zu übermitteln“. Artikel 13 und 14 DSGVO verpflichten dazu, betroffene Personen über die Verarbeitung ihrer Daten zu informieren. Dies soll zum einen einer transparenten und fairen

Verarbeitung dienen. Zum anderen sollen Betroffene dadurch ihre Rechte wahrnehmen können. Daran anschließend wird durch Artikel 15 DSGVO das Auskunftsrecht für betroffene Personen sichergestellt. So haben sie das Recht auf Auskunft darüber, ob, wie und in welchem Umfang personenbezogene Daten verarbeitet werden. Für alle drei Artikel gilt zudem, dass beim *Profiling* durch automatisierte Entscheidungsprozesse Betroffene gemäß Artikel 22 DSGVO „aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung“ erhalten müssen (Art. 13 (2f), Art. 14 (2f), Art. 15 (1h) DSGVO) (Busch 2018: 38f.; Hacker 2018a: 27f.).

Wenngleich diese Verordnungen die Position von Einzelpersonen stärken, merkt Hacker kritisch an, dass für den Nachweis einer Diskriminierung der Zugang zu Vergleichsdaten notwendig ist, dies jedoch durch die DSGVO nicht ermöglicht wird (Hacker 2018a: 27). So wird argumentiert, dass mit Verweis auf Artikel 15 (1h) DSGVO Verantwortliche zusammengefasste Informationen über das Bestehen eines *bias* im ADM-Prozess bereitstellen müssen. Zudem könnten hier die oben genannten Grundsätze aus Artikel 5 DSGVO zum Tragen kommen. Jedoch gibt Hacker (2018a) zu bedenken, dass es auch bei diesen Regelungen für von (potenzieller) Diskriminierung betroffene Personen eher unrealistisch sei, dass diese nach spezifischen technischen Details fragen. Zudem können Verantwortliche zwar Betroffene über diskriminierende Effekte informieren und damit ihrer Auskunftspflicht nachkommen. Dies bedeutet aber nicht, dass gleichzeitig auch der Grund (z.B. verzerrte Daten) benannt wird. Um jedoch eine individuelle Betroffenheit von Diskriminierung geltend zu machen, ist es notwendig, diesen Grund zu kennen. Hacker (2018a) empfiehlt daher, die Informationspflichten und Auskunftsrechte mit den Durchsetzungsbefugnissen der Aufsichtsbehörden zu komplementieren.

Artikel 58 DSGVO ermöglicht den öffentlichen und unabhängigen Aufsichtsbehörden korrigierende und investigative Maßnahmen, wie die Durchführung von Datenschutzprüfungen (Art. 58 (1b) DSGVO), und den Rechtsanspruch auf alle relevanten – auch personenbezogenen – Informationen der Verantwortlichen, (Art. 58 (1a, 1e) DSGVO). Zudem sind gemäß Artikel 35 DSGVO Verantwortliche verpflichtet, Datenschutzfolgenabschätzungen durchzuführen, wenn die Verarbeitung von Daten „voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge“ hat. Bei der Durchführung dieser Verfahren ist die Aufsichtsbehörde bzw.

¹² Eine rechtskonforme Ausnahme ist bspw. der erläuterte automatisierte Steuerbescheid. Auch im Sozialrecht sind rechtlich mittlerweile vollautomatisierte Bescheide möglich (Busch 2018: 16f.).

die*der Datenschutzbeauftragte zu konsultieren (Artikel 35 (1) DSGVO). Artikel 58 (2) DSGVO gibt den Aufsichtsbehörden verschiedene Eskalationsstufen vor, die von der Verpflichtung der Verantwortlichen zu einer Korrektur von Verarbeitungsprozessen bis hin zur Verhängung eines Verbots der Verarbeitung reichen können. Dadurch könnten Verzerrungen in Datensätzen, die zur Benachteiligung einer geschützten Merkmalsgruppe führen, zunächst entdeckt und ggf. begrenzt werden. Hacker (2018a) geht sogar davon aus, dass „Kläger*innen möglicherweise endlich imstande sind, Dokumente von der Datenschutzbehörde zu erhalten, um einen Anscheinsbeweis für einen Diskriminierungsfall zu schaffen“¹³ (Hacker 2018a: 30).

Bei Verstößen gegen die Verordnung können gemäß Artikel 83 DSGVO „Geldbußen von bis zu 20.000.000 EUR oder im Fall eines Unternehmens von bis zu 4% seines gesamten weltweit erzielten Jahresumsatzes“ verhängt werden (Art. 83 (5) DSGVO)¹⁴. Neben diesen empfindlichen Strafen sieht Hacker die Datenschutzprüfungen und die Datenschutzfolgeabschätzungen als wichtige Werkzeuge an, um Gleichbehandlung durch Konstruktion (*equal treatment by design*) voranzubringen (Hacker 2018a: 31). So können Datenschutzprüfungen genutzt werden, um Diskriminierung bzw. den

Verdacht auf Diskriminierungen im Nachhinein aufzudecken. Datenschutzfolgeabschätzungen können wiederum genutzt werden, um im Vorhinein einen ADM-Prozess auf eventuell resultierende Schäden zu bewerten (ebd.).

In der Verbindung von Antidiskriminierungsrecht und Datenschutzrecht liegen sicherlich viele Chancen zur Begrenzung von Diskriminierung durch ADM-Prozesse, jedoch ergeben sich dadurch auch Probleme (Hacker 2018a: 31ff.). Da die DSGVO erst seit Mai 2018 in Kraft ist, werden sich die Herausforderungen und der Umfang ihrer Rechtsdurchsetzung noch zeigen müssen. Hacker (2018a) gibt zudem zu bedenken, dass sich Verantwortliche unter Berufung auf das Geschäftsgeheimnis gegen die Auskunftsrechte zur Wehr setzen könnten. Ein weiteres Problem liegt in dem Umgang mit nicht-intendierter bzw. indirekter Diskriminierung durch die Auswertung von nicht-geschützten Merkmalen. Zudem ist nicht klar, wie sich die Förderung von Minderheiten im Antidiskriminierungsrecht in die oben geschilderten Möglichkeiten der Kontrolle und des Datenschutzes eingliedern lässt (ebd.: 31ff.).

- Das Allgemeine Gleichbehandlungsgesetz (AGG) ist die wichtigste Rechtsgrundlage gegen Diskriminierungen.
- Es gibt eine Diskussion, ob das AGG an die neuen technischen Möglichkeiten angepasst werden muss oder ob es in der jetzigen Form ausreicht.
- Die Rechtsdurchsetzung mit dem AGG bei ADM-Prozessen ist kaum möglich.
- Die EU-Datenschutzgrundverordnung (EU-DSGVO) bietet einen weiteren Ansatzpunkt zur Durchsetzung von Gleichbehandlung im digitalen Raum.
- Der Verarbeitung von personenbezogenen Daten werden Grenzen gesetzt.
- Entscheidungen, die allein auf automatisierten Entscheidungsprozessen beruhen, sind rechtswidrig (inkl. *Profiling*).
- Es gibt Informationspflichten der Verantwortlichen und ein Auskunftsrecht für Betroffene.
- Aufsichtsbehörden sind mit verschiedenen Eskalationsstufen ausgestattet.
- Datenschutzprüfungen und Datenschutzfolgeabschätzungen sind wichtige Instrumente.

13 Übersetzung durch die Autorin. Bei einem Anscheinsbeweis wird bei bestimmten Ereignissen auf die Auswirkungen aufgrund von bisherigen Erfahrungen geschlossen.

14 Rebigier (2019) berichtet von einer ersten Strafe gegen Google wegen eines Verstoßes gegen die DSGVO in Frankreich. Wenngleich sich der Fall nicht auf Diskriminierung bezieht, wird deutlich, dass die DSGVO „erste Zähne“ zeigt.

2.5. Übersicht der Forderungen nach Handlungsbedarfen

Die bisherige Bestandsaufnahme hat die Vielschichtigkeit der Thematik verdeutlicht, die bei Regulierungsansätzen in Gänze berücksichtigt werden müsste. Grundsätzlich sollte in die Erforschung und Analyse von ADM-Prozessen und ihren Auswirkungen investiert sowie die Öffentlichkeit entsprechend sensibilisiert werden (f/21 2017; Ulbricht 2017). Die Diskussionen um Handlungsbedarfe, politische Forderungen und die Anpassung des rechtlichen Regulierungsrahmens lassen sich grob in vier Themenfelder einteilen: (1) Informationen, (2) Transparenz, (3) Anforderungen zur technischen Gestaltung und (4) Kontrollen.

(1) Die Forderungen nach mehr Informationen beim Einsatz von ADM-Prozessen beziehen sich vor allem auf die Sensibilisierung von Anwender*innen, Verbraucher*innen und Betroffenen (Shadowen 2017: 18; vzbv 2017: 15). Diese sehen eine Kennzeichnungspflicht beim Einsatz von ADM-Prozessen, insbesondere in sensiblen Bereichen, vor. Zudem soll durch eine Erläuterungspflicht bzw. ein Auskunftsrecht gewährleistet werden, dass den Betroffenen erklärt wird, wie die Ergebnisse des ADM-Prozesses zustande kommen bzw. gekommen sind (Busch 2018: 58ff.; vzbv 2017: 10). Grundsätzlich kann dabei zwischen der Information vor dem Einsatz eines ADM-Prozesses (*ex ante*) und der Information nach der Entscheidung (*ex post*) unterschieden werden (Busch 2018: 57). Die Idee eines Beipackzettels analog zu Medikamenten kann insbesondere Anwender*innen, die an der Entwicklung des ADM-Systems nicht beteiligt waren, dabei unterstützen zu verstehen, wie, wann und mit welchen Nebenwirkungen das System Verwendung finden kann. Dabei ist für Zweig insbesondere von Bedeutung, „wie der gesellschaftlich relevante Prozess auf das mathematische Problem reduziert wurde“ (Zweig 2018: 31). Daran schließt sich die Forderung nach standardisierten Schulungen für Anwender*innen an.

(2) Die in der Literatur geforderte Transparenz bezieht sich auf die Veröffentlichung des Quellcodes bzw. des Algorithmus und teilweise auch der Trainingsdaten. Dies soll eine externe Beforschung und Überprüfung ermöglichen (Shadowen 2017: 17; Zweig 2018: 31f.). Zudem soll dadurch die „Entscheidungssouveränität“ von Verbraucher*innen gewahrt werden (vzbv 2017: 10). Durch eine externe Beforschung könnten Fehler in den verschiedenen Entwicklungsphasen, wie verzerrte Datensätze oder unterschiedliche Gewichtungen einzelner Kriterien im Algorithmus, aber auch die Auswirkungen des ADM-Prozesses auf die Gesellschaft analysiert werden (Zweig 2018: 31). Es gibt aller-

dings auch Kritik an der Forderung nach Offenlegung der Codes. Während Zweig die Gefahr von Manipulationen durch Kriminelle sieht (ebd.: 29f.), könnte eine solche Transparenz der Daten für Busch den Grundsätzen des Schutzes von Geschäftsgeheimnissen widersprechen (Busch 2018: 60). Jaume-Palasi (2017) gibt zu bedenken, dass mehr Transparenz aufgrund der hohen Komplexität der Systeme nicht automatisch zu mehr Mündigkeit der Bürger*innen führen würde. Sie fordert daher, dass die Unternehmen, die ADM-Prozesse entwickeln und nutzen, in die Verantwortung genommen werden. Schließlich „verlangen [wir] [auch] nicht von Bürgern, dass sie zu Biochemikern werden, um das Essen im Supermarkt vor jedem Kauf zu überprüfen“ (Jaume-Palasi 2017). Für die Einhaltung und Überprüfung von gesetzlichen Vorgaben sollen folglich auch Produzent*innen und Behörden zuständig sein.

