

WIE ALGORITHMEN VERSTÄNDLICH WERDEN

Ideen für Nachvollziehbarkeit von
algorithmischen Entscheidungs-
prozessen für Betroffene

Leonie Beining

INHALT

1. EINFÜHRUNG	6
2. WOZU? WARUM TRANSPARENZ UND NACHVOLLZIEHBARKEIT FÜR DAS GEMEINWOHL WICHTIG SIND	9
3. WAS? FÜNF DIMENSIONEN VON TRANSPARENZ FÜR MEHR NACHVOLLZIEHBARKEIT FÜR BETROFFENE	11
4. ABER WIE? HERAUSFORDERUNGEN FÜR TRANSPARENZ UND NACHVOLLZIEHBARKEIT VON ADM FÜR BETROFFENE	12
5. TRANSPARENZ UND NACHVOLLZIEHBARKEIT IN DER PRAXIS	15
Recruiting – Bewerber:innen den Durchblick geben	16
Predictive Policing – weniger Hollywood, mehr Bürgerorientierung durch Transparenz	19
Gesundheits-Apps – Transparenz gegen Risiken und Nebenwirkungen	24
6. KEIN „ONE SIZE FITS ALL“ – EINE CHECKLISTE FÜR MEHR NACHVOLLZIEHBARKEIT VON ADM	28
7. EMPFEHLUNGEN	31
8. FAZIT	35

VORWORT

Algorithmen sind oft unsichtbar, aber alles andere als unwichtig für unser Leben. Sie erleichtern uns nicht nur den Alltag, suchen die schnellste Route von A nach B aus, schlagen Filme auf Netflix oder den nächsten Einkauf auf Amazon vor. Algorithmische Systeme können auch Muster in großen Datenmengen erkennen, wenn wir Menschen den Überblick verlieren, effizient berechnen, wofür wir Menschen viel mehr Zeit bräuchten, konsistent entscheiden, wo wir Menschen mit Vorurteilen kämpfen. Klug eingesetzt können Algorithmen die Welt besser machen, für mehr Chancengerechtigkeit und eine nachhaltigere Gesellschaft sorgen. Damit das gelingt, müssen wir uns ihrer Grenzen bewusst werden. Technologie ist kein Allheilmittel und wird nie fehlerfrei sein. Zum Beispiel können sich über die Daten, die Programmierung, die vorgegebenen Ziele oder die konkrete Software-Nutzung diskriminierende Muster einschleichen. Dann drohen Algorithmen im Gegenteil soziale Ungleichheit zu reproduzieren und gar zu verstärken.

Die gute Nachricht ist: Ob sich die Chancen oder die Risiken algorithmischer Systeme verwirklichen, liegt an uns. Wir als Gesellschaft müssen uns unserer Verantwortung stellen, dem Einsatz der Systeme klare Schranken setzen und zugleich ihre gemeinwohlorientierte Nutzung fördern. Im gesellschaftlichen Algorithmen-Diskurs dominieren bislang zu oft die Extreme: Verheißungsszenarien auf der einen oder Schreckensszenarien auf der anderen Seite.

Das Kooperationsprojekt „Algorithmen fürs Gemeinwohl“ der Bertelsmann Stiftung und der Stiftung Neue Verantwortung verfolgt daher bewusst einen komplementären Weg, der die gesellschaftlichen Chancen algorithmischer Systeme in den Blick nimmt, ohne vor den Risiken die Augen zu verschließen. In den vergangenen zwei Jahren haben wir drei exemplarische Anwendungsbereiche untersucht, in denen automatisierte Softwareentscheidungen (englisch: automated decision-making, kurz: ADM) besonders einflussreich für unser Leben sind.

Wir haben genau geprüft, wo die Chancen und Risiken algorithmischer Systeme liegen und wie sie gemeinwohlorientiert gestaltet werden können. Im Rahmen von vielfältig besetzten Workshops haben wir die transdisziplinäre Problemanalyse geschärft und den unmittelbaren Austausch über den gemeinwohlförderlichen Einsatz algorithmischer Systeme ermöglicht. Der erste intersektorale Workshop fand im April 2018 statt und widmete sich dem Thema „Vorausschauende Polizeiarbeit“ (Predictive Policing).¹ Im Juli 2018 kamen Entwickler:innen, Personalmanager:innen, Wissenschaftler:innen, Vertreter:innen von Arbeitgeber:innen und des öffentlichen Sektors sowie Akteur:innen der Zivilgesellschaft zu einem zweiten Workshop zum

¹ Das daraus hervorgegangene Policy Paper findet sich unter <https://www.stiftung-nv.de/de/publikation/vor-die-lage-kommen-predictive-policing-deutschland>.

Thema „Automatisierte Personalauswahl“ (Robo Recruiting) zusammen.² Beim dritten Workshop im November 2018 standen Gesundheits-Apps im Mittelpunkt.³

Das Projekt bewegte sich auf diese Weise vom Konkreten zum Allgemeinen: Durch den Blick auf spezifische Anwendungsbeispiele identifizierten wir wichtige Handlungsfelder der Algorithmengestaltung, verglichen diese miteinander und abstrahierten so übergreifende Erkenntnisse und Handlungsbedarfe, wie beispielsweise die Etablierung geeigneter Kontrollmechanismen, die Förderung von Interdisziplinarität bei der Entwicklung und Implementierung algorithmischer Systeme oder die Stärkung proaktiver und offener Kommunikation über den Einsatz der Technologie. In einer abschließenden Phase nimmt das Projekt ein spezifisches Handlungsfeld für eine gemeinwohlorientierte Algorithmengestaltung genauer in den Fokus: Die Schaffung von Transparenz und Nachvollziehbarkeit für Betroffene. Im Mittelpunkt steht die Frage, wie algorithmische Entscheidungsprozesse auch den Menschen verständlich gemacht werden können, die von ihnen betroffen sind.

Transparenz und Nachvollziehbarkeit sind Kernforderungen in der Debatte über den ethischen Umgang mit Algorithmen. Doch nur selten werden diese Forderungen bislang konkretisiert. Insbesondere mit Blick auf diejenigen, die von Entscheidungen betroffen sind – Bewohner:innen einer Stadt mit vorhersagender Polizeiarbeit (Predictive Policing), Bewerber:innen, deren Lebensläufe algorithmisch analysiert werden, oder Patient:innen, die Gesundheitsapps nutzen – bleiben viele Fragen offen. Das vorliegende Impulspapier möchte daher vor allem dazu beitragen, die Debatte zu strukturieren, indem es zunächst aufzeigt, *warum Transparenz und Nachvollziehbarkeit für das Gemeinwohl wichtig sind* (Kapitel 2), *fünf verschiedene Dimensionen von Transparenz und Nachvollziehbarkeit konkretisiert* (Kapitel 3) und *verschiedene Herausforderungen bei der Schaffung von Transparenz und Nachvollziehbarkeit beleuchtet* (Kapitel 4). Darüber hinaus wirft Kapitel 5 *Nachvollziehbarkeit in der Praxis* auch einen Blick in die drei konkreten Themenfelder, die während des Kooperationsprojekts im Mittelpunkt standen. Wie kann Nachvollziehbarkeit für Bürger:innen, Bewerber:innen und Patient:innen hergestellt werden? Und was lässt sich aus den Anwendungsbeispielen für andere Einsatzfelder lernen? Ziel ist es, das Thema Transparenz und Nachvollziehbarkeit anhand praktischer Beispiele greifbar zu machen. Am Ende stehen eine *Checkliste für mehr Nachvollziehbarkeit von ADM* (Kapitel 6) sowie *Empfehlungen für politische Entscheidungsträger und Stellen, die algorithmische Systeme einsetzen* (Kapitel 7). Auf diese Weise gestaltet sich unser abschließendes Impulspapier gleichermaßen praxisnah und lösungsorientiert über die Grenzen spezifischer Anwendungsfälle hinweg.

² Das daraus hervorgegangene Policy Paper findet sich unter https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/SNV_Robo_Recruiting_final.pdf.

³ https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Impulspapier_Gesund-Dank-Algorithmen.pdf

Wie nutzen wir „Algorithmen fürs Gemeinwohl“? Wenn es unserem Kooperationsprojekt gelungen ist, für dieses programmatische Ziel einige Antwortmöglichkeiten zu entwickeln, hat es sich für uns als Initiatoren sehr gelohnt. Wir danken insbesondere Leonie Beining, Tobias Knobloch und Carla Hustedt, die das Projekt mit ihrem großen Engagement und Weitblick, aber auch der nötigen Liebe zum Detail maßgeblich gestaltet haben. Unser gemeinsames Leitbild war dabei ein gedeihliches Zusammenspiel von Mensch und Maschine – dafür werden sich die Stiftung Neue Verantwortung und das Projekt Ethik der Algorithmen der Bertelsmann Stiftung auch künftig mit aller Kraft einsetzen!



A handwritten signature in black ink that reads "Anna Wohlfarth".

Anna Wohlfarth
Sprecherin des Vorstands
Stiftung Neue Verantwortung e. V.



A handwritten signature in black ink that reads "R. Müller-Eiselt".

Ralph Müller-Eiselt
Director Programm Megatrends
Bertelsmann Stiftung

1. EINFÜHRUNG

In der US-amerikanischen Stadt Boston musste die Schulbehörde ein Projekt zur Umgestaltung der Schulbusrouten und Unterrichtszeiten wieder beenden nachdem es von Seiten der Eltern zu großen Protesten gekommen war. Ein Team des Massachusetts Institute of Technology (MIT) hatte den Auftrag bekommen, mithilfe von Algorithmen ein neues System zu entwickeln, das eine bessere Auslastung der Busse, kürzere Fahrzeiten, weniger CO₂-Emissionen, Anpassungen in den Unterrichtszeiten, insbesondere einen späteren Schulbeginn für ältere Schüler und früheres Unterrichtsende für die jüngeren sowie insgesamt eine gerechtere Verteilung der Abfahrtszeiten zwischen privilegierten und ärmeren Schulbezirken versprach. Obwohl sich die Stadt während des Projekts stets darum bemühte, die Öffentlichkeit auf dem Laufenden zu halten, waren die Familien von den durch das algorithmische Entscheidungssystem erzeugten Veränderungen vollkommen überrascht. Es kam zu Demonstrationen und eine Petition wurde gestartet, um das neue System abzuwenden.⁴

Das Beispiel unterstreicht, dass es für die gesellschaftliche Akzeptanz von algorithmischen Entscheidungssystemen (Algorithmic Decision Making, ADM) äußerst wichtig sein kann, ihren Einsatz und ihre Funktionsweise auch für diejenigen transparent und nachvollziehbar zu machen, die von

⁴ Goodman, Ellen P. (2019): Defining Equity in Algorithmic Change. In: The Regulatory Review. <https://www.theregreview.org/2019/02/11/goodman-defining-equity-algorithmic-change/> (Stand: 13.11.2019)

ihnen betroffen sind.⁵ Sicher gibt es Anwendungsbereiche, in denen Algorithmen auch ohne größere Aufklärung bedenkenlos zum Einsatz kommen können, wie etwa Navigation oder Spracherkennung. Und gewiss sind Transparenz und Nachvollziehbarkeit kein Allheilmittel, das die Einhaltung ethischer Standards bei der Verwendung von ADM garantiert oder eine Kontrolle durch Expert:innen ersetzt. Doch wenn algorithmische Systeme über die Verteilung von Lebenschancen entscheiden, muss es für Betroffene ersichtlich sein, wo Algorithmen eingreifen und welchen Einfluss sie mitunter auf das eigene Leben haben können.⁶ Das ist Grundbedingung für einen gemeinwohlorientierten Einsatz von ADM und der Fokus des vorliegenden Papiers.

Forderungen nach Transparenz und Nachvollziehbarkeit bestimmen längst die politische Debatte über ADM in Deutschland sowie auf europäischer und internationaler Ebene. So legen etwa die im April 2019 vorgestellten Ethik-Leitlinien für vertrauenswürdige Künstliche Intelligenz (KI) der Hochrangigen Expertengruppe der Europäischen Kommission fest, dass (End-)Nutzer über den Einsatz von KI informiert werden müssen. Auch die KI-Strategie der Bundesregierung benennt „transparente Verfahren und Nachvollziehbarkeit für Bürgerinnen und Bürger“ als eine Voraussetzung für den Erfolg von KI. Das im Oktober 2019 vorgestellte Gutachten der Datenethikkommission identifiziert konkrete Anforderungen an die Transparenz und Nachvollziehbarkeit für Betroffene. Und auch in verschiedenen rechtlichen Bestimmungen und Gesetzesvorhaben schlagen sich Transparenzforderungen bereits nieder. Nicht nur aus der im Mai 2018 in Kraft getretenen Datenschutzgrundverordnung (DSGVO) ergeben sich bestimmte Informationspflichten.⁷ Die im Frühjahr 2019 beschlossene EU-Richtlinie zur Reform des europäischen Verbraucherschutzes sieht z. B. Informationen über Preispersonalisierung und Produktrankings vor, während im Entwurf des Digital Service Act u.a. Transparenzanforderungen für Algorithmen in der Content Moderation

5 Das Papier verwendet die Begriffe „algorithmisches System“ und „Prozesse algorithmischer Entscheidungsfindung“ (ADM-Prozesse). Der Begriff „algorithmisches System“ bezieht sich auf die technische Ebene und meint das System aus Ein- und Ausgabedaten, Operationalisierungen des zu lösenden Problems und Modelle für die Anwendung von Lösungswegen (Algorithmen). Der „Prozess algorithmischer Entscheidungsfindung“ hingegen, bezieht außerdem die gesellschaftliche Einbettung mit ein, d. h. umfasst auch, wie das algorithmische System eingesetzt wird, von wem und wofür es genutzt wird. Transparenz und Nachvollziehbarkeit können grundsätzlich auf beiden Ebenen ansetzen. Siehe Krüger, Julia/Lischka, Konrad (2018): Damit Maschinen den Menschen dienen. Lösungsansätze, um algorithmische Prozesse in den Dienst der Gesellschaft zu stellen, 9f. https://algorithmenethik.de/wp-content/uploads/sites/10/2018/05/Algorithmenethik_Lösungspanorama_final_online.pdf (Stand: 18.11.2019)

6 Wie sich das Teilhabewirkungspotenzial von algorithmischen Entscheidungssystemen vergleichen lässt, haben Vieth und Wagner (2017) beispielhaft skizziert: Wichtige Anhaltspunkte ihrer Kategorisierung sind beispielsweise diese Fragen: Werden Menschen bewertet? Wie viel politische und ökonomische Macht hat der Betreiber des algorithmischen Systems? Wie abhängig sind die Bewerteten von dem Ergebnis? Wie groß ist die Reichweite des Systems? Siehe Vieth, Kilian/Wagner, Ben (2017): Teilhabe, ausgerechnet. Wie algorithmische Prozesse Teilhabechancen beeinflussen können. https://www.bertelsmann-stiftung.de/fileadmin/files/BSf/Publikationen/GrauePublikationen/Teilhabe_ausgerechnet.pdf (Stand: 14.11.2019)