(3) Als Vorbild für die technischen Anforderungen kann das in den 90er Jahren entwickelte Konzept des Datenschutzes durch Technikgestaltung (*privacy by design*)¹⁵ herangezogen werden (Busch 2018: 60). Für ADM-Prozesse sind vor allem die Schlagwörter der algorithmischen Rechenschaftspflicht (*Algorithmic Accountability*) und Ethik in der Technikgestaltung (*ethics by design*) von Relevanz (ebd.; Jaume-Palasi/Spielkamp 2017; Hacker 2018a: 31; Zweig 2018: 30). Neben der grundsätzlichen Auseinandersetzung mit den gesellschaftlichen Auswirkungen durch den Einsatz von ADM-Prozessen können diese Konzepte „die Gefahr von algorithmischen Diskriminierungen durch Technikgestaltung (...) reduzieren“ (Busch 2018: 62). *Algorithmic Accountability* bedeutet zum einen, dass die Funktionsweise des ADM-Prozesses erklärbar ist und die Gründe für die Entscheidung des Systems nachvollziehbar sind. Zum anderen kann der Begriff auch die automatische Berücksichtigung von Rechtmäßigkeitsanforderungen bei der Entwicklung eines ADM-Prozesses beinhalten (*legality by design*) (Busch 2018: 60ff.; vzbv 2017: 12). *Ethics by design*¹⁶ geht dahingehend weiter, dass ethische Standards im gesamten ADM-Prozess gefordert werden bzw. eine Berufsethik für die Entwickler*innen herausgearbeitet werden soll (vzbv 2017: 12; Busch 2018: 62; Zweig 2018: 30). So hat beispielsweise die Ethikkommission des Bundesverkehrsministeriums Leitlinien für den Einsatz von Algorithmen beim automatisierten Fahren entwickelt (vzbv 2017: 12). Zudem hat die Bundesregierung eine Datenethikkommission ins Leben gerufen, die bis Herbst 2019 ethische Leitlinien und Regulierungsempfehlungen für den Einsatz von Algorithmen und Künstlicher Intelligenz entwickeln und vorschlagen soll (BMI 2018). Auf supranationaler Ebene hat die

¹⁵ Dieses Konzept wurde in Artikel 25 DSGVO festgehalten (Hacker 2018a: 31).

¹⁶ Für eine ausführliche Übersicht zur Diskussion um *ethics by design* siehe Jaume-Palasi/Spielkamp (2017).

Europäische Kommission eine hochrangige Expert*innengruppe für Künstliche Intelligenz¹⁷ einberufen, damit diese bei Gesetzgebungsverfahren zu dieser Thematik beratend eingreifen kann (European Commission 2019).

(4) Die Forderungen nach Kontrollen, insbesondere für sensible, teilhaberelevante ADM-Prozesse, lassen sich grob in allumfassende Zulassungskontrollen von ADM-Prozessen sowie in verschiedene Einzelmaßnahmen an unterschiedlichen Zeitpunkten der Entwicklungsphasen unterscheiden. Präventive Zulassungskontrollen wie ein Algorithmen-TÜV müssten durch eine unabhängige Institution verabschiedet werden. Zweig fordert, dass „[i]nsbesondere wenn Algorithmen menschliche Experten in gesellschaftlich relevanten Prozessen ersetzen, (...) ein experimentell überprüfbarer Beweis vorliegen [muss], dass die algorithmisch unterstützte Lösung den gesellschaftlich relevanten Prozess verbessert“ (Zweig 2018: 32). Auch die Berliner Beauftragte für Datenschutz und Informationsfreiheit (2018b) fordert u.a. einen Algorithmen-TÜV (ebd.: 5). Angelehnt an die Forderungen der Datenschutzbeauftragten könnten auch betrieblich verbindliche Algorithmenbeauftragte zur Kontrolle von ADM-Prozessen eingesetzt werden. Zertifizierungen nach einheitlichen Standards bzw. Normen, wie bspw. bei DIN oder ISO, hätten ähnlich wie Gütesiegel eher selbstregulierende Wirkungen (vzbv 2017: 8; Busch 2018: 63ff.; Zweig 2018: 29f.).

Ein *one-fits-all*-Ansatz für ADM-Prozesse wird jedoch als problematisch angesehen, da die Systeme zu komplex sind und in zu vielen verschiedenen Lebensbereichen eingesetzt werden (Jaume-Palasi 2017). So gibt es bereits in einigen Bereichen wie Gesundheit und Finanzen Kontrollinstanzen wie das Bundesinstitut für Arzneimittel und Medizinprodukte oder die Bundesanstalt für Finanzmarktaufsicht. Jaume-Palasi (2017) spricht sich daher eher dafür aus, die Kompetenzen und Regelungen

dieser Institutionen entsprechend des jeweiligen Einsatzbereichs von ADM-Prozessen zu stärken und zu erweitern. Einzelne Maßnahmen sind u.a. Monitoring-Instrumente, die den Input auf Richtigkeit, Vollständigkeit und Fairness überprüfen, oder Black-Box-Experimente, die den Output des ADM-Prozesses systematisch mit Daten testen, bei denen das richtige Ergebnis bekannt ist (Zweig 2018: 30f.). Shadowen (2017) fordert, dass bei selbstlernenden Systemen regelmäßige Prüfungen auf Korrektheit und Fairness der Ergebnisse durchgeführt werden sollen (Shadowen 2017: 17).

Zur Beurteilung, welche Maßnahme für welchen ADM-Prozess angemessen ist, fordert die Verbraucherzentrale Bundesverband zunächst eine „Prüfung der Relevanz des ADM-Prozesses“ (vzbv 2017: 11). Auch Zweig konzentriert sich in ihrer Analyse auf „teilhaberelevante“ Algorithmen, „die Entscheidungen unterstützen oder treffen [und] eine Auswirkung auf das Leben von Personen haben“ (Zweig 2019: 13). In dem Projekt Algo.Rules wurden neun Regeln für die Gestaltung von ADM-Prozessen entwickelt, die sich an die verschiedenen Akteur*innen richten, die an der Entstehung beteiligt sind. Dieses zivilgesellschaftliche Projekt kann als eine Entwicklung eines Berufskodex gedeutet werden und beinhaltet einige der in diesem Kapitel genannten Forderungen (iRights.Lab/Bertelsmann Stiftung 2019) (s. Anhang B).

Zusammenfassend fordert die Berliner Beauftragte für Datenschutz und Informationsfreiheit (2018b), dass „[s]owohl der Gesetzgeber als auch die öffentlichen Stellen (...) vor dem Hintergrund der Grundrechtsbindung der Verwaltung dafür Sorge tragen [müssen], dass (...) Maßstäbe für den öffentlichen Bereich verbindlich festgelegt und umgesetzt werden. Darüber hinaus ist der Gesetzgeber dazu angehalten, entsprechende Vorgaben auch für die Privatwirtschaft zu normieren“ (ebd.: 5).

Übersicht über Handlungsforderungen

(1) Informationen

- Kennzeichnungspflicht von ADM-Prozessen; ‚Beipackzettel‘ und Schulungen für Anwender*innen

(2) Transparenz

- Offenlegung von Codes und Trainingsdaten und dadurch Ermöglichung der Beforschung von Algorithmen

(3) Anforderungen zur technischen Gestaltung

- Algorithmische Rechenschaftspflicht; ethische Standards; Berücksichtigung von Rechtmäßigkeitsanforderungen

(4) Kontrollen

- Präventive Zulassungskontrollen von Algorithmen: Algorithmen-TÜV
- Betriebliche Algorithmenbeauftragte; externe Algorithmenbeauftragte; Zertifizierungen

¹⁷ Übersetzung durch die Autorin.

3. Relevanz von ADM-Prozessen für das Land Berlin

Für den öffentlichen Dienst des Landes Berlin gewinnt das Thema Digitalisierung zunehmend an Bedeutung. In den Richtlinien der Regierungspolitik 2016–2021 heißt es in dem Kapitel ‚Leistungsfähige Verwaltung für eine funktionierende Stadt‘, dass die „Digitalisierung als Antrieb der Modernisierung“ genutzt werden soll und dabei „der Einsatz moderner Informations- und Kommunikationstechnik zur Beschleunigung der Verwaltungsabläufe“ für Bürger*innendienste und interne Verwaltungsabläufe „[w]ichtigstes Ziel des Senats ist“ (Senatskanzlei 2017). Als gesetzliche Grundlage dient das Berliner E-Government-Gesetz (eGovG), das nach § 2(1) zum Ziel hat, „die Verwaltungsverfahren und -strukturen aller Verwaltungsebenen und -bereiche der Berliner Verwaltung unter Nutzung der Möglichkeiten der Informations- und Kommunikationstechnik auf E-Government umzustellen“. Dadurch sollen unter anderem Transparenz, Partizipationsmöglichkeiten und Wirtschaftlichkeit gewährleistet werden.

Wie vielfältig die Vorhaben rund um das Thema Digitalisierung im Land Berlin sind, zeigen die Richtlinien der Regierungspolitik, in denen dieses Stichwort in verschiedenen Politikfeldern erwähnt wird: In der Entwicklung einer Digitalisierungsstrategie für die Wirtschaft, im Ausbau der digitalen Infrastruktur, in der Weiterentwicklung der *Smart City*-Strategie, in der Stadtentwicklung, im Verkehr und in der Mobilität, in der Kultur und Medienpolitik, im Schulunterricht und an Hochschulen.

Wenngleich der Einsatz von ADM-Prozessen in den Regierungsrichtlinien nicht erwähnt wird, so ist davon auszugehen, dass dieser in gesteigertem Maße stattfindet wird. Bisher steht die Nutzung von ADM-Prozessen in der Berliner Verwaltung jedoch noch am Anfang (Senatsverwaltung für Inneres und Sport 2019b). Die Berliner Beauftragte für Datenschutz und Informationsfreiheit (2018b) verfasste mit weiteren Datenschutzbeauftragten ein Positionspapier mit dem Titel ‚Transparenz der Verwaltung beim Einsatz von Algorithmen für gelebten Grundrechtsschutz unabdingbar‘, das die Risiken von ADM-Prozessen in (internen) Verwaltungsabläufen thematisiert und konkrete Forderungen nach einer Minimierung derselben formuliert. Neben grundsätzlichen Ansprüchen wie Transparenz, Rechtskonformität und Nachvollziehbarkeit der Entscheidungen, die durch ADM-Prozesse generiert werden, wird das Diskriminierungsverbot und die Achtung der Menschenwürde als grundlegend erachtet. Daher wird unter anderem gefordert, dass sich beim Einsatz von ADM-Prozessen „unter keinen Umständen eine diskriminierende Wirkung“ entfalten darf. Daraus ergeben sich „hohe Anforderungen an Auswahl und Entwicklung von Algorithmen und KI-Verfahren, beispielsweise bei der Auswahl von Trainingsdaten für selbstlernende Systeme“ (ebd.: 4).

- Die Digitalisierung dient der Verwaltungsmodernisierung im Land Berlin und ist im Berliner E-Government-Gesetz (EGovG) geregelt.
- Ziel ist die Gewährleistung von Transparenz, Partizipationsmöglichkeit und Wirtschaftlichkeit.
- Der Berliner Beauftragten für Datenschutz und Informationsfreiheit sind die Risiken von ADM-Prozessen in der Verwaltung bewusst.

3.1. Einsatz von ADM-Prozessen in der Berliner Verwaltung

Im Folgenden werden einige verwaltungsinterne¹⁸ Einsatzorte sowie verwaltungsexterne¹⁹ Lebensbereiche bzw. Politikfelder im Land Berlin umrissen, bei denen der Einsatz von ADM-Prozessen möglich ist und ein erhöhtes Diskriminierungsrisiko darstellt.

Arbeitsmarktzugang

Im Bereich des Arbeitsmarktzugangs ergeben sich durch die Digitalisierung neue Chancen, aber auch Diskriminierungsrisiken. ADM-Prozesse spielen daher auch im Personalwesen²⁰ eine wichtige Rolle. Vor dem Hintergrund des Fachkräftemangels werden ADM-Prozesse zunehmend auch in deutschen Firmen zur Optimierung der Personalgewinnung in Bewerbungsverfahren eingesetzt. Diese Entwicklung gibt es in den USA schon länger, dementsprechend ist die Datenlage dort vergleichsweise besser als in Deutschland. Im Jahr 2016 nutzten 6 Prozent der 1.000 größten Unternehmen in Deutschland computergesteuerte Auswahlverfahren – Tendenz steigend. Zugleich gaben vier von zehn Bewerber*innen an, bei der Stellensuche an automatisierten Verfahren teilgenommen zu haben (Lischka/Klingel 2017: 22, 24).