7 Ob sie ein konkretes Recht auf Erklärung enthält, ist umstritten, aber grundsätzlich sieht der Text verschiedene Informationspflichten und Auskunftsrechte vor, die das Potenzial haben, in ihrem Geltungsbereich für mehr Transparenz und Nachvollziehbarkeit zu sorgen. Der Geltungsbereich der DSGVO für ADM-Entscheidungen hat allerdings enge Grenzen. Für eine Analyse der Bedeutung der DSGVO für ADM siehe Schulz, Wolfgang/Dreyer, Stephan (2018): Datenschutz-Grundverordnung reicht für Kontrolle algorithmischer Systeme nicht aus. <https://algorithmenethik.de/2018/04/13/studie-dsgvo/> (Stand: 12.11.2019)

geplant sind.⁸ In Deutschland sieht der aktuelle Entwurf des neuen Medienstaatsvertrags für Nutzer eine höhere Transparenz der Funktionsweise von Algorithmen in sozialen Netzwerken vor.⁹

Dennoch bleiben die Forderungen und Bestimmungen oftmals abstrakt. Das vorliegende Papier möchte zur Vertiefung der Debatte beitragen, indem es die Begriffe „Transparenz“ und „Nachvollziehbarkeit“ weiter konkretisiert. Dazu wirft es auch einen Blick in drei verschiedene Anwendungsbereiche von algorithmischer Entscheidungsfindung, die für unser gesellschaftliches Zusammenleben zentral sind: öffentliche Sicherheit/Polizeiarbeit, Personalwesen/Recruiting und Gesundheit. Wie kann Nachvollziehbarkeit hier für Bürger:innen, Bewerber:innen und Patient:innen hergestellt werden? Ziel ist es, das Thema Transparenz und Nachvollziehbarkeit an praktischen Beispielen zu erörtern und damit greifbarer zu machen.

Die Autorin bedankt sich beim Team Ethik der Algorithmen der Bertelsmann Stiftung, insbesondere bei Carla Hustedt für Kommentare, Feedback und die gemeinsame Konzipierung der Graphiken sowie bei Viktoria Grzymek für die Unterstützung bei der Erstellung der Graphiken. Ein großer Dank geht auch an das Team der Stiftung Neue Verantwortung, insbesondere an Maria Jacob, Johanna Famulok und Sebastian Rieger.

⁸ Europäische Kommission (2019). <https://cdn.netzpolitik.org/wp-upload/2019/07/Digital-Services-Act-note-DG-Connect-June-2019.pdf>. (Stand: 18.11.2019)

⁹ Rundfunkkommission der Länder (2019): Diskussionsentwurf für einen „Medienstaatsvertrag“, 60. https://www.rlp.de/fileadmin/rlp-stk/pdf-Dateien/Medienpolitik/MStV-E_Synopse_2019-07_Online_.pdf (Stand: 18.11.2019)

2. WOZU? WARUM TRANSPARENZ UND NACHVOLLZIEHBARKEIT FÜR DAS GEMEINWOHL WICHTIG SIND

ADM-Prozesse für Betroffene transparent und nachvollziehbar zu machen bedeutet, ihnen aktiv Informationen zur Verfügung zu stellen (Transparenz) und diese Informationen in einer Weise aufzubereiten, die es Betroffenen ermöglicht, zu verstehen, wie ein eingesetztes algorithmisches System das eigene Leben betreffen kann (Nachvollziehbarkeit). Transparenz schafft damit die Grundlage für Nachvollziehbarkeit. Dabei trägt Transparenz zunächst dazu bei, *Aufmerksamkeit zu schaffen und das Bewusstsein* für den Einsatz von algorithmischen Systemen zu stärken, das notwendig ist, um sich mit ihnen auseinanderzusetzen. Bislang fehlt dieses Bewusstsein oft: Wie eine Umfrage der Bertelsmann Stiftung zeigt, wissen weniger als die Hälfte der Europäer, dass Algorithmen in Bereichen wie Kreditentscheidungen oder Personalwesen eingesetzt werden.¹⁰

Werden ADM-Systeme für Entscheidungen über gesellschaftliche Teilhabe genutzt, ohne dass eine betroffene Person etwas darüber weiß, verfügt sie über keinerlei Handlungsmöglichkeiten: In zentralen Lebensbereichen wird sie zum Objekt algorithmischer Entscheidungsfindung. Hier bietet Nachvollziehbarkeit Betroffenen Schutz vor Informationsasymmetrien und sichert die – schon für die Entwicklung des Datenschutzrechts maßgebliche – informationelle Selbstbestimmung auch im Zuge algorithmischer Entscheidungsfindung ab. Die Schaffung von Transparenz und Nachvollziehbarkeit ist damit eine zentrale Voraussetzung für die *Wahrung von Autonomie und Menschenwürde* im digitalen Zeitalter.

Gleichzeitig eröffnet Nachvollziehbarkeit Betroffenen Handlungs- und Gestaltungsspielräume: Sie sorgt nicht nur dafür, dass sie informiert *am gesellschaftlichen Diskurs teilhaben* und sich so für ihre Interessen stark machen können. Die Möglichkeit, sich mit ADM-Prozessen auseinanderzusetzen, bietet auch die Grundlage dafür, dass Betroffene – oder zivilgesellschaftliche Wächter und Verbraucherverbände, die sie vertreten – *Fehler aufdecken, Entscheidungen anfechten und korrigieren* können. Beispielsweise wird das Allgemeine Gleichbehandlungsgesetz (AGG) zukünftig nur geringen Schutz vor Diskriminierung durch Algorithmen bieten, z.B. wenn es für Bewerber:innen nicht klar ist, dass und auf welcher Grundlage ein algorithmisches System eine Bewerbung analysiert.¹¹

¹⁰ Grzymek, Viktoria/Puntschuh, Michael (2019): Was Europa über Algorithmen weiß und denkt. Ergebnisse einer repräsentativen Bevölkerungsumfrage. <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WasEuropaUEberAlgorithmenWeissUndDenkt.pdf>. (Stand: 13.11.2019)

¹¹ Busch, Christoph (2018): Gutachten: Algorithmic Accountability: 48. <http://www.abida.de/sites/default/files/ABIDA%20Gutachten%20Algorithmic%20Accountability.pdf> (Stand: 14.11.2019)

Nachvollziehbarkeit hilft Betroffenen außerdem dabei, die *richtige Anwendung und bewussten Umgang mit einem algorithmischen System und dessen Ergebnissen sicherzustellen*. Das heißt auch, dass Betroffene ein System – z. B. die algorithmische Sortierung von Nachrichten in sozialen Netzwerken – normativ einordnen, bewerten und hinterfragen können. Für Betroffene in der Rolle des Verbrauchers ist Nachvollziehbarkeit vor allem auch wichtig, um eine bewusste Konsumententscheidung zu treffen.¹² Darüber hinaus können Auskünfte über algorithmische Entscheidungen auch als Feedback zum eigenen Verhalten funktionieren: Wenn Menschen die Gründe für eine Entscheidung erfahren, werden sie *in die Lage versetzt zu lernen*, wie andere Ergebnisse erreicht werden können und wie sie Einfluss auf Entscheidungen nehmen können.

Insgesamt schaffen Transparenz und Nachvollziehbarkeit auf diese Weise die Basis, auf der *Vertrauen in Technologie und Akzeptanz von algorithmischen Systemen* entstehen kann. So zeigen etwa verhaltenswissenschaftliche Untersuchungen, dass Menschen selbst unerfreuliche Entscheidungen eher akzeptieren, wenn sie den Prozess der Entscheidungsfindung als fair und angemessen wahrnehmen.¹³ Dies wird nicht gelingen, wenn algorithmische Systeme für Betroffene eine Black Box sind, über die sie nichts wissen und deren Bedeutung sie nicht einschätzen können.¹⁴



12 Verbraucherzentrale Bundesverband e.V. (2019): Algorithmenkontrolle. Positionspapier der Verbraucherzentrale Bundesverband (vzbv). https://www.vzbv.de/sites/default/files/downloads/2019/05/02/19-05-02_vzbv_positionspapier_algorithmenkontrolle.pdf (Stand: 09.11.2019)

13 Hustedt, Carla (2016): A Public Value Perspective on Predictive Policing in the US. https://www.researchgate.net/publication/327155597_A_Public_Value_Perspective_on_Predictive_Policing_in_the_US (Stand: 18.11.2019); Tyler, Tom R. (2003): Procedural Justice, Legitimacy, and the Effective Rule of Law. In: Crime & Justice 30: 283, 291.

14 Kizilcec, René F. (2016): How Much Information? Effects of Transparency on Trust in an Algorithmic Interface. <http://rene.kizilcec.com/wp-content/uploads/2016/01/kizilcec2016information.pdf> (Stand: 15.11.2019)

3. WAS? FÜNF DIMENSIONEN VON TRANSPARENZ FÜR MEHR NACHVOLLZIEHBARKEIT FÜR BETROFFENE

Noch bleibt oft unklar, was es braucht, um Nachvollziehbarkeit für Betroffene zu schaffen. Was sollte transparent gemacht, d. h. welche Informationen sollten bereitgestellt werden? Die Bandbreite ist groß und verschiedene Stakeholder haben ganz unterschiedliche Ansprüche an den Gehalt der Informationen. Während es für Aufsichtsbehörden mit der Aufgabe algorithmische Systeme zu kontrollieren mitunter sinnvoll sein kann, Einblick in den Quellcode zu erhalten, ist dies für Betroffene nicht zielführend. Für sie ist ein anderer Ansatz notwendig, um algorithmische Systeme nachvollziehbar zu machen. Fünf Dimensionen sind für Betroffene dabei besonders zentral:

- Kommt ein algorithmisches System zum Einsatz?

In teilhaberelevanten Anwendungsbereichen sollte es für Betroffene grundsätzlich transparent gemacht werden, wenn algorithmische Systeme in Entscheidungsprozessen zum Einsatz kommen. Je nach Kontext und Anwendungsbereich kann das durch eine Kennzeichnung zu Beginn einer Interaktion, z. B. durch ein Symbol, geschehen. In Situationen, in denen Betroffene nicht mit dem algorithmischen System interagieren, muss über andere Formate nachgedacht werden. Der Informationsgehalt ist für Betroffene zwar klar beschränkt und weist lediglich die Existenz von ADM aus, sorgt aber gleichzeitig für Aufmerksamkeit, trägt zur Bewusstseinsbildung bei und ist Grundlage für die anderen Dimensionen von Nachvollziehbarkeit.

- Wie funktioniert das algorithmische System?

Um die in Abschnitt 2 aufgeführten Funktionen zu erfüllen, reicht eine einfache Information über den Einsatz eines ADM jedoch nicht aus. Zusätzlich müssen Informationen, die sich auf die Funktionsweise von ADM beziehen, bereitgestellt werden, um die Black Box ein Stück zu öffnen. Betroffene müssen erfahren, welche Daten das Entscheidungssystem verwendet, anhand welcher Daten es ggf. trainiert wurde, welche Entscheidungskriterien sich aus dem Lernprozess ergeben oder von Menschen festgelegt wurden und wie die Entscheidungskriterien gewichtet werden. Neben dieser *Systemtransparenz* ist es wichtig zu kommunizieren, was das System leisten kann und wo die Grenzen liegen, z. B. welche Fehlerquoten vorliegen, um damit auch Transparenz über die Qualität der Systeme zu schaffen.

- Warum gibt es das algorithmische System?

Mitunter ist es wichtig, Betroffenen den Kontext des Systems verständlich zu machen und Informationen über den gesamten Prozess algorithmischer

Entscheidungsfindung zur Verfügung zu stellen. Dann geht es darum, die Wertentscheidungen sichtbar zu machen, die der Entwicklung und dem Einsatz des Systems zugrunde liegen. Es kann z. B. wichtig sein zu erfahren, wer das System entwickelt hat, mit welcher Motivation es entwickelt und eingesetzt wird, was das Geschäftsmodell ist, wofür die Entscheidungen genutzt werden, was die erwartete Wirkung ist und welchen Stellenwert das System im gesamten Entscheidungsprozess hat. Aufzuzeigen, dass bestimmte Werte mit dem algorithmischen System verbunden sind und dass es nicht neutral und objektiv ist, hilft bei der realistischen Einordnung und normativen Bewertung von ADM, ob es geltenden gesellschaftlichen Maßstäben entspricht (*Kontexttransparenz*).

- Wie kommt die konkrete Entscheidung über mich zustande?

Neben der abstrakten Darstellung des Entscheidungsprozesses sollten Betroffene darüber informiert werden, welche Faktoren für die sie persönlich betreffende Entscheidung relevant waren. Die *Entscheidungstransparenz* liefert damit eine Erklärung des Ergebnisses, zu dem das algorithmische System gekommen ist, und kann Verhaltensänderungen sowie Kontrolle der Systeme ermöglichen.

- Über welche Handlungsoptionen verfüge ich?

Bestimmte Funktionen von Transparenz und Nachvollziehbarkeit, wie z. B. Beschwerde einlegen oder die Korrektur einer Entscheidung verlangen, werden nur wirksam, wenn Betroffene auch über konkrete Möglichkeiten verfügen, um auf Basis der Informationen über ein System oder eine Entscheidung konkrete Schritte zu unternehmen. Es ist wichtig, Betroffene klar über ihre Handlungsoptionen zu informieren und diese dann leicht zugänglich und benutzerfreundlich zur Verfügung zu stellen.

4. ABER WIE? HERAUSFORDERUNGEN FÜR TRANSPARENZ UND NACHVOLLZIEHBARKEIT VON ADM FÜR BETROFFENE

So wichtig es ist, algorithmische Entscheidungsprozesse für Betroffene transparent und nachvollziehbar zu machen, so schwierig gestaltet sich dies mitunter in der Praxis. Oftmals wird zunächst auf konkrete technische Herausforderungen und rechtliche Hindernisse aufmerksam gemacht: Etwa die auch für Expert:innen nicht zu durchschauende Komplexität bestimmter Methoden des Maschinellen Lernens und der Zielkonflikt, dass verständliche Modelle zum Teil weniger gute Ergebnisse liefern, sowie die Rolle von Geschäftsgeheimnissen, die vor Offenlegung geschützt werden sollen, und

etwa das Datenschutzrecht, welches Grenzen setzt, wenn es z. B. um personenbezogene Daten von Dritten geht.¹⁵ Während diese Aspekte in erster Linie Einfluss darauf haben, was überhaupt transparent gemacht werden kann, bestehen darüber hinaus noch weitere Herausforderungen, die sich vor allem auf die praktische Umsetzung und die Frage des konkreten Nutzens für die Betroffenen beziehen.