Insbesondere im Dienstleistungssektor werden zunehmend automatisierte Vorauswahlen von Bewerber*innen durch Online-Persönlichkeitstests durchgeführt. Durch Skalenfragen werden Bewerber*innen in Persönlichkeitsmodelle einsortiert und somit als passend oder weniger passend eingestuft. Noch bevor die Personalere*innen die Bewerbungen gesehen haben, wird ein von dem Unternehmen bestimmter Prozentsatz automatisch aussortiert. Der Einsatz, die Bewertungen und die Reichweite dieser Form von ADM-Prozessen sind intransparent und es gibt Hinweise auf Benachteiligung. Ein weiteres Problem ist, dass es für Personen, die aussortiert werden, mitunter nicht nur um eine einzelne Stelle geht, sondern um den grundsätzlichen Zugang zum Arbeitsmarkt. Die Auswahl an entsprechender Software ist jedoch überschaubar. Viele Unterneh-

men nutzen daher den gleichen Persönlichkeitstest. Dies führt dazu, dass „die Software zum Torwächter nicht für einen[sic], sondern für die Mehrheit potenzieller Stellen wird“ (Lischka/Klingel 2017: 22).

Lischka und Klingel (2017) gehen davon aus, dass durch diese Tests gängige Formen von Diskriminierung auf dem Arbeitsmarkt, wie zum Beispiel ein nicht als deutsch interpretierter Name, durch andere Formen der Diskriminierung ersetzt werden können, wie ein Beispiel aus den USA zeigt. Ein Student mit einer erfolgreich abgeschlossenen Therapie seiner bipolaren Störung war nicht in der Lage, einen Nebenjob im Niedriglohnssektor zu finden, da er immer vom gleichen System aussortiert wurde. Gleichzeitig besteht aber auch die Möglichkeit, dass bestehende Formen der Diskriminierung weitergeschrieben werden, wenn der Algorithmus anhand der bisher genutzten Daten trainiert wird. Da diese Systeme vor allem im Niedriglohnssektor eingesetzt werden, ist das Ziel die effizientere Auswahl von passendem Personal und nicht etwa eine Bestenauslese. Folglich haben die Unternehmen wenig Interesse an einer Verbesserung der Online-Persönlichkeitstests, solange sie ihren Zweck erfüllen. Wenn jedoch die Testergebnisse mehr gewichtet werden als die formelle Qualifikation, besteht auch die Möglichkeit, Niedrigqualifizierten mehr Zugang zum Arbeitsmarkt zu gewähren (Lischka/Klingel 2017: 22ff.).

Ein weiterer Einsatz von ADM-Prozessen im Bewerbungsprozess findet durch *Matching*-Algorithmen statt, die die formellen Anforderungen einer Stelle mit dem Bewerber*innenprofil abgleichen und eine entsprechende Vorauswahl für die Personalere*innen treffen (Busch 2018: 13). *Matching*-Algorithmen werden zudem von Unternehmen eingesetzt, um die kulturelle Übereinstimmung der Bewerber*innen mit der Unternehmenskultur abzugleichen. Diese sogenannten „psychometrische[n] Algorithmen“ werden bisweilen auch schon vor der Überprüfung der formellen Kriterien eingesetzt (Busch 2018: 14). So geben Online-Karriereportale mittlerweile Tipps, welche Schlagwörter in eine Bewerbung einfließen sollten, um nicht direkt vom Algorithmus aussortiert zu werden (Zeitler o.J.).

18 Weitere mögliche Einsatzfelder von ADM-Prozessen im öffentlichen Dienst sind Gesichtserkennungssoftwares in Sicherheitssystemen, wie am Südkreuz von der Bundespolizei getestet und vielfach kritisiert (Berliner Beauftragte für Datenschutz und Informationsfreiheit 2018a) oder auch Lernsoftwares im Bildungsbereich (Aupperle 2018).

19 Ein weiteres Einsatzgebiet, das viele Möglichkeiten, aber auch Risiken birgt, ist das Gesundheitswesen. Der Einsatz von medizinischen Apps zur Therapie und die Auswertung von *Big-Gesundheits-Data* zur Diagnostik schreitet voran (Busch 2018: 15f.).

20 Im Personalwesen gibt es auch außerhalb des Bewerbungsverfahrens Anwendungsbeispiele, die ADM-Prozesse einsetzen. So werden in den USA Lehrkräfte durch einen ADM-Prozess detailgenau nach ihren Leistungen bewertet und die US-amerikanische Bank JP Morgan Chase plant eine automatisierte Auswertung der digitalen Kommunikation ihrer Mitarbeiter*innen zur Betrugsprävention (Busch 2018: 15).

Auch Online-Karriereportale setzen diese Form von Algorithmen ein, um Bewerber*innen den Service zu bieten, vor einer zeitintensiven Bewerbung herauszufinden, ob sie zu der entsprechenden Unternehmenskultur passen. Bei diesen Formen des Abgleichs potenzieller neuer Mitarbeiter*innen werden solche gefunden, die zu dem bisherigen Personal des Unternehmens passen. Unkonventionellere Personen, die die Innovation eines Unternehmens stärken oder fehlende Eigenschaften komplementieren können, werden dadurch schon im Vorhinein aussortiert oder bewerben sich erst gar nicht (Busch 2018: 14).

Des Weiteren werden Such- und Filteralgorithmen vereinzelt von Unternehmen eingesetzt, um passendes Personal auf Online-Karriereportalen gezielt und direkt anzusprechen – auch *Headhunting 2.0* genannt. Diese sogenannten *Active-Sourcing*-Anwendungen bzw. automatisierten Empfehlungssysteme greifen direkt auf Online-Profile zu und gleichen diese durch die beschriebenen *Matching*-Algorithmen mit den entsprechenden formellen oder kulturellen Vorgaben des Unternehmens ab. Dazu müssen die Nutzer*innen dieser Plattformen jedoch vorher einwilligen, dass ihre Profile entsprechend gescannt werden dürfen (ebd.). Insbesondere im IT-Sektor ist diese Form der Personalakquise mittlerweile gängig (ADS 2018: 30).

Zwar geht Busch davon aus, dass ADM-Prozesse in „den meisten Fällen (...) der Bewerbervorauswahl [dienen], die abschließende Einstellungsentscheidung [aber] [...] in der Hand der menschlichen Entscheider [verbleibt]“ (Busch 2018: 13). Trotzdem sei an dieser Stelle darauf hingewiesen, dass diese Formen der Bewerber*innenauswahl die Antidiskriminierungsarbeit im Handlungsfeld Arbeitsmarkt vor neuen Herausforderungen stellt, insbesondere da es für Deutschland keine unabhängigen Studien zur Diskriminierung durch ADM-Prozesse auf dem Arbeitsmarkt gibt.

Als Diskriminierungsrisiko kann hier grundsätzlich benannt werden, dass ADM-Prozesse möglicherweise Personen benachteiligen, denen vom AGG geschützte Merkmale zugeschrieben werden, da beispielsweise die Trainingsdaten verzerrt sind oder sich die Gewichtung gewisser Persönlichkeitsmerkmale im ADM-Prozess nachteilig auswirkt. Fraglich ist auch, inwiefern die Ergebnisse im Nachhinein von Personaler*innen genutzt und kontrolliert sowie ggf. korrigiert werden können.

Ein weiteres Diskriminierungsrisiko im Bewerbungsverfahren ist die Stellenausschreibung als Werbung auf sozialen Medien für nur eine bestimmte Altersgruppe oder ein Ge-

schlecht. Gleichwohl besteht auch die Chance, Bewerbungsverfahren durch ADM-Prozesse fairer zu gestalten, indem die Vorauswahl nur nach Kompetenzen bzw. *skills* erfolgt (Ciesielski 2017).

Im Land Berlin sollen Bewerbungsverfahren zunehmend durch digitale *Recruiting*-Methoden abgewickelt werden. Unter Federführung der Senatsverwaltung für Finanzen soll landesweit das *e-Recruiting*-Verfahren rexx eingesetzt werden (Senatsverwaltung für Finanzen 2018). Inwiefern in rexx *decision support systems* oder sogar *automated decision-making systems* eingebettet sind und Diskriminierungspotenziale in den ADM-Prozessen stecken, ist zu überprüfen. Zudem ist zu klären, inwiefern Online-Persönlichkeitstests zum Einsatz kommen bzw. in Zukunft eingesetzt werden sollen. Eine Untersuchung auf diskriminierende Auswirkungen ist ebenfalls naheliegend.

Legal Tech

Unter *Legal Tech* wird allgemein die Digitalisierung im juristischen Bereich verstanden. Um Kosten einzusparen und die Effizienz zu steigern, können rechtliche Arbeitsprozesse durch ADM-Prozesse unterstützt werden oder sogar automatisiert ablaufen (*legal-tech* o.J.). Die Anwendungsbereiche sind vielfältig: Anwält*innen können den Erfolg eines Rechtsstreits im Vorhinein berechnen lassen und Richter*innen können bei der Rechtsprechung unterstützt werden, beispielsweise indem ein Prüfprogramm die Richtigkeit von ALG II-Bescheiden kontrolliert (Suliak 2018a, 2018b).

Die Entwicklungen in Deutschland stehen dazu noch am Anfang (Forschungsstelle Legal Tech 2019), weshalb ein konkretes Diskriminierungsrisiko in diesem Bereich derzeit nicht benannt werden kann. Es ist davon auszugehen, dass der Einsatz von ADM-Prozessen in deutschen Gerichten noch nicht stattfindet. Im nicht-justiziellen Bereich gibt es aber bereits Anwendungen, wie *Ratisbot* – der erste deutsche Roboter-Anwalt, der bei juristischen Problemen Fragen beantwortet (Ratis 2019).

Mit der Forschungsstelle Legal Tech an der Humboldt-Universität zu Berlin ist die erste wissenschaftliche Einrichtung entstanden, die sich in Deutschland mit den Auswirkungen und der Bedeutung des Einzugs von Algorithmen für die Justiz auseinandersetzt (Forschungsstelle Legal Tech 2019).²¹ Zudem wird in einer länderübergreifenden Arbeitsgruppe ein Bericht verfasst, der die Entwicklungen in diesem Bereich dis-

21 Es findet jährlich in Berlin eine internationale Konferenz zum Thema Legal Tech statt (Berlin Legal Tech 2019).

kutiert und eine rechtspolitische Einschätzung zum Einsatz von Algorithmen ermöglichen soll (88. Konferenz der Justizministerinnen & Justizminister 2017).

Predictive Policing

Beim schon aufgegriffenen *Predictive Policing* soll die Polizeiarbeit unterstützt werden, indem Softwares zukünftige Verbrechen wie beispielweise Einbrüche vorhersagen. Dies geschieht durch die Einspeisung von Kriminalitätsstatistiken und soziodemographischen Daten in einen Algorithmus, der die Wahrscheinlichkeit einer neuen Tat in einem Gebiet errechnet. Dort könnte die Polizei dann vermehrt Streife fahren und beispielsweise Anwohner*innen informieren (Diehl/Kartheuser 2018).

Das Diskriminierungsrisiko wurde durch das Beispiel der Berechnung von Rückfallquoten in den USA verdeutlicht. Bisher wird in Deutschland jedoch anders als in den USA nicht mit persönlichen Daten von Personen gearbeitet. Die Quellcodes der Software-Lösungen, die in Deutschland eingesetzt werden, werden jedoch nicht offen gelegt. Dadurch wird eine Kontrolle hinsichtlich problematischer Ergebnisse nicht möglich und verzerrte Datenbanken können nicht aufgedeckt werden. Zudem können sich gesellschaftliche Nachteile durch den Einsatz solcher algorithmischen Vorhersagen ergeben, indem aufgrund von vermehrten Kontrollen das Gefühl unter den Anwohner*innen entsteht, in einer gefährlichen Gegend zu wohnen (Kartheuser 2018).

Bei der Berliner Polizei läuft derzeit eine zweijährige Testphase mit der von Microsoft entwickelten Software KrimPro zur Prognose von Wohnraumeinbrüchen. Der Einsatz wird positiv bewertet, da die Software „als Ergänzung und ggf. auch Weiterentwicklung [einen positiven Beitrag zu] bereits vorhandene[n] polizeiliche[n] Bekämpfungsansätze[n] polizeilicher Bekämpfungsansätze und -strategien leisten kann“ (Senatsverwaltung für Inneres und Sport 2019a: 1). Da das Programm für die Berechnung von Eintrittswahrscheinlichkeiten von Einbrüchen in bestimmten Gebieten eingesetzt wird, ist die Evaluation personenbezogener Daten nicht notwendig. Vielmehr werden „fallbezogene und raumbezogene Informationen“ genutzt (ebd.: 4). Dies soll auch bei der Erweiterung um andere Delikte so verbleiben.