Um algorithmische Prozesse für Betroffene nachvollziehbar zu machen, besteht auf der einen Seite die Herausforderung, dass komplexe Sachverhalte so vermittelt werden müssen, dass sie verständlich sind, ohne dabei eine betroffene Person mit Informationen zu überlasten. Die Tatsache, dass Betroffene über unterschiedlich viel und meist nur begrenztes Wissen und Zeit verfügen, erschwert es zudem, das richtige Maß an Information zu finden. Für Unternehmen oder öffentliche Stellen, die ADM entwickeln oder einsetzen, stellt sich damit auf der anderen Seite eine sehr anspruchsvolle Aufgabe. Die Schaffung von Transparenz und Nachvollziehbarkeit bedeutet dabei erst einmal zusätzliche Kosten und mehr Arbeit, zunächst ohne dass sich daraus offensichtliche, direkt messbare Vorteile für die einsetzenden Stellen ergeben. Abgesehen von eventuell bestehenden rechtlichen Anforderungen fehlt so mitunter der klare Anreiz, sich um wirksame Transparenz zu bemühen. Wenn Nachvollziehbarkeit aber lediglich umgesetzt wird, um gesetzliche Mindestanforderungen zu erfüllen, Unternehmen aber keinen direkten Mehrwert erkennen, wird dies dazu führen, dass entsprechende Maßnahmen nicht mit der notwendigen Sorgfalt implementiert werden und Betroffene daraus keinen Nutzen ziehen. Im schlimmsten Fall könnten Informationsangebote so gestaltet sein, dass sie Personen über die Funktionsweise des Algorithmus in die Irre führen und zu von ihnen nicht beabsichtigtem Verhalten führt (sog. Dark Pattern).¹⁶ Die Umsetzung von datenschutzrechtlichen Informationspflichten liefert hier abschreckende Beispiele, aber auch wertvolle Erfahrungen, wenn es um die Frage geht, wie man technisch und rechtlich komplexe Sachverhalte an Betroffene kommuniziert.¹⁷

Hinzu kommt, dass es oftmals unklar und bisweilen umstritten ist, welchen Nutzen Transparenz und Nachvollziehbarkeit für Betroffene überhaupt

¹⁵ Auf technischer Ebene ist die Erklärbarkeit solcher Modelle aber zentraler Gegenstand der wissenschaftlichen Forschung, die unter dem Stichwort Explainable AI an mathematischen Methoden zur Verbesserung der Erklärbarkeit arbeitet. Gleiches gilt für den oftmals genannten Zielkonflikt zwischen Präzision der algorithmischen Systeme und der Interpretierbarkeit. Auch hier werden Methoden entwickelt, die versprechen, die Interpretierbarkeit zu verbessern, ohne auf Kosten der Leistungsfähigkeit zu gehen. Selbst, Andrew D./Barocas, Solon (2018): The Intuitive Appeal of Explainable Machines. In: Fordham Law Review, Vol. 87. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3126971 (Stand: 18.11.2019)

¹⁶ Zur Gefahr von Dark Pattern bei der Schaffung von Transparenz und Kontrollmöglichkeiten von intelligenten Systeme für Nutzer:innen, siehe Chromik, Michael/Eiband, Malin/Völkel, Sarah T./Buschek, Daniel (2019): Dark Patterns of Explainability, Transparency, and User Control for Intelligent Systems. In: IUI Workshops '19. March 20, 2019. LA, USA. <http://www.medien.ifi.lmu.de/pubdb/publications/pub/chromik2019iuiworkshop/chromik2019iuiworkshop.pdf> (Stand: 15.11.2019)

¹⁷ Siehe z. B. Analysen von Schaub, Florian/Balebako, Rebecca/Durity, Adam/Cranor, Lorrie F. (2015): A Design Space for Effective Privacy Notices. In: Symposium on Usable Privacy and Security (SOUPS), July 22-24, 2015, Ottawa, Canada. <https://www.usenix.org/system/files/conference/soups2015/soups15-paper-schaub.pdf> (Stand: 18.11.2019) und Renaud, Karen/Shepherd, Lynsay A. (2018): How to Make Privacy Policies both GDPR-Compliant and Usable. <https://arxiv.org/pdf/1806.06670.pdf> (Stand: 13.11.2019)

bringen.¹⁸ Bezogen z.B. auf den Einfluss von Transparenz auf das Vertrauen von Betroffenen in ADM-Systeme kommen verschiedene Studien in unterschiedlichen Anwendungsbereichen zu ganz gemischten Ergebnissen.¹⁹ Ein Experiment in den USA, bei dem Studierende eine durch einen Algorithmus errechnete Note erhielten, zeigte beispielsweise, dass ein gewisses Maß an Transparenz tatsächlich zu mehr Vertrauen in die Benotung durch das algorithmische System führen kann. Das Vertrauen bei den Studierenden schwand jedoch, wenn sie viele Informationen über das System bekamen.²⁰

Transparenz und Nachvollziehbarkeit können potenziell auch den sogenannten Automation Bias verstärken. Das bedeutet, dass Menschen blind vertrauen und einfach nur das tun, was das ADM-System sagt.²¹ Die Gefahr besteht, wenn die bereitgestellten Informationen die algorithmische Entscheidung nicht angemessen einordnen, sondern als Wahrheit transportieren, weil Unsicherheiten oder Wahrscheinlichkeiten nicht erläutert werden. Beispielsweise kann der algorithmenbasierte Hinweis auf eine Erkrankung durch eine Gesundheits-App nicht in allen Fällen den Gang zum Arzt mit fachgerechter Diagnostik ersetzen und sollte entsprechend kommuniziert werden.

Eine andere Herausforderung besteht in der Befürchtung, dass die Erklärung einer algorithmischen Entscheidung dazu führt, dass sich Menschen auf die Logik von ADM-Systemen einstellen. Also beispielsweise an die Anforderungen, anhand derer ein solches System im Personalrecruiting Bewerbungsunterlagen analysiert. Stellen, die Algorithmen anwenden, sehen in diesem "Gaming" ein Argument gegen mehr Transparenz bei algorithmischen Entscheidungen.

Wenn algorithmische Entscheidungsprozesse Betroffenen verständlich gemacht werden sollen, ist es wichtig, nicht nur die technischen und rechtlichen Herausforderungen anzugehen, sondern auch die Frage der praktischen Umsetzung und den Auswirkungen von Transparenz zu adressieren. Andernfalls besteht die Gefahr, dass Bemühungen um mehr Transparenz und Nachvollziehbarkeit ins Leere laufen und keinen Nutzen für die von algorithmischen Entscheidungen Betroffenen bringen.

18 Springer, Aaron/Whittaker, Steve (2018): Progressive Disclosure: Designing for Effective Transparency <https://arxiv.org/ftp/arxiv/papers/1811/1811.02164.pdf> (Stand: 12.11.2019)

19 Einen Überblick über verschiedene Studien zu Auswirkungen von Transparenz auf Nutzer:innen von algorithmenbasierten Anwendungen geben Felzmann, Heike/Fosch Villaronga, Eduard/Lutz, Christopher/Tamo-Larrieux, Aurelia (2019): Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. In: Big Data & Society January-June 2019: 6. <https://journals.sagepub.com/doi/pdf/10.1177/2053951719860542> (Stand: 18.11.2019)

20 Hosanagar, Kartik/Jair, Vivian (2018): We Need Transparency in Algorithms, But Too Much Can Backfire. In: Harvard Business Review. 25.07.2018. <https://hbr.org/2018/07/we-need-transparency-in-algorithms-but-too-much-can-backfire> (Stand: 09.11.2019)

21 Binns, Reuben/Gallo, Valeria (2019): Automated Decision Making: the role of meaningful human reviews. <https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/> (Stand: 12.11.2019)

5. TRANSPARENZ UND NACHVOLLZIEHBARKEIT IN DER PRAXIS

Der politische und gesellschaftliche Diskurs über Transparenz und Nachvollziehbarkeit von ADM-Prozessen bleibt bislang oft abstrakt. Was aber bedeuten die Forderungen in der Praxis? Das vorliegende Papier untersucht daher drei konkrete teilhaberelevante Bereiche, in denen in Deutschland ADM-Systeme zum Einsatz kommen,

- welche Rolle Transparenz und Nachvollziehbarkeit bereits heute spielen,
- warum Transparenz und Nachvollziehbarkeit gerade im Zuge von algorithmischer Entscheidungsfindung in diesen Bereichen von Bedeutung ist,
- welche Herausforderungen bestehen,
- wie jeweils für größere Transparenz und Nachvollziehbarkeit gesorgt werden kann.