Anhand der öffentlichen Informationen zum Einsatz von KrimPro lassen sich ein konkretes individuelles Diskriminie-

rungsrisiko oder auch ggf. problematische gesellschaftliche Auswirkungen derzeit nicht evaluieren.

Verbraucher*innenschutz

Wie schon mehrfach angedeutet, ergeben sich durch den Einsatz von ADM-Prozessen im Verbraucher*innenschutz neue Handlungsfelder. So haben ADM-Prozesse im Online-Handel für personalisierte Kaufempfehlungen eine wichtige Funktion. Durch den Online-Kauf von Produkten ‚füttern‘ Kund*innen den Algorithmus mit stetig neuen Daten. Der ist wiederum in der Lage, personalisierte Produkte und Werbung vorzuschlagen, die von anderen Kund*innen, die diese Produkte gekauft haben, ebenfalls erworben wurden (*Profiling*). Die Genauigkeit dieser Empfehlungen wird durch ein Beispiel deutlich, bei dem durch den Kauf von Nahrungsergänzungsmitteln in Kombination mit parfümfreier Körperlotion nicht nur die Schwangerschaft einer Kundin, sondern auch der entsprechende errechnete Geburtstermin bestimmt wurde.

Ein weiterer Einsatz im Online-Handel ist die dynamische bzw. automatische Preisanpassung an Angebot und Nachfragesituation (Busch 2018: 12f.). Der Einsatz von *Scoring*-Algorithmen ist bei der Berechnung der Kreditwürdigkeit²² weit verbreitet – zum Beispiel bei der Schufa. Ein Schufa-Nachweis zur eigenen Finanzhistorie ist für die Vergabe von Krediten, aber auch bei der Anmietung von Wohnungen oder anderen Vertragsabschlüssen erforderlich. Ein negativer Schufa-Eintrag kann sich demnach erheblich auf Verbraucher*innen auswirken. Netzpolitische Initiativen kritisieren die intransparente Datensammlung und Auswertung der Bonität von Privatpersonen (Palmetshofer/Semsrott 2018a). Um den verdächtigen Algorithmus kritisch analysieren zu können, bat das Projekt OpenSchufa Verbraucher*innen um Datenspenden zur Auswertung der Schufa-Ergebnisse. Die Analyse ergab, dass der Schufa-Algorithmus erstens grundsätzlich schlechte *Scores* verteilt, obwohl keine negativen Einträge vorliegen. Zweitens macht die Schufa Aussagen zur Kreditwürdigkeit, obwohl die Datenbasis dafür zu gering ist. Drittes wirken sich Merkmale wie Geschlecht, Alter und häufige Umzüge negativ auf den *Score* aus. Junge Männer würden demnach öfter schlechter bewertet. Viertens hängt der *Score* von der Schufa-Version ab, die die jeweilige Bank nutzt. Ältere Versionen des Systems zeigen demnach schlechtere Ergebnisse an (Palmetshofer/Semsrott 2018b). Dies sind allesamt Ein

²² Ein weiteres denkbare Einsatzgebiet von ADM-Prozessen im Verbraucher*innenschutz sind Versicherungen (Vieth/Wagner 2017: 19).

fallstore für Diskriminierungen und den Verlust der eigenen (Verbraucher*innen-)Souveränität, die zwar mit der bisherigen Literatur und dem bisherigen Stand der Forschung nicht zweifelsfrei belegbar, aber trotzdem denkbar sind.

Die Auseinandersetzung mit ADM-Prozessen im Verbraucher*innenschutz erfolgt sowohl in der Zivilgesellschaft (vzbv 2017) als auch im Land Berlin. So fand im Februar 2019 die gemeinsame Veranstaltung zu ‚Chancen und Risiken der Nutzung von Algorithmen – wie können wir die digitale Souveränität der Verbraucher*innen stärken?‘ der Senatsverwaltung für Justiz, Verbraucherschutz und Antidiskriminierung und des Think Tanks iRights.Lab (iRights.Lab o.J. a) statt.

Wohnungsmarkt

Auch in der Wohnungswirtschaft kommt es zu Digitalisierungsprozessen wie beispielsweise *Smart-Home*-Anwendungen, die Sicherheitskontrollen, Heizungsregulierungen oder automatische Einkäufe ermöglichen. In einer Studie aus dem Bereich der Immobilienwirtschaft der Unternehmensberatung Ernst and Young aus dem Jahr 2016 wurde die sogenannte PropTech (*Property Technology*²³) als ein wichtiger Trend in der Branche verzeichnet. Insbesondere

beim Zugang zu Wohnraum werden vermehrt ADM-Prozesse eingesetzt, die für die Verwaltung von Mieten oder den Ankauf von Wohnraum genutzt werden. Junge *Startup*-Unternehmen²⁴ aus Städten mit angespanntem Wohnungsmarkt bieten Softwarelösungen an, die Vermieter*innen bzw. gewerblichen Immobilienfirmen ein automatisiertes Interessensmanagement von potenziellen Mieter*innen ermöglicht. Ziel ist es, die Vermittlungsprozesse effizienter und besser zu gestalten. Ähnlich wie beim Bewerbungsverfahren werden *Matching*-Algorithmen eingesetzt, die eine entsprechende Vorauswahl bzw. Filterung für das Immobilienunternehmen ermöglichen (Rodeck et al. 2016: 30).

Das Diskriminierungsrisiko liegt offensichtlich bei der Anwendung von Kriterien zur Vorauswahl, die nicht AGG-konform sind. Es besteht aber auch die Chance, Diskriminierungen auf dem Wohnungsmarkt entgegenzuwirken, da beispielsweise durch ADM-Prozesse Kriterien wie das Windhundprinzip²⁵ und Einkommen als neutrale Faktoren gewertet werden können. Für das Land Berlin und hier im konkreten für die Fachstelle gegen Diskriminierung auf dem Wohnungsmarkt ‚Fair mieten – Fair wohnen‘ kann bei Testverfahren der Einsatz von *Matching*-Algorithmen mitgedacht werden.

Übersicht der Bereiche, in denen der Einsatz von ADM-Prozessen für das Land Berlin und die Antidiskriminierungsarbeit relevant werden (könnte):

- Im Arbeitsmarktzugang durch das *Matching*-Tool
- Im Bereich *Legal Tech*, dessen Relevanz in Deutschland stetig wächst
- *Predictive Policing* zur Unterstützung der Berliner Polizeiarbeit
- Im Verbraucher*innenschutz durch *Scoring*-Algorithmen und personalisierte Werbung
- Auf dem Wohnungsmarkt durch *PropertyTechnology*

23 Der Begriff bezeichnet den Einsatz digitaler Technologie in der Immobilienwirtschaft.

24 Für eine Übersicht solcher spezifischer Unternehmen siehe Rodeck et. al. 2016: 43f.

25 Auch als *first-come-first-served*-Prinzip bekannt.

3.2. Handlungsmöglichkeiten für das Land Berlin und die LADS

Diese Bestandsaufnahme hat einen ersten Überblick über Risiken algorithmenbasierter Diskriminierung gegeben. Daraus ergeben sich für das Land Berlin und die Landesstelle für Gleichbehandlung – gegen Diskriminierung (LADS) verschiedene Handlungsoptionen und -bedarfe. Konkrete Handlungsbedarfe deuten sich an, können aber nicht abschließend zusammenge-

stellt und formuliert werden, da weitere Forschungserkenntnisse über eine sich schnell entwickelnde Technologie notwendig sind.

Zur Anwendung klassischer Antidiskriminierungsarbeit auf algorithmenbasierte Diskriminierung muss ein Kompetenzaufbau bei den behördlichen Akteur*innen und zivilgesellschaftlichen Beratungsangeboten stattfinden. Folgende behördliche Institutionen im Land Berlin zum Thema Digitalisierung sind dabei wichtig:

Organisation	Aufgaben
Berliner Beauftragte für Datenschutz und Informationsfreiheit	Auf Berliner Landesebene u.a.: <ul style="list-style-type: none"> • Kontrolle und Beratung von öffentlichen Stellen und Behörden • Kontrolle und Beratung von privaten Unternehmen • Informationen für Bürger*innen
Senatsverwaltung für Inneres und Sport	<ul style="list-style-type: none"> • Grundsatzangelegenheiten der Informations- und Kommunikationstechnik • Digitalisierung der Verwaltung und Bürger*innendienst • Angelegenheiten des E-Government-Gesetzes
Senatsverwaltung für Wirtschaft, Energie und Betriebe	<ul style="list-style-type: none"> • Entwicklung der Digitalisierungsstrategie für das Land Berlin
IT Dienstleistungszentrum Berlin (ITDZ)	<ul style="list-style-type: none"> • Kommunales IT-Unternehmen: zentraler Dienstleister für die Digitalisierung der Berliner Verwaltung

Zusätzlich ist eine Sensibilisierung der Berliner Verwaltung für den Einsatz von ADM-Prozessen notwendig. Die Abwägung von Chancen der Effizienzsteigerung darf nicht zu Lasten von Antidiskriminierungsvorgaben gehen. Die geforderte Festlegung von Standards für den öffentlichen Dienst in diesem Bereich muss Antidiskriminierung in jedem Fall mitdenken. Softwares, die ADM-Prozesse beinhalten, müssen nicht nur datenschutzkonform sein, sondern den Anforderungen an Diskriminierungs- und Barrierefreiheit genügen. Eine entsprechende Standardisierung von Ausschreibungen könnte dem entgegenwirken. Dabei sind nicht nur eventuelle direkte,

sondern auch mittelbare Diskriminierungsrisiken mitzudenken sowie Folgenabschätzungen einzufordern.

Obgleich diese Bestandsaufnahme eine defizitorientierte Perspektive auf die Digitalisierung gezeichnet hat, sei an dieser Stelle nochmals auf die Möglichkeiten, insbesondere zum Abbau von Barrieren, hingewiesen. Eine strukturierte Analyse der Chancen, die die Digitalisierung bietet, ist die andere Seite der Medaille und kann verdeutlichen, welche Potenziale für die Antidiskriminierungsarbeit noch auszuschöpfen sind.

- Es besteht eine Notwendigkeit der Sensibilisierung von Öffentlichkeit, Politik und Verwaltung bezüglich der Risiken der algorithmenbasierten Diskriminierung.
- Die Anwendung von klassischer Antidiskriminierungsarbeit auf algorithmenbasierte Diskriminierungsrisiken wird durch einen Kompetenzaufbau ermöglicht.
- Die Software für den öffentlichen Dienst muss antidiskriminierungsrechtlich konform sein.
- Die Chancen der Digitalisierung für die Antidiskriminierungsarbeit erfordern eine weitere Analyse.

4. Initiativen zur Auseinandersetzung mit ADM-Prozessen

Die folgende Übersicht bietet einen Einblick in zivilgesellschaftliche Initiativen und Einrichtungen sowie Projekte, die an der Schnittstelle Ethik und ADM-Prozesse im deutschsprachen

chigen Raum arbeiten und forschen. Eine Auseinandersetzung mit Diskriminierungen erfolgt mitunter nicht explizit; oft werden allgemeinere ethische Fragen verhandelt.

Zivilgesellschaftliche Initiativen:

Organisation	Ansatz
AlgorithmWatch	„AlgorithmWatch ist eine gemeinnützige Organisation mit dem Ziel, Prozesse algorithmischer Entscheidungsfindung zu betrachten und einzuordnen, die eine gesellschaftliche Relevanz haben – die also entweder menschliche Entscheidungen vorhersagen oder vorbestimmen, oder Entscheidungen automatisiert treffen“ (AlgorithmWatch 2019a). Die Organisation hat eine Datenbank zu Akteur*innen, Richtlinien etc. in Bezug auf ADM-Prozesse und gesellschaftliche Teilhabe erstellt: https://atlas.algorithmwatch.org/datenbank
iRights.Lab Projekt ‚Algo.Rules – Regeln für die Gestaltung algorithmischer Systeme‘ (in Kooperation mit der Bertelsmann Stiftung)	<i>Think Tank</i> für die digitale Welt: „Wir unterstützen öffentliche Einrichtungen, Stiftungen, Unternehmen, Wissenschaft und Politik dabei, die Herausforderungen der Digitalisierung zu meistern und die vielschichtigen Potenziale effektiv und positiv zu nutzen. Dazu verknüpfen wir rechtliche, technische, ökonomische und gesellschaftspolitische Expertise.“ (iRights.Lab o.J. b)
Verbraucherzentrale Bundesverband e.V.	Forderungen nach Transparenz, Nachvollziehbarkeit und Kontrolle von ADM-Prozessen (Verbraucherzentrale Bundesverband e.V. o.J.).
re:publica	Größte Konferenz zu den Themen Internet und (digitale) Gesellschaft in Europa. Findet einmal jährlich in Berlin statt (re:publica o.J.).