Die drei untersuchten Bereiche – Recruiting, Gesundheitswesen und Polizeiarbeit – stellen ganz spezifische Anwendungsfelder von algorithmischen Systemen dar, die sich klar voneinander unterscheiden. Während beim sogenannten Predictive Policing ADM-Systeme von staatlicher Seite eingesetzt werden, stammen Anwender:innen algorithmenbasierter Recruitinginstrumente auch aus dem Privatsektor. Sowohl Bürger:innen als auch Bewerber:innen können sich ihrem Einsatz nicht – oder nur unter sehr hohen Kosten – entziehen. Ob im Zuge der Gesundheitsversorgung ADM-Systeme in Form von Gesundheits-Apps genutzt werden bleibt hingegen bislang jeder/m Einzelnen überlassen. Auch der Fokus der jeweiligen algorithmischen Entscheidung unterscheidet sich. In Deutschland werden mithilfe von Predictive Policing etwa Prognosen für ganze Städte bzw. Stadtteile erstellt, während Gesundheits-Apps und ADM-Systeme im Recruiting individuellen Analysen dienen. Durch den Blick auf die drei verschiedenen Bereiche sollen Unterschiede und Gemeinsamkeiten im Einsatz algorithmischer Entscheidungssysteme aufgezeigt werden, um so übergreifende Erkenntnisse und Handlungsbedarfe für eine bessere Nachvollziehbarkeit von ADM für Betroffene zu identifizieren.²²

22 Die Praxisbeispiele basieren maßgeblich auf Gesprächen und Interviews mit Expert:innen aus den unterschiedlichen Anwendungsbereichen und den Teilnehmer:innen einer Workshopreihe, die die Stiftung Neue Verantwortung gemeinsam mit der Bertelsmann Stiftung im Rahmen des Projekts „Algorithmen fürs Gemeinwohl“ im Jahr 2018 durchgeführt hat.

RECRUITING – BEWERBER:INNEN DEN DURCHBLICK GEBEN

Bei der Suche und Gewinnung neuer Mitarbeiter kommen in Deutschland zunehmend algorithmenbasierte Systeme zum Einsatz, z. B. zur Analyse des Arbeitsmarktes, zur Talentsuche auf Karriereplattformen, zum Vorfiltrieren von Bewerbungen oder zur Unterstützung von Assessment Centern. In einer Studie der Universität Bamberg und des Karriereportals Monster geben 70 Prozent der Top-1.000-Unternehmen an, dass die Vorauswahl von Bewerber:innen in Zukunft immer häufiger automatisiert geschehen wird.²³

Nachvollziehbarkeit im Recruiting – die Ausgangslage

Die Forderung nach Transparenz und Nachvollziehbarkeit trifft beim Recruiting auf einen Bereich, in dem Entscheidungen für Betroffene traditionell eher intransparent sind. Rechtlich sind Arbeitgeber, die eine Stelle vergeben wollen, nicht dazu verpflichtet, Gründe für die Ablehnung anzugeben. Intransparenz wird aus arbeitsrechtlicher Sicht sogar meist empfohlen, da Ablehnungsgründe nach dem Allgemeinen Gleichbehandlungsgesetz (AGG) diskriminierend ausgelegt werden können. Lediglich für Bewerber:innen, die schwerbehindert oder Schwerbehinderten gleichgestellt sind, gelten andere Informationspflichten.

In der gelebten Praxis sieht es jedoch mitunter anders aus: Beispielsweise gehört es für Personaler zum Abschluss eines längeren, mehrstufigen Verfahrens unter Umständen dazu, ein persönliches Gespräch zu führen, um die Möglichkeit für ein ausführlicheres Feedback zu bieten. Bestandteil des Feedbacks sind dabei durchaus auch Empfehlungen mit Blick auf die Entwicklung der abgelehnten Bewerber:innen und die Frage, welche Kompetenzen notwendig wären, um bei der nächsten Bewerbung erfolgreich zu sein. Zur Praxis gehört es mitunter auch, zu Beginn des Prozesses detailliert über den zeitlichen Verlauf des Bewerbungsverfahrens zu informieren und beispielsweise den Bearbeitungsstand der Bewerbung per Statusanzeige in Echtzeit zu dokumentieren.²⁴ Die bisweilen schwierige Suche nach passendem Personal und die Entwicklung von Plattformen für berufliche Netzwerke und Arbeitgeberbewertungen führen dazu, dass Einstellungsprozesse in jüngster Zeit oftmals transparenter werden. Offenheit ist ein Hebel, um für potenzielle Kandidat:innen attraktiv zu sein. Ein klarer Vorteil gegenüber Wettbewerbern, da sich Bewerber:innen heute mehr Transparenz wünschen (Stichwort „Employer Branding“).²⁵ Transparenz schafft sowohl Glaubwürdigkeit als auch Authentizität und hilft beim Aufbau langfristiger

23 Einen Einblick in den Einsatz von ADM im Recruiting geben Knobloch, Tobias/Hustedt, Carla (2019): Der maschinelle Weg zum passenden Personal. Zur Rolle algorithmischer Systeme in der Personalauswahl. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/SNV_Robo_Recruiting_final.pdf (Stand: 18.11.2019)

24 Benhof, Andreas (2018): Transparenz statt Blackbox im Bewerbungsprozess. In: Medium. 20.06.2018. <https://medium.com/recruiting-essentials/transparenz-statt-blackbox-im-bewerbungsprozess-fb53390865eb> (Stand: 18.11.2019)

25 Personalwirtschaft (2016): Aus Kandidatensicht noch optimierbar. In: Personalwirtschaft. 03.06.2016. <https://www.personalwirtschaft.de/recruiting/bewerbermanagement/artikel/aus-kandidatensicht-noch-optimierbar.html> (Stand: 15.11.2019)

Beziehungen über ein konkretes Bewerbungsverfahren hinaus, auf die bei zukünftiger Stellenbesetzung zurückgegriffen werden kann. Die Hintergründe des eigenen Abschneidens besser zu verstehen bietet Ansatzpunkte für die individuelle Weiterentwicklung, wovon auch Arbeitgeber in zukünftigen Verfahren profitieren können.

Robo-Recruiting: neue Ansprüche an Nachvollziehbarkeit?

ADM-Systeme im Recruiting bergen die Gefahr, Bewerbungsverfahren für Bewerber:innen noch undurchsichtiger zu machen als sie ohnehin oft sind, indem Entscheidungsprozesse auf ein System verlagert werden, von dessen Einsatz die Bewerber:innen keine Kenntnis haben oder dessen Funktionsweise nicht verständlich ist. Der Einsatz von ADM-Systemen steht damit im Widerspruch zum Wunsch vieler Kandidat:innen nach mehr Transparenz in Bewerbungsverfahren. Zudem sehen Bewerber:innen im Einsatz von ADM-gestützten Recruitingmethoden mitunter eine „Entmenschlichung“ des Recruitingprozesses.²⁶ Und nicht zuletzt besteht das Risiko, dass ADM-Systeme problematische Entscheidungsverfahren weiter verschleiern. Um für Vertrauen und Akzeptanz zu sorgen und dem bestehenden Diskriminierungspotenzial entgegenzuwirken, ist es daher wichtig, den Einsatz von algorithmischen Systemen gut zu begründen, verständlich zu machen, und Bewerber:innen die Möglichkeit zu geben, Widerspruch einzulegen oder Korrekturen zu verlangen.

Nachvollziehbarkeit von algorithmischen Recruitingentscheidungen liegt damit auch im Interesse der Stellen, die entsprechende Systeme einsetzen. Nur so können sie sich als verantwortungsvoller Arbeitgeber präsentieren: Der gute Eindruck fängt im digitalen Zeitalter bereits bei den Maschinen an, die sie einsetzen. Die Nutzung von algorithmischen Systemen und die sich daraus ergebende Notwendigkeit, die Ziele und Kriterien, die Entscheidungsprozessen zugrunde liegen sollen, vorher zu definieren, stellt prinzipiell eine gute Basis für transparentere Personalentscheidungen dar. Der Einsatz von Algorithmen für die Personalgewinnung eröffnet somit die Chance, den Trend zu mehr Transparenz durch konsequente Übertragung auch auf algorithmenbasierte Recruitingverfahren zu befördern.

Unter Arbeitgeber besteht jedoch teilweise die Befürchtung, dass die genaue Erklärung der algorithmischen Entscheidung dazu führt, dass Bewerber:innen nicht mehr ihre individuelle Qualifikation für eine Stelle hervorheben, sondern sich auf die Funktionsweise der Systeme einstellen und versuchen, sie auszutricksen (sogenanntes Gaming).²⁷

²⁶ Laut einer Umfrage der Bertelsmann Stiftung aus dem Jahr 2018 bevorzugen 49 Prozent der Befragten, dass bei der Vorauswahl von Bewerber:innen nach festgelegten Kriterien allein ein Mensch entscheidet. Fischer, Sarah/Petersen, Thomas (2018): Was Deutschland über Algorithmen weiß und denkt. Ergebnisse einer repräsentativen Bevölkerungsumfrage: 28. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Was_die_Deutschen_ueber_Algorithmen_denken.pdf (Stand: 18.11.2019)

²⁷ Knobloch, Tobias/Hustedt, Carla (2019): Der maschinelle Weg zum passenden Personal. Zur Rolle algorithmischer Systeme in der Personalauswahl. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/SNV_Robo_Recruiting_final.pdf (Stand: 18.11.2019)

Je komplexer die eingesetzten Systeme werden, umso schwieriger wird es mitunter, ihre Funktionsweise für Betroffene verständlich zu machen. Wenn ADM-Systeme für Menschen schwer interpretierbare Faktoren wie beispielsweise die Analyse der Stimme in einem Telefoninterview in die Entscheidungsfindung einbeziehen, kann die Information über diesen Entscheidungsfaktor das Misstrauen in das System zudem verstärken.

Eine weitere Herausforderung bei der Herstellung von Nachvollziehbarkeit ist, dass dem eigentlichen Bewerbungsverfahren oft bereits viele technisch unterstützte Auswahlentscheidungen vorgelagert sind. Zum Beispiel wenn es darum geht, welche potenziellen Bewerber:innen ein Jobangebot zu sehen bekommen und welche nicht.

Dies zeigt beispielhaft die Grenzen von Transparenz und Nachvollziehbarkeit gegenüber Betroffenen auf. Schutz kann hier nur durch eine entsprechende Regulierung und unabhängige Kontrolle der Systeme gewährleistet werden.

Wege für mehr Nachvollziehbarkeit

Ein erster essenzieller Schritt ist es, Bewerber:innen im Rahmen des Bewerbungsverfahrens darüber zu informieren, dass algorithmische Systeme zur Anwendung kommen. Nachvollziehbarkeit sollte nicht erst nach dem jeweiligen Verfahrensende hergestellt werden, sondern bereits zu Beginn größtmögliche Klarheit über die eingesetzten Verfahren gewährleistet sein. Eine Option, die ermöglicht, dass sich Bewerber:innen frühzeitig mit dem Einsatz von Algorithmen im Bewerbungsprozess auseinandersetzen können, sind bereits Ausschreibungen. Im Sinne einer „realistic job preview“ können Stellenausschreibungen dafür genutzt werden, um auch über algorithmische Recruitingverfahren Klartext zu reden. Da Bewerbungsprozesse mittlerweile zu einem Großteil online erfolgen, kann die Information über eine gut sichtbare Einblendung/ein Infofenster auf der entsprechenden Website erfolgen. Um sicherzustellen, dass diese Information Bewerber:innen nicht abschreckt, kann zusätzlich anhand von Beispielen aufgezeigt werden, was ein System misst, wie Ergebnisse grundsätzlich zustande kommen und welchen Stellenwert sie bei der letztlichen Entscheidung haben. Auch Personaler:innen können, wenn sie hierfür geschult werden, die Aufklärung über den Einsatz des algorithmischen Systems übernehmen. Die Information im persönlichen Gespräch hat den klaren Vorteil, dass Betroffene nachfragen können.

Um zu verhindern, dass Bewerber:innen die Information nutzen, um das System auszutricksen („Gaming“), gilt es, das Verfahren so konkret wie möglich und so allgemein wie nötig zu erklären. Auch Visualisierungen können für ein besseres Verständnis zur Hilfe gezogen werden. Darüber hinaus ist es von zentraler Bedeutung, den Einsatz des ADM in den Gesamtkontext zu stellen und zu erklären, warum man die Technologie einsetzt („Wir machen xy, weil wir damit gut yz beurteilen können“). Idealerweise kann hierbei auch verdeutlicht werden, wie die Bewerber:innen vom Einsatz des Systems profitieren (Wandel von der Erfahrungs- hin zur Fähigkeitsmessung

und dadurch besseres Matching von Stelle und Bewerber:innen). Allein die Information, dass ein ADM-System eingesetzt wird, ist nicht ausreichend, sondern würde dazu führen, dass das Misstrauen gegenüber dem gesamten Bewerbungsprozess und im Ergebnis auch gegenüber dem potenziellen Arbeitgeber steigt.

Sobald eine Entscheidung getroffen wurde, gilt es, die allgemeinen Informationen über die Funktionsweise des Systems für die betroffene Person zu personalisieren, d. h. sie sollte die Möglichkeit erhalten zu verstehen, welche Faktoren für die Entscheidung relevant waren. Insbesondere für diejenigen, für die ein algorithmengestütztes Verfahren mit einer Absage endet, ist diese Information zentral. Für diesen Zweck bietet sich für Arbeitgeber der Ansatz der „counterfactual explanation“ an, der Bewerber:innen Einblick gibt, welche kleinstmögliche Veränderung im Lebenslauf oder der Kompetenzen mindestens nötig wäre, um Aussicht auf die Stelle zu haben. Darüber hinaus gilt es, konkrete Kontaktmöglichkeiten anzugeben, an die sich Bewerber:innen unkompliziert bei Fragen und Beanstandungen rund um das eingesetzte ADM-System wenden können.

PREDICTIVE POLICING – WENIGER HOLLYWOOD, MEHR BÜRGERORIENTIERUNG DURCH TRANSPARENZ

„Predictive Policing“ oder vorausschauende Polizeiarbeit bedeutet, dass Polizeibehörden Software einsetzen, die auf der Grundlage historischer Kriminalitätsdaten und aktueller kriminologisch relevanter Daten über das Wetter, die Verkehrslage oder auch anstehende Großveranstaltungen Prognosen über zukünftige Verbrechen erstellt. Beim sogenannten ortsbezogenen Predictive Policing bezieht sich die Prognose auf geographische Gebiete mit erhöhter Verbrechenswahrscheinlichkeit, beim personenbezogenen Predictive Policing auf die Wahrscheinlichkeit, dass eine bestimmte Person oder Personengruppe eine Straftat verüben wird. In Deutschland kommen bislang lediglich die sogenannten ortsbezogenen Systeme zum Einsatz. Sechs Bundesländer setzen sie zurzeit ein, vor allem um Wohnungseinbrüche zu verhindern.²⁸

Nachvollziehbare Polizeiarbeit – die Ausgangslage

In Deutschland ist Polizei Ländersache, d.h. Anforderungen an Transparenz und entsprechende Informations- und Dokumentationspflichten hängen von den Regelungen in den einzelnen Bundesländern ab. Transparenz und Bürgerorientierung gehören oftmals zum Selbstverständnis der Polizei, dem sie sich als Trägerin des staatlichen Gewaltmonopols besonders verpflichtet sieht.²⁹ Etwaige Auskunftsrechte, die z. B. für Transparenz über die Arbeit