Forschungseinrichtungen und -projekte:

Organisation	Ansatz
ABIDA (Assessing Big Data)	„Das interdisziplinäre Projekt ABIDA (Assessing Big Data), gefördert vom Bundesministerium für Bildung und Forschung, lotet gesellschaftliche Chancen und Risiken der Erzeugung, Verknüpfung und Auswertung großer Datenmengen aus und entwirft Handlungsoptionen für Politik, Forschung und Entwicklung.“ (ABIDA o.J.)
Bertelsmann Stiftung Projekt ‚Ethik der Algorithmen‘	„Im Projekt ‚Ethik der Algorithmen‘ setzen wir uns mit den gesellschaftlichen Folgen algorithmischer Entscheidungsfindung auseinander. Wir wollen zu einer Gestaltung algorithmischer Systeme beitragen, die zu mehr Teilhabe für alle führt. Nicht das technisch Mögliche, sondern das gesellschaftlich Sinnvolle muss Leitbild sein“ (Bertelsmann Stiftung 2019). Weiteres unter dem Blog: https://algorithmenethik.de/
Forschungsstelle Legal Tech (HU Berlin)	Forschungsstelle zu den Auswirkungen und der Bedeutung von Legal Tech in den Justizwissenschaften (Forschungsstelle Legal Tech 2019)
Hasso-Plattner-Institut Fachgebiet Algorithm Engineering	Theoretische Grundlagenforschung zu Design und Analyse von Algorithmen (Hasso-Plattner-Institut 2019)
TU Kaiserslautern Algorithm Accountability Lab	Die Arbeitsgruppe erforscht komplexe Netzwerke, um die Funktionsweise von Algorithmen besser nachvollziehen zu können (Algorithm Accountability Lab o.J.).
TU München Institute for Ethics in Artificial Intelligence	Die TU München gründet 2019 das Institute for Ethics in Artificial Intelligence zur Erforschung von ethischen Grundsätzen in der Entwicklung von Künstlicher Intelligenz. Das Institut wird von Facebook finanziert (Dornis 2019).
Weizenbaum-Institut für die vernetzte Gesellschaft	„Die Aufgabe des Weizenbaum-Instituts für die vernetzte Gesellschaft ist es, aktuelle gesellschaftliche Veränderungen, die sich im Zusammenhang mit der Digitalisierung abzeichnen, zu untersuchen und künftige politische und wirtschaftliche Handlungsoptionen zu skizzieren“ (Weizenbaum-Institut o.J.).
Wissenschaftszentrum Berlin für Sozialforschung (WZB) Digitalisierung und gesellschaftlicher Wandel	„Die Forschungsgruppe ‚Politik der Digitalisierung‘ (POLDI) erforscht die Deutung, Aushandlung und Regelung der digitalen Transformation. Ihr Ziel ist es, die Strategien, Kompetenzen und Praktiken der Gestaltung der Digitalisierung und ihrer gesellschaftlichen Folgen zu untersuchen und zu bewerten“ (WBZ o.J.).
Fraunhofer-Institut für Offene Kommunikationssysteme Kompetenzzentrum Öffentliche IT	„Das Kompetenzzentrum Öffentliche IT (ÖFIT) versteht sich als Ansprechpartner und Denkfabrik für Fragen der öffentlichen IT, mit besonderem Augenmerk auf gesellschaftliche Trends und den sich daraus ergebenden Handlungsfeldern. (...) Unter öffentlicher IT versteht man Informationstechnologien, die in einem öffentlichen Raum durch die gesamtgesellschaftliche Relevanz unter besonderer Berücksichtigung der staatlichen Verantwortung stehen“ (Fraunhofer-Institut für Offene Kommunikationssysteme o.J.).

5. Literatur

88. Konferenz der Justizministerinnen & Justizminister (2017): Herbstkonferenz am 9. November 2017. Beschluss. TOP I.9 Legal Tech: Herausforderungen für die Justiz. Online: https://www.justiz.nrw.de/JM/jumiko/beschluesse/2017/Herbstkonferenz-2017/TOP-I_-9.pdf
(Zugriff: 07.03.2019)

ABIDA (o.J.): ABIDA – Das Projekt. Online: <http://abida.de/de/content/abida-das-projekt>
(Zugriff 29.01.19)

Algorithm Accountability Lab (o.J.): Forschung. Online: <http://aalab.informatik.uni-kl.de/forschung/>
(Zugriff: 29.01.2019).

AlgorithmWatch (2019a): Was wir tun. Online: <https://algorithmwatch.org/was-wir-tun/>
(Zugriff: 29.01.2019)

AlgorithmWatch (2019b): Datenbank. Atlas der Automatisierung. Automatisierte Entscheidungen und Teilhabe in Deutschland. Online: <https://atlas.algorithmwatch.org/datenbank>
(Zugriff: 20.06.2019)

Antidiskriminierungsstelle des Bundes (2018): Divers, fair und digital: Personalauswahl diskriminierungsfrei gestalten. Dokumentation der Fachtagung im Quadriga Forum, Berlin am 10. Oktober 2018. Online: http://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/publikationen/Dokumentationen/Dokumentation_FT_Digital_Recruiting.pdf?__blob=publicationFile&v=3
(Zugriff: 22.02.2019)

ARD/ZDF (2019): ARD/ZDF-Onlinestudie 2018. Online: <http://www.ard-zdf-onlinestudie.de/ardzdf-onlinestudie-2018/>
(Zugriff: 20.02.2019)

Aupperle, Astrid (2017): Integration durch Sprache: Schlaumäuse-Sprachlernprogramm für Berliner Grundschulen. In: Microsoft Berlin. Online: <https://www.microsoft-berlin.de/artikel/integration-durch-sprache-schlaumause-sprachlernprogramm-fur-berliner-grundschulen>
(Zugriff: 19.02.2019)

Baldauf, Johannes et al. (2018): Einleitung. In: Dies. (Hrsg.): Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht. Online Civil Courage Initiative. Online: <https://www.isdglobal.org/wp-content/uploads/2018/09/ISD-NetzDG-Report-German-FINAL-26.9.18.pdf>
(Zugriff: 20.02.2019)

Baldauf, Johannes/Guhl, Jakob (2018): Lösungsvorschläge: „Hassrede und Extremismus im Kontext des NetzDG – Empfehlungen an Politik, soziale Netzwerke und Zivilgesellschaft“. In: Baldauf, Johannes/Ebner, Julia/Guhl, Jakob (Hrsg.): Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht. Online Civil Courage Initiative. Online: <https://www.isdglobal.org/wp-content/uploads/2018/09/ISD-NetzDG-Report-German-FINAL-26.9.18.pdf>
(Zugriff: 20.02.2019)

Ley, Hannes/Böckem, Jörg (2019): „Politiker sollen endlich begreifen, wie das Netz funktioniert“. In: Zeit Magazin. Online: <https://www.zeit.de/zeit-magazin/2019/05/hannes-ley-unternehmensberater-ichbinhier-hashtag-traum>
(Zugriff: 20.02.2019)

Berlin Legal Tech (2019): Das Legal Tech Event 2019. Online: <https://berlinlegal.tech/2019/>
(Zugriff: 20.02.2019)

Berliner Beauftragte für Datenschutz und Informationsfreiheit (2018a): Datenschutz und Informationsfreiheit. Jahresbericht 2017. Online: <https://www.datenschutz-berlin.de/infotek-und-service/veroeffentlichungen/jahresberichte/>
(Zugriff: 12.02.2019)

Berliner Beauftragte für Datenschutz und Informationsfreiheit (2018b): Transparenz der Verwaltung beim Einsatz von Algorithmen für gelebten Grundrechtsschutz unabdingbar. Positionspapier im Rahmen der 36. Konferenz der Informationsfreiheitsbeauftragten in Deutschland am 16. Oktober 2018 in Ulm. Online: <https://www.datenschutz-berlin.de/infotek-und-service/veroeffentlichungen/beschluesse-informationsfreiheit/>
(Zugriff: 12.02.2019)

- Bertelsmann Stiftung (2019): Unsere Projekte. Ethik der Algorithmen. Online: <https://www.bertelsmann-stiftung.de/de/unsere-projekte/ethik-der-algorithmen/> (Zugriff: 12.02.2019)
- Biselli, Anna (2017): Syrien oder Ägypten? Software zur Dialektanalyse ist fehleranfällig und intransparent. In: Netzpolitik.org. Online: <https://netzpolitik.org/2017/syrien-oder-aegypten-software-zur-dialektanalyse-ist-fehleranfaellig-und-intransparent/> (Zugriff: 21.02.2019)
- Biselli, Anna (2018a): Eine Software des BAMF bringt Menschen in Gefahr. In: Motherboard. Vice. Online: <https://motherboard.vice.com/de/article/a3q8wj/fluechtlinge-bamf-sprachanalyse-software-entscheidet-asyl> (Zugriff: 21.02.2019)
- Biselli, Anna (2018b): Das BAMF will seine Probleme mit Technik lösen – und macht alles noch schlimmer. In: Netzpolitik.org. Online: <https://netzpolitik.org/2018/das-bamf-will-seine-probleme-mit-technik-loesen-und-macht-alles-noch-schlimmer/> (Zugriff: 21.02.2019)
- Bundesministerium des Inneren, Bau und Heimat (BMI) (2018): Datenethikkommission. Online: <https://www.bmi.bund.de/DE/themen/it-und-digitalpolitik/datenethikkommission/datenethikkommission-node.html> (Zugriff: 07.02.2019)
- Busch, Christoph (2018): Algorithmic Accountability. ABIDA-Assessing Big Data. Online: <http://www.abida.de/sites/default/files/ABIDA%20Gutachten%20Algorithmic%20Accountability.pdf> (Zugriff: 28.01.2019)
- Ciesielski, Rebecca (2017): Digitale Diskriminierung: Wenn Computer Bewerber auswählen. In: Der Tagesspiegel. Online: <https://www.tagesspiegel.de/themen/diversity/recruiting-digitale-diskriminierung-wenn-computer-bewerber-auswaehlen/20583772.html> (Zugriff: 07.03.2019)
- Committee of Ministers (2019): Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes. Decl (13/02/2019)1. Online: https://search.coe.int/cm/pages/result_details.aspx?objectId=090000168092dd4b (Zugriff: 05.03.2019)
- Diehl, Jörg/Kartheuser, Boris (2018): Predictive Policing. Ich weiß, was du heute tun wirst. In: Spiegel Online. Online: <http://www.spiegel.de/panorama/justiz/kriminalitaet-in-deutschland-polizei-setzt-auf-computer-vorhersagen-a-1188350.html> (Zugriff: 28.01.2019)
- Dornis, Valentin (2019): Facebook finanziert Ethik-Institut an der TU München. In: Süddeutsche Zeitung. Online: <https://www.sueddeutsche.de/muenchen/facebook-tu-muenchen-kuenstliche-intelligenz-1.4295434> (Zugriff: 28.01.2019)
- European Commission (2019): High-Level Expert Group on Artificial Intelligence. Online: <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> (Zugriff: 18.02.2019)
- f/21 Büro für Zukunftsfragen (2017): Digitale Diskriminierung. In: f/21 Quarterly, Q3. Online: https://www.f-21.de/downloads/f21_quarterly_q3_2017.pdf (Zugriff: 28.01.2019)
- Forschungsstelle Legal Tech (2019): Die Forschungsstelle Legal Tech. Online: <https://www.forschungsstelle-legal-tech.de/> (Zugriff: 19.02.2019)
- Fraunhofer-Institut für Offene Kommunikationssysteme (o. J.): Kompetenzzentrum Öffentliche IT. Online: https://www.fokus.fraunhofer.de/de/dps/about_oeffit (Zugriff: 29.01.2019)
- Fröhlich, Wiebke/Spiecker, Indra (2018): Können Algorithmen diskriminieren? In: Verfassungsblog. Online: <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/> (Zugriff: 30.01.2019)
- Grzymek, Viktoria/Puntschuh, Michael (2019): Was Europa über Algorithmen weiß und denkt. Ergebnisse einer repräsentativen Bevölkerungsumfrage. Bertelsmann Stiftung (Hrsg). Online: <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WasEuropaUEberAlgorithmenWeissUndDenkt.pdf> (Zugriff: 20.02.2019)
- Hacker, Philipp (2018a): Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law. In: Common Market Law Review. Online: <https://ssrn.com/abstract=3164973> (Zugriff: 30.01.2019)

Hacker, Philipp (2018b): From Algorithmic Discrimination to Algorithmic Fairness. In: Rails Robotics & AI Law Society. Online: <http://ai-laws.org/2018/10/from-algorithmic-discrimination-to-algorithmic-fairness-dr-philipp-hacker-ll-m/> (Zugriff: 30.01.2019)

Hasso-Plattner-Institut (2019): Algorithm Engineering. Online: <https://hpi.de/forschung/fachgebiete/algorithm-engineering.html> (Zugriff: 28.03.2019)

iRights.Lab (o.J. a): Einladung zur Veranstaltung „Chancen und Risiken der Nutzung von Algorithmen – wie können wir die digitale Souveränität der Verbraucher*innen stärken?“. Online: <https://irights-lab.de/chancenrisikenalgorithmen/> (Zugriff: 18.04.2019)

iRights.Lab (o.J. b): Über uns. Online: <https://irights-lab.de/ueber-uns-2/> (Zugriff: 18.02.2019)

iRights.Lab/Bertelsmann Stiftung (2019): Algo.Rules. Regeln für die Gestaltung algorithmischer Systeme. Online: <https://algorules.org> (Zugriff: 13.03.2019)

Jaume-Palásí, Lorena (2017): Diskriminierung hängt nicht vom Medium ab. In: AlgorithmWatch. Online: <https://algorithmwatch.org/de/diskriminierung-haengt-nicht-vom-medium-ab/> (Zugriff: 25.01.2019)

Jaume-Palásí, Lorena/Spielkamp, Matthias (2017): Ethik und algorithmische Prozesse zur Entscheidungsfindung oder –vorbereitung. In: AlgorithmWatch. Arbeitspapier Nr. 4. Online: https://algorithmwatch.org/wp-content/uploads/2017/06/AlgorithmWatch_Arbeitspapier_4_Ethik_und_Algorithmen.pdf (Zugriff: 18.04.2019).

Kartheuser, Boris (2018): Forscherin zu Predictive Policing „Gefahr der Diskriminierung“. Interview mit Lorena Jaume-Palásí. In: Spiegel Online. Online: <http://www.spiegel.de/panorama/justiz/predictive-policing-wo-die-gefahren-in-der-arbeit-mit-der-neuen-software-liegen-a-1189340.html> (Zugriff: 19.02.2019).

Köver, Chris (2019): Gesichtserkennung: Kritik macht Algorithmen genauer, nicht nur für weiße Männer. In: Netzpolitik.org. Online: <https://netzpolitik.org/2019/gesichtserkennung-kritik-macht-algorithmen-genauer-nicht-nur-fuer-weisse-maenner/> (Zugriff: 26.03.2019).

legal-tech (o. J.): Was ist Legal Tech? Online: <https://www.legal-tech.de/was-ist-legal-tech-ffi/> (Zugriff: 29.02.2019).

Lischka, Konrad/Klingel, Anita (2017): Wenn Maschinen Menschen bewerten. Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung. Arbeitspapier. Impuls Algorithmenethik #1. Bertelsmann Stiftung (Hrsg.). Online: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/ADM_Fallstudien.pdf (Zugriff: 28.01.2019).

Maas, Heiko (2017): Rede des Bundesministers der Justiz und für Verbraucherschutz Heiko Maas „Zusammenleben in der digitalen Gesellschaft – Teilhabe ermöglichen, Sicherheit gewährleisten, Freiheit bewahren“. Konferenz „Digitales Leben – Vernetzt. Vermessen. Verkauft? #Werte #Algorithmen #IoT“ am 3. Juli 2017 in Berlin. Online: https://www.bmjv.de/SharedDocs/Reden/DE/2017/07032017_digitales_Leben.html (Zugriff: 25.01.2019).

Martini, Mario (2019): Blackbox Algorithmus. Grundfragen einer Regulierung Künstlicher Intelligenz. Berlin/Heidelberg: Springer. Online: <https://www.springer.com/gp/book/9783662590096> (Zugriff: 11.03.2019).

Mohabbat Kar, Resa et al. (Hrsg.) (2018): (Un)Berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft. Berlin: Kompetenzzentrum Öffentliche IT. Online: <https://www.oeffentliche-it.de/documents/10181/76866/%28Un%29berechenbar+-+Algorithmen+und+Automatisierung+in+Staat+und+Gesellschaft.pdf> (Zugriff: 18.04.2019).

- Moorstedt, Michael (2017): Wie Algorithmen Hass und Vorurteile zementieren. In: Süddeutsche Zeitung. Online: <https://www.sueddeutsche.de/digital/kuenstliche-intelligenz-wie-algorithmen-hass-und-vorurteile-zementieren-1.3620668> (Zugriff: 25.01.2019).
- Neudert, Lisa-Marie (2017): Angriff der Algorithmen – Wahlmanipulation im Internet. Impulsvortrag im Rahmen der FES-Veranstaltung denkraum_medien „(DES)INFORMATION?! Politische Meinungs- & Willensbildung in sozialen Netzwerken“, in Köln am 08. Juni 2017. Friedrich-Ebert-Stiftung. Online: <https://www.fes.de/index.php?eID=dumpFile&t=f&f=15002&token=cffa557f222b899c1454a74ef3ca3a583375de6b> (Zugriff: 25.03.2019).
- Palmeshofer, Walter/Semsrott, Arne (2018a): OpenSchufa. Wir knacken die Schufa. Open Knowledge Foundation Deutschland. Online: <https://okfn.de/projekte/openschufa/> (Zugriff: 21.02.2019).
- Palmeshofer, Walter/Semsrott, Arne (2018b): OpenSchufa: Die Ergebnisse (Updates). Open Knowledge Foundation Deutschland. Online: <https://okfn.de/blog/2018/11/openschufa-ergebnisse/> (Zugriff: 21.02.2019).
- Ratis (2019): RATISBOT – Der erste deutsche Roboter-Anwalt. Online: <https://ratis.de/chatbot/> (Zugriff: 07.03.2019).
- Rebiger, Simon (2019): Die DSGVO zeigt erste Zähne: 50-Millionen-Strafe gegen Google verhängt. In: Netzpolitik.org. Online: <https://netzpolitik.org/2019/die-dsgvo-zeigt-erste-zaehne-50-millionen-strafe-gegen-google-verhaengt/> (Zugriff: 05.02.2019).
- re:publica (o.J.): Über die re:publica GmbH. Online: <https://re-publica.com/de/page/ueber-republica-gmbh> (Zugriff: 29.01.2019).
- Rodeck, Martin et al. (2016): Einsatz digitaler Technologie in der Immobilienwirtschaft. ZIA Zentraler Immobilien Ausschuss e. V./Ernst and Young Real Estate (Hrsg.). Online: https://www.zia-deutschland.de/fileadmin/Redaktion/Positionen/Studie_Digitalisierung_27.09.16.pdf (Zugriff: 07.03.2019).
- Sachverständigenrat für Verbraucherfragen (Hrsg.) (2017): Digitale Souveränität. Gutachten des Sachverständigenrats für Verbraucherfragen. Online: http://www.svr-verbraucherfragen.de/wp-content/uploads/Gutachten_Digitale_Souver%C3%A4nit%C3%A4t_.pdf (Zugriff: 18.04.2019).
- Schallbruch, Martin (2018): Schwacher Staat im Netz. Wie die Digitalisierung den Staat in Frage stellt. Wiesbaden: Springer.
- Senatskanzlei (2017): Richtlinien der Regierungspolitik 2016-2021. Online: <https://www.berlin.de/rbmskzl/regierender-buergermeister/senat/richtlinien-der-politik/> (Zugriff: 12.02.2019).
- Senatsverwaltung für Finanzen (2018): Personalpolitisches Aktionsprogramm 2019/2020. Online: https://www.berlin.de/sen/finanzen/personal/personalmanagement/ppap_2019_2020_public.pdf (Zugriff: 19.02.2019).
- Senatsverwaltung für Inneres und Sport (2019a): Schriftliche Anfrage des Abgeordneten Bernd Schlömer (FDP) zum Thema: Predictive Policing in Berlin – Chancen und Probleme datenanalytischer Prognosetechnik in der Berliner Polizei. Drucksache 18/17 562. Abgeordnetenhaus Berlin. Online: <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/18/SchrAnfr/s18-17562.pdf> (Zugriff: 19.02.2019).
- Senatsverwaltung für Inneres und Sport (2019b): Schriftliche Anfrage des Abgeordneten Bernd Schlömer (FDP) zum Thema: Künstliche Intelligenz in der Berliner Verwaltung. Drucksache 18/18 008. Abgeordnetenhaus Berlin. Online: <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/18/SchrAnfr/S18-18008.pdf> (Zugriff: 26.03.2019).
- Shadowen, Ashley Nicole (2017): Ethics and Bias in Machine Learning: A Technical Study of What Makes Us “Good”. In: CUNY Academic Works. Online: https://academicworks.cuny.edu/jj_etds/44 (Zugriff: 18.04.2019).

Spiegel Online (2019): Digitalpakt für Schulen. Bund und Länder einigen sich auf Grundgesetzänderung. Online: <http://www.spiegel.de/lebenundlernen/schule/digitalpakt-bund-und-laender-einigen-sich-auf-grundgesetzanderung-a-1254310.html> (Zugriff: 05.03.2019).

Suliak, Hasso (2018a): Vorhersehen, wie der Rechtsstreit ausgeht. In: Legal Tribune Online. Online: <https://www.lto.de/recht/zukunft-digitales/l/predictive-analysis-legal-tech-digital-rechtsprechung-anwaelte-rechtsschutzversicherung-rechtsanwender/> (Zugriff: 19.02.2019).

Suliak, Hasso (2018b): Interview: erste Forschungsstelle für Legal Tech „Großer Nachholbedarf im Rechtssystem“. In: Legal Tribune Online. Online: <https://www.lto.de/recht/zukunft-digitales/l/forschungsstelle-legal-tech-berlin-humboldt-universitaet-digitalisierung/> (Zugriff: 19.02.2019).

Tischbirek, Alexander (erscheint vsl. 2019): Artificial Intelligence and Discrimination. In: Rademacher, Timo/Wischmeyer, Thomas (Hrsg.): Regulating Artificial Intelligence. Berlin: Springer.

Ulbricht, Lena (2017): Der aktuelle Gastkommentar. Big Data und Diskriminierung – Impulse aus den USA für eine deutsche Debatte. In: ads aktuell 01 | 2017. Newsletter der Antidiskriminierungsstelle des Bundes vom 15. Februar 2017. Online: https://www.antidiskriminierungsstelle.de/SharedDocs/Downloads/DE/newsletter_pdf/newsletter_01_2017.pdf?__blob=publicationFile&v=2 (Zugriff: 25.01.2019).

Verbraucherschutzministerkonferenz (2018): Ergebnisprotokoll der 14. Verbraucherschutzministerkonferenz am 15. Juni 2018 in Saarbrücken. Online: https://www.verbraucherschutzministerkonferenz.de/documents/14-vsmk-protokoll_final_ext_1531310548.pdf (Zugriff: 07.03.2019).

Vieth, Kilian/Wagner, Ben (2017): Teilhabe, ausgerechnet. Arbeitspapier. Bertelsmann Stiftung (Hrsg.). Online: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/teilhabe-ausgerechnet> (Zugriff: 28.01.2019).

Verbraucherzentrale Bundesverband e.V. (vzbv) (o.J.): Digitale Welt. Online: <https://www.vzbv.de/themen/digitale-welt> (Zugriff: 24.01.2019).

Verbraucherzentrale Bundesverband e.V. (vzbv) (Hrsg.) (2017): Algorithmenbasierte Entscheidungsprozesse. Thesenpapier. Online: https://www.vzbv.de/sites/default/files/downloads/2018/05/22/dm_17-12-07_vzbv_thesenpapier_algorithmen.pdf (Zugriff: 25.01.2019).

Wissenschaftszentrum Berlin für Sozialforschung (WZB) (o.J.): Forschungsgruppe Politik der Digitalisierung. Wissenschaftszentrum Berlin für Sozialforschung. Online: <https://www.wzb.eu/de/forschung/digitalisierung-und-gesellschaftlicher-wandel/politik-der-digitalisierung> (Zugriff: 29.01.2019).

Weizenbaum-Institut für die vernetzte Gesellschaft (o.J.): Das Institut. Online: <https://weizenbaum-institut.de/das-institut/> (Zugriff: 18.04.2019).

Wolfangel, Eva (2018): Algorithmen. Programmierter Rassismus. In: Zeit Online. Online: <https://www.zeit.de/digital/internet/2018-05/algorithmen-rassismus-diskriminierung-daten-vorurteile-alltagsrassismus> (Zugriff: 25.01.2019).

Zeitler, Maria (o.J.): Auswahl per Algorithmus. So wichtig sind Keywords. In: Bewerbung.com. Ein Service von Xing. Online: <https://bewerbung.com/keywords-bewerbung/> (Zugriff: 25.02.2019).

Zweig, Katharina A. (2018): Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung. Arbeitspapier. Bertelsmann Stiftung (Hrsg.). Online: <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WoMaschinenIrrenKoennen.pdf> (Zugriff: 28.01.2019).

Anhang A: Normtext EU-DSGVO (Auszug)

Datenschutzgrundverordnung DSGVO

Art. 5 DSGVO Grundsätze für die Verarbeitung personenbezogener Daten

1. Personenbezogene Daten müssen
 - a.) auf rechtmäßige Weise, nach Treu und Glauben und in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden („Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz“);

Art. 9 DSGVO Verarbeitung besonderer Kategorien personenbezogener Daten

1. Die Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person ist untersagt.
2. Absatz 1 gilt nicht in folgenden Fällen:
 - a.) Die betroffene Person hat in die Verarbeitung der genannten personenbezogenen Daten für einen oder mehrere festgelegte Zwecke ausdrücklich eingewilligt, es sei denn, nach Unionsrecht oder dem Recht der Mitgliedstaaten kann das Verbot nach Absatz 1 durch die Einwilligung der betroffenen Person nicht aufgehoben werden,

Art. 12 DSGVO Transparente Information, Kommunikation und Modalitäten für die Ausübung der Rechte der betroffenen Person

1. Der Verantwortliche trifft geeignete Maßnahmen, um der betroffenen Person alle Informationen gemäß den Artikeln 13 und 14 und alle Mitteilungen gemäß den Artikeln 15 bis 22 und Artikel 34, die sich auf die Verarbeitung beziehen, in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache zu übermitteln;

Art. 13 DSGVO Transparente Information, Kommunikation und Modalitäten für die Ausübung der Rechte der betroffenen Person

1. Der Verantwortliche trifft geeignete Maßnahmen, um der betroffenen Person alle Informationen gemäß den Artikeln 13 und 14 und alle Mitteilungen gemäß den Artikeln 15 bis 22 und Artikel 34, die sich auf die Verarbeitung beziehen, in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache zu übermitteln:
 - a.) den Namen und die Kontaktdaten des Verantwortlichen sowie gegebenenfalls seines Vertreters;

Art. 14 DSGVO Informationspflicht, wenn die personenbezogenen Daten nicht bei der betroffenen Person erhoben wurden

1. Werden personenbezogene Daten nicht bei der betroffenen Person erhoben, so teilt der Verantwortliche der betroffenen Person Folgendes mit:
 - a.) den Namen und die Kontaktdaten des Verantwortlichen sowie gegebenenfalls seines Vertreters;

Art. 15 DSGVO Auskunftsrecht der betroffenen Person

1. Die betroffene Person hat das Recht, von dem Verantwortlichen eine Bestätigung darüber zu verlangen, ob sie betreffende personenbezogene Daten verarbeitet werden; ist dies der Fall, so hat sie ein Recht auf Auskunft über diese personenbezogenen Daten und auf folgende Informationen:
 - a.) die Verarbeitungszwecke;

Art. 22 DSGVO Automatisierte Entscheidungen im Einzelfall einschließlich Profiling

1. Die betroffene Person hat das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profiling – beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.
2. Absatz 1 gilt nicht, wenn die Entscheidung
 - a.) für den Abschluss oder die Erfüllung eines Vertrags zwischen der betroffenen Person und dem Verantwortlichen erforderlich ist,
 - b.) aufgrund von Rechtsvorschriften der Union oder der Mitgliedstaaten, denen der Verantwortliche unterliegt, zulässig ist und diese Rechtsvorschriften angemessene Maßnahmen zur Wahrung der Rechte und Freiheiten sowie der berechtigten Interessen der betroffenen Person enthalten oder
 - c.) mit ausdrücklicher Einwilligung der betroffenen Person erfolgt.

Art. 35 DSGVO Datenschutz-Folgenabschätzung

1. Hat eine Form der Verarbeitung, insbesondere bei Verwendung neuer Technologien, aufgrund der Art, des Umfangs, der Umstände und der Zwecke der Verarbeitung voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge, so führt der Verantwortliche vorab eine Abschätzung der Folgen der vorgesehenen Verarbeitungsvorgänge für den Schutz personenbezogener Daten durch. Für die Untersuchung mehrerer ähnlicher Verarbeitungsvorgänge mit ähnlich hohen Risiken kann eine einzige Abschätzung vorgenommen werden.
2. Der Verantwortliche holt bei der Durchführung einer Datenschutz-Folgenabschätzung den Rat des Datenschutzbeauftragten, sofern ein solcher benannt wurde, ein.

Art. 58 DSGVO Befugnisse

1. Jede Aufsichtsbehörde verfügt über sämtliche folgenden Untersuchungsbefugnisse, die es ihr gestatten,
 - a.) den Verantwortlichen, den Auftragsverarbeiter und gegebenenfalls den Vertreter des Verantwortlichen oder des Auftragsverarbeiters anzuweisen, alle Informationen bereitzustellen, die für die Erfüllung ihrer Aufgaben erforderlich sind,
 - b.) Untersuchungen in Form von Datenschutzüberprüfungen durchzuführen,
 - c.) eine Überprüfung der nach Artikel 42 Absatz 7 erteilten Zertifizierungen durchzuführen,
 - d.) den Verantwortlichen oder den Auftragsverarbeiter auf einen vermeintlichen Verstoß gegen diese Verordnung hinzuweisen,
 - e.) von dem Verantwortlichen und dem Auftragsverarbeiter Zugang zu allen personenbezogenen Daten und Informationen, die zur Erfüllung ihrer Aufgaben notwendig sind, zu erhalten, gemäß dem Verfahrensrecht der Union oder dem Verfahrensrecht des Mitgliedstaats Zugang zu den Räumlichkeiten, einschließlich aller Datenverarbeitungsanlagen und -geräte, des Verantwortlichen und des Auftragsverarbeiters zu erhalten.

Erwägungsgrund 71 Profiling*

¹ Die betroffene Person sollte das Recht haben, keiner Entscheidung – was eine Maßnahme einschließen kann – zur Bewertung von sie betreffenden persönlichen Aspekten unterworfen zu werden, die ausschließlich auf einer automatisierten Verarbeitung beruht und die rechtliche Wirkung für die betroffene Person entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt, wie die automatische Ablehnung eines Online-Kreditantrags oder Online-Einstellungsverfahrens ohne jegliches menschliche Eingreifen.

² Zu einer derartigen Verarbeitung zählt auch das „Profiling“, das in jeglicher Form automatisierter Verarbeitung personenbezogener Daten unter Bewertung der persönlichen Aspekte in Bezug auf eine natürliche Person besteht, insbesondere zur Analyse oder Prognose von Aspekten bezüglich Arbeitsleistung, wirtschaftliche Lage, Gesundheit, persönliche Vorlieben oder Interessen, Zuverlässigkeit oder Verhalten, Aufenthaltsort oder Ortswechsel der betroffenen Person, soweit dies rechtliche Wirkung für die betroffene Person entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.

³ Eine auf einer derartigen Verarbeitung, einschließlich des Profilings, beruhende Entscheidungsfindung sollte allerdings erlaubt sein, wenn dies nach dem Unionsrecht oder dem Recht der Mitgliedstaaten, dem der für die Verarbeitung Verantwortliche unterliegt, ausdrücklich zulässig ist, auch um im Einklang mit den Vorschriften, Standards und Empfehlungen der Institutionen der Union oder der nationalen Aufsichtsgremien Betrug und Steuerhinterziehung zu überwachen und zu verhindern und die Sicherheit und Zuverlässigkeit eines von dem Verantwortlichen bereitgestellten Dienstes zu gewährleisten, oder wenn dies für den Abschluss oder die Erfüllung eines Vertrags zwischen der betroffenen Person und einem Verantwortlichen erforderlich ist oder wenn die betroffene Person ihre ausdrückliche Einwilligung hierzu erteilt hat.

⁴ In jedem Fall sollte eine solche Verarbeitung mit angemessenen Garantien verbunden sein, einschließlich der spezifischen Unterrichtung der betroffenen Person und des Anspruchs auf direktes Eingreifen einer Person, auf Darlegung des eigenen Standpunkts, auf Erläuterung der nach einer entsprechenden Bewertung getroffenen Entscheidung sowie des Rechts auf Anfechtung der Entscheidung.

⁵ Diese Maßnahme sollte kein Kind betreffen.

⁶ Um unter Berücksichtigung der besonderen Umstände und Rahmenbedingungen, unter denen die personenbezogenen Daten verarbeitet werden, der betroffenen Person gegenüber eine faire und transparente Verarbeitung zu gewährleisten, sollte der für die Verarbeitung Verantwortliche geeignete mathematische oder statistische Verfahren für das Profiling verwenden, technische und organisatorische Maßnahmen treffen, mit denen in geeigneter Weise insbesondere sichergestellt wird, dass Faktoren, die zu unrichtigen personenbezogenen Daten führen, korrigiert werden und das Risiko von Fehlern minimiert wird, und personenbezogene Daten in einer Weise sichern, dass den potenziellen Bedrohungen für die Interessen und Rechte der betroffenen Person Rechnung getragen wird und unter anderem verhindern, dass es gegenüber natürlichen Personen aufgrund von Rasse, ethnischer Herkunft, politischer Meinung, Religion oder Weltanschauung, Gewerkschaftszugehörigkeit, genetischer Anlagen oder Gesundheitszustand sowie sexueller Orientierung zu diskriminierenden Wirkungen oder zu einer Verarbeitung kommt, die eine solche Wirkung hat.

⁷ Automatisierte Entscheidungsfindung und Profiling auf der Grundlage besonderer Kategorien von personenbezogenen Daten sollten nur unter bestimmten Bedingungen erlaubt sein.

*Dieser Titel ist eine inoffizielle Beschreibung des Erwägungsgrundes.

Anhang B: Algo.Rules – Regeln für die Gestaltung algorithmischer Systeme

Erarbeitet von iRights.Lab und der Bertelsmann Stiftung in einem offenen und interdisziplinären Prozess

Präambel

Algorithmische Systeme werden in immer mehr Bereichen angewendet. Mit ihrer Hilfe werden Entscheidungen getroffen, die erhebliche Auswirkungen auf unser Leben haben. Sie bringen ebenso Chancen wie Risiken mit sich. Es liegt in unserer Hand, gemeinsam dafür zu sorgen, dass algorithmische Systeme zum Wohle der Gesellschaft gestaltet werden. Die in den Menschenrechten abgebildeten individuellen und kollektiven Freiheiten und Rechte sollen durch algorithmische Systeme gestärkt, nicht eingeschränkt werden. Regulierungen zum Schutz dieser Normen müssen durchsetzbar bleiben. Um das zu erreichen, haben wir gemeinsam mit Expert*innen und der interessierten Öffentlichkeit die vorliegenden Algo.Rules entwickelt.

Was sind die Algo.Rules?

Bei den Algo.Rules handelt es sich um einen Katalog an formalen Kriterien, die beachtet werden müssen, um eine gesellschaftlich förderliche Gestaltung und Überprüfung von algorithmischen Systemen zu ermöglichen und zu erleichtern. Sie legen eine Grundlage für ethische Erwägungen und für die Um- und Durchsetzung rechtlicher Rahmenbedingungen. Diese Kriterien sollen bereits bei der Entwicklung der Systeme mitgedacht und „by design“ implementiert werden. Die einzelnen Algo.Rules sind interdependent und als Gesamtkomplex zu betrachten. Interessierte sind dazu eingeladen, die Algo.Rules gemeinsam mit uns weiterzugestalten, sie zu übernehmen, anzupassen, zu erweitern und vor allem Wege zu suchen, sie in der Praxis anzuwenden. Die Algo.Rules sind dynamisch angelegt und sollen insbesondere hinsichtlich ihrer konkreten Umsetzung weiter verfeinert werden.

An wen richten sich die Algo.Rules?

Die Algo.Rules richten sie sich an alle Personen, die einen wesentlichen Einfluss auf die Entstehung, die Entwicklung und Programmierung, den Einsatz und die Auswirkungen algorithmischer Systeme haben, sowie an Personen, die deren Entwicklung beauftragt haben. Das sind insbesondere, aber nicht ausschließlich, Personen, die:

- zur Entwicklung und zum Einsatz von algorithmischen Systemen forschen,
- Daten erheben und sammeln, kategorisieren, bündeln und pflegen,
- über den Einsatz und die Ziele eines algorithmischen Systems entscheiden,
- auf Leitungsebene in Institutionen oder Unternehmen arbeiten, welche algorithmische Systeme einsetzen oder entwickeln,
- algorithmische Systeme programmieren,
- algorithmische Systeme testen, weiterentwickeln und anpassen,
- designen, wie Anwender*innen algorithmische Ergebnisse präsentiert bekommen, sowie
- in ihrem Arbeitsalltag mithilfe von algorithmischen Systemen Entscheidungen treffen oder über die Art der Anwendung entscheiden.

Was ist ein „algorithmisches System“?

Ein Algorithmus ist eine eindeutige Handlungsvorschrift zur Lösung eines vorab definierten Problems. Ein algorithmisches System ist ein System einer oder mehrerer Algorithmen, die in Software implementiert wurden, um Daten zu erfassen, zu analysieren und Schlüsse zu ziehen, die zur Lösung eines vorher definierten Problems beitragen sollen. Das System kann dabei selbstlernend sein oder vorprogrammierten Entscheidungsregeln folgen. Die Bewertung eines algorithmischen Systems im Rahmen der Algo.Rules umfasst auch den soziotechnischen Gesamtkontext, in den die Software eingebettet ist, wie zum Beispiel die Deutung und Interpretation des Ergebnisses und die Ableitung einer Entscheidung durch Anwender*innen des Systems. Der Geltungsanspruch der Algo.Rules umfasst den gesamten Prozess der Entwicklung und Einbettung algorithmischer Systeme in den jeweiligen sozialen Kontext.

Im Fokus der Algo.Rules stehen jene algorithmischen Systeme, die einen direkten oder mittelbaren, immer jedoch signifikanten Einfluss auf das Leben der Menschen oder die Gesellschaft haben. Es geht also nicht um Gestaltungsregeln für alle algorithmischen Systeme. Wohl aber für solche, die entsprechend

relevant sind. Um zu bestimmen, ob ein algorithmisches System dieses Erfordernis erfüllt, sollte vor seiner Gestaltung eine Abschätzung über seine Wirkung getroffen werden. Je stärker der mögliche Einfluss eines algorithmischen Systems auf das Leben der Menschen oder die Gesellschaft ist, desto genauer sollte die Umsetzung der Algo.Rules geprüft werden.

1. Kompetenz aufbauen

Die Funktionsweise und die möglichen Auswirkungen eines algorithmischen Systems müssen verstanden werden.

Diejenigen, die algorithmische Systeme entwickeln, betreiben und/oder über ihren Einsatz entscheiden, müssen über die erforderliche Fachkompetenz und ein entsprechend abgestuftes Verständnis der Funktionsweisen und potenziellen Auswirkungen der Technologie verfügen. Das Teilen von individuellem und institutionellem Wissen sowie der interdisziplinäre Austausch zwischen den Aufgabenbereichen sind dafür ebenso zentral wie qualifizierende Maßnahmen. Diese sind in die Ausbildung bzw. das Onboarding neuer Mitarbeiter*innen zu integrieren. Der interdisziplinäre Austausch sollte verstetigt werden und für andere Interessierte und/oder betroffene Akteure offenstehen.

2. Verantwortung definieren

Für die Auswirkungen des Einsatzes eines algorithmischen Systems muss stets eine natürliche oder juristische Person verantwortlich sein.

Die Verantwortung bedarf einer eindeutigen Zuteilung. Die damit verbundenen Aufgaben müssen der zuständigen Person bewusst sein. Dies gilt auch für geteilte Verantwortlichkeiten bei mehreren Personen oder Organisationen. Die Zuteilung muss lückenlos sowie nach innen und außen transparent dokumentiert werden. Die Verantwortung darf weder auf das algorithmische System noch auf die Anwender:innen oder betroffenen Personen abgewälzt werden.

3. Ziele und erwartete Wirkung dokumentieren

Die Ziele und die erwartete Wirkung des Einsatzes eines algorithmischen Systems müssen vor dessen Einsatz dokumentiert und abgewogen werden.

Die Ziele des algorithmischen Systems müssen klar definiert

und Informationen zu dessen Einsatz dokumentiert werden. Dazu zählen etwa die zugrunde liegenden Daten und Berechnungsmodelle. Vor dem Einsatz des algorithmischen Systems muss eine dokumentierte Folgenabschätzung durchgeführt werden. Insbesondere bei lernenden Systemen und in dynamischen Einsatzfeldern mit viel Veränderung ist die Folgenabschätzung in regelmäßigen Abständen zu wiederholen. Hierbei sind Risiken für Diskriminierungen und weitere das Individuum und das Gemeinwohl berührende Folgen im Blick zu behalten. Werteabwägungen bei der Zielsetzung und dem Einsatz von algorithmischen Systemen müssen festgehalten werden.

4. Sicherheit gewährleisten

Die Sicherheit eines algorithmischen Systems muss vor dessen Einsatz getestet und fortlaufend gewährleistet werden.

Zuverlässigkeit und Robustheit eines algorithmischen Systems und der zugrunde liegenden Daten gegenüber Angriffen, Zugriffen und Manipulationen sind unbedingt zu gewährleisten. Dafür muss Sicherheit von Anfang an festes Element der Gestaltung des algorithmischen Systems sein (*Security by Design*). Das System ist vor seinem Einsatz in einer geschützten Umgebung zu testen. Die Sicherheitsvorkehrungen sind zu dokumentieren.

5. Kennzeichnung durchführen

Der Einsatz eines algorithmischen Systems muss gekennzeichnet sein.

Beim Einsatz algorithmischer Systeme muss durch eine entsprechende Kennzeichnung für Personen, die mit ihnen interagieren, erkennbar sein, dass der Entscheidung oder Prognose ein algorithmisches System zugrunde liegt. Das gilt besonders dann, wenn das System einen Menschen in Art und Weise der Interaktion (Sprache, Aussehen etc.) imitiert.

6. Nachvollziehbarkeit sicherstellen

Die Entscheidungsfindung eines algorithmischen Systems muss stets nachvollziehbar sein.

Ein algorithmisches System muss mit seinen direkten oder mittelbaren Wirkungen und seiner Funktionsweise für Menschen leicht verständlich gemacht werden, damit die-

se es hinterfragen und überprüfen können. Dafür müssen Informationen über die dem System zugrunde liegenden Daten und Modelle, seine Architektur sowie die möglichen Auswirkungen veröffentlicht und in leicht verständlicher Sprache dargestellt werden. Es ist stets zu prüfen, ob sich ein Ziel ohne wesentliche Qualitätseinbußen auch mit einem algorithmischen System erreichen lässt, das weniger komplex und in seiner Funktionsweise leichter nachvollziehbar ist.

7. Beherrschbarkeit absichern

Ein algorithmisches System muss während seines gesamten Einsatzes gestaltbar sein und bleiben.

Damit ein algorithmisches System gestaltbar bleibt, müssen alle Personen, die an seiner Entwicklung und seinem Einsatz beteiligt sind, gemeinsam stets die Kontrolle über das System behalten. Dabei muss sichergestellt werden, dass der Gesamtüberblick über das System stets gewahrt bleibt, auch wenn Aufgaben zwischen unterschiedlichen Personen und Arbeitsbereichen verteilt sind. Die Arbeitsweise eines Systems darf niemals so komplex und undurchschaubar werden, dass es von Menschen nicht mehr beherrschbar ist oder nicht mehr geändert werden kann. Das gilt insbesondere für selbstlernende Systeme. Kann diese Beherrschbarkeit nicht sichergestellt werden, so ist auf den Einsatz eines algorithmischen Systems zu verzichten.

8. Wirkung überprüfen

Die Auswirkungen eines algorithmischen Systems müssen regelmäßig überprüft werden.

Ein algorithmisches System muss einer aktiven Kontrolle unterliegen, ob die beabsichtigten Ziele tatsächlich verfolgt werden und der Einsatz des Systems bestehendes Recht nicht verletzt. Externe Prüfstellen sollten unter Wahrung legitimer Geschäftsgeheimnisse durch entsprechende tech-

nische Vorkehrungen in die Lage versetzt werden, ein algorithmisches System tatsächlich und umfassend unabhängig prüfen zu können. Wird eine negative Wirkung festgestellt, muss die Fehlerursache ermittelt und das algorithmische System entsprechend angepasst werden.

9. Beschwerden ermöglichen

Fragwürdige oder die Rechte einer betroffenen Person beeinträchtigende Entscheidungen eines algorithmischen Systems müssen erklärt und gemeldet werden können.

Die ein algorithmisches System einsetzende Stelle muss leicht zugängliche Wege zur Kontaktaufnahme zur Verfügung stellen. Betroffene Personen müssen erstens eine qualifizierte und detaillierte Auskunft zur konkreten Entscheidung und der dahinter liegenden Abwägung einfordern können. Diese Möglichkeit sollte auch in ihrem Namen handelnden oder eigene legitime Interessen geltend machenden Organisationen eingeräumt sein. Zweitens muss eine einfache, niedrighschwellige und wirksame Beschwerdemöglichkeit zur Verfügung stehen. Beschwerden und eingeleitete Maßnahmen sind zu dokumentieren.

Raum für Notizen



Impressum

Stand der Recherchen: März 2019

Autorin:

Judith Scheer ist im Land Berlin als Trainee im Nachwuchskräfteprogramm der Senatsverwaltung für Inneres und Sport angestellt. Im Rahmen des Traineeships absolvierte sie u.a. eine 6-monatige Station an der Landestelle für Gleichbehandlung – gegen Diskriminierung. Zuvor war sie in verschiedenen NGOs für benachteiligte Schüler*innen und für Geflüchtete als Projektmanagerin tätig. Judith Scheer verfügt über einen Bachelor in European Studies der Universität Maastricht und einen Master in Soziokulturellen Studien der Europa-Universität Viadrina. Sie ist Vorstandsmitglied von Humanity in Action Deutschland e.V.

Herausgeberin:

Senatsverwaltung für Justiz, Verbraucherschutz und Antidiskriminierung
Landesstelle für Gleichbehandlung – gegen Diskriminierung (LADS)
Salzburger Straße 21–25 | 10825 Berlin
Telefon: 030 9013-3460
Internet: www.berlin.de/lads
E-Mail: antidiskriminierungsstelle@senjustva.berlin.de



Redaktion:

Sofia Ratsitska

Gestaltung:

ariadne an der spree GmbH

Druck:

Newprint Blue GmbH

ViSdP.:

Pressestelle der Senatsverwaltung für Justiz, Verbraucherschutz und Antidiskriminierung
Telefon: 030 9013-3622
E-Mail: pressestelle@senjustva.berlin.de
Auflage: 1000 Stück
Juli 2019

Bildnachweis:

Seite 6: © arno
Titelbild: © Gerd Altmann (Pixabay)

Diese Veröffentlichung ist Teil der Öffentlichkeitsarbeit des Landes Berlin. Sie ist nicht zum Verkauf bestimmt und darf nicht zur Werbung von politischen Parteien verwendet werden. Die Meinung der Autorin muss nicht in jedem Fall mit der Meinung der Herausgeberin übereinstimmen.

ISBN:

978-3-947001-02-6



Senatsverwaltung
für Justiz, Verbraucherschutz
und Antidiskriminierung



Landesstelle für Gleichbehandlung – gegen Diskriminierung (LADS)
Salzburger Straße 21-25 | 10825 Berlin
Telefon: 030 9013-3460
www.berlin.de/lads
antidiskriminierungsstelle@senjustva.berlin.de