28 Siehe Knobloch, Tobias (2018): Vor die Lage kommen: Predictive Policing in Deutschland. Chancen und Gefahren datenanalytischer Prognosetechnik und Empfehlungen für den Einsatz in der Polizeiarbeit. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/Graue_Publikationen/predictive.policing.pdf (Stand: 15.11.2019)

29 Polizei Nordrhein-Westfalen (2019): Rolle und Selbstverständnis. <https://polizei.nrw/artikel/rolle-und-selbstverstaendnis> (Stand: 15.11.2019); Polizei Rheinland-Pfalz (2019): Unser Leitbild. <https://www.polizei.rlp.de/de/die-polizei/unser-leitbild/> (Stand: 15.11.2019)

der Polizei sorgen könnten, sehen die entsprechenden Gesetze aber nicht vor.³⁰ Auch bestehende Transparenz- und Informationsfreiheitsgesetze, die eigentlich den Zugang einer Person zu amtlichen Informationen von Bundes- und Landesbehörden regeln, sehen Belange der inneren Sicherheit meist als Ausnahme, die den Zugang zu Informationen über die Polizei entsprechend einschränken können.³¹ In der Regel ist die Presse- und Öffentlichkeitsarbeit das zentrale Instrument, um polizeiliche Arbeit transparent zu machen. Hier liegt es allerdings im Ermessen der Behörde, worüber in welchem Umfang sie berichtet. Bisweilen mangelt es aber an professioneller Pressearbeit oder es wird versucht, die Öffentlichkeit aus polizeilichen Belangen herauszuhalten.³²

Dabei besteht für gewöhnlich ein sehr großes öffentliches Interesse an der Arbeit der Polizei. Transparenz und Nachvollziehbarkeit polizeilichen Handelns gehören zu den zentralen Anliegen in der Diskussion über öffentliche Sicherheit. Dies zeigen etwa Forderungen nach einer Kennzeichnungspflicht von Polizeibeamt:innen durch Namensschilder oder Dienstnummern, einer Nutzung von Bodycams auch zum Nachhalten polizeilichen Verhaltens oder nach umfassenderen Berichts- und Rechenschaftspflichten von Polizist:innen über den Einsatz von Zwangsmitteln.³³ Transparenz und Nachvollziehbarkeit in der Polizeiarbeit schaffen größere Offenheit in einem Bereich, der Kernbestandteil unseres Rechtsstaats ist, aber für Bürger:innen oft undurchsichtig bleibt. Transparenz ist wichtig, um das Machtgefälle zwischen Bürgern und Polizei auszugleichen und damit Akzeptanz und Vertrauen in die Polizei und den Rechtsstaat zu stärken. Daneben können sich auch ganz konkrete Vorteile für die Polizei ergeben: Gerade bei der Einbruchsprävention ist die Polizei mitunter auf die Wachsamkeit von Bürger:innen angewiesen. Hier spielt Nachvollziehbarkeit eine große Rolle, um Bürger:innen zu sensibilisieren, die soziale Kontrolle und auf diese Weise das Informationsaufkommen für die Polizei zu erhöhen. Auch in den USA gibt es Beispiele, die zeigen, dass Transparenz dabei helfen kann, Bürger:innen in die Kriminalitätsprävention einzubinden.³⁴

30 Ausnahmen sind Anfragen darüber, welche Daten Polizeibehörden über eine Person gespeichert haben. Siehe Gruber, Angela (2018): Was weiß die Polizei über mich? In: Der Spiegel. 02.05.2018 <https://www.spiegel.de/netzwelt/web/polizei-datenbanken-welche-infos-stehen-da-ueber-mich-a-1205527.html> (Stand: 14.11.2019)

31 Siehe <https://transparenzranking.de/> oder konkret in Hessen Semsrott, Arne (2018): Hessen: Schwarzgrün beschließt schlechtestes Informationsfreiheitsgesetz Deutschlands. In: netzpolitik.org. 25.04.2018. <https://netzpolitik.org/2018/hessen-schwarz-gruen-beschliesst-schlechtestes-informationsfreiheitsgesetz-deutschlands/#spendenleiste> (Stand: 13.11.2019)

32 Diehl, Jörg (2016): Die Verunsicherten. In: Der Spiegel. 02.03.2016. <https://www.spiegel.de/panorama/justiz/polizei-in-der-oeffentlichkeitsarbeit-mangelt-es-an-professionalitaet-a-1080303.html> (Stand: 14.11.2019)

33 Steinke, Ronen (2019): Polizisten haben Namen. In: Süddeutsche Zeitung. 12.03.2019. <https://www.sueddeutsche.de/politik/rechtsstaat-polizisten-haben-namen-1.4364361> (Stand: 18.11.2019); Reuter, Markus (2019): Nie Täter, immer Opfer: Die Polizei muss endlich lernen, mit Kritik umzugehen. In: Netzpolitik.org. 17.09.2019. <https://netzpolitik.org/2019/nie-taeter-immer-opfer-die-polizei-muss-endlich-lernen-mit-kritik-umzugehen/> (Stand: 15.11.2019); Winkler, Daniela/Zeccola, Marc (2019): Keine Stigmatisierung, sondern eine Selbstverständlichkeit. In: Legal Tribune Online. 25.09.2019. <https://www.lto.de/recht/hintergruende/h/bverwg-2-c-32-18-polizei-gewalt-kennzeichnungspflicht-buerger/> (Stand: 18.11.2019); Lupke, Sebastian (2019): PM: Transparenz im rechtsstaat ist keine Einbahnstraße. In: Amnesty International Sachsen. 12.03.2019. <https://amnesty-sachsen.de/2019/03/pm-transparenz-im-rechtsstaat-ist-keine-einbahnstrasse/> (Stand: 18.11.2019)

34 Hustedt, Carla (2016): A Public Value Perspective on Predictive Policing in the US. https://www.researchgate.net/publication/327155597_A_Public_Value_Perspective_on_Predictive_Policing_in_the_US (Stand: 18.11.2019)

Predictive Policing: neue Ansprüche an Nachvollziehbarkeit?

Die Einführung von ADM-Systemen in einem Bereich, der traditionell wegen fehlender Transparenz und Offenheit in der Kritik steht und verdächtigt wird, bürgerliche Freiheiten zugunsten der Sicherheit einzuschränken, birgt das Potenzial, Ängste zu wecken und zusätzliches Misstrauen zu schaffen. Daher wird insbesondere der Einsatz von digitalen Technologien durch Strafverfolgungs- und Sicherheitsbehörden meist kritisch vonseiten der Zivilgesellschaft und Medien begleitet. Transparenz und Nachvollziehbarkeit sind deswegen gerade hier besonders wichtig. Wenn der Einsatz von ADM-Systemen erst nach und nach öffentlich würde, ginge Vertrauen weiter verloren.

Transparenz und Nachvollziehbarkeit zu schaffen, ist auch deswegen von Bedeutung, weil in der gesellschaftlichen Debatte zu dem Thema häufig nicht ausreichend differenziert wird: Diejenigen Formen von Predictive Policing, die eine ganz konkrete Diskriminierungsgefahr bergen, weil sie personenbezogene Daten verwenden und auf die Identifikation potenzieller Straftäter abzielen, kommen in Deutschland bislang nicht zum Einsatz. Diese fehlende Trennschärfe kann weitere Ängste schüren. Deswegen ist es wichtig, offen mit dem Einsatz von Predictive Policing umzugehen und Bürger:innen aufzuklären. Mehr Transparenz kann dabei helfen, Predictive-Policing-Systeme zu entmystifizieren sowie irreführende Meldungen, Spekulationen und damit „Hollywood-Vorstellungen“ von Predictive Policing zu verhindern. Allerdings dringt auch das in Deutschland praktizierte ortsbezogene Predictive Policing in die Privatsphäre von Bürgern ein und kann diskriminierende Muster verstärken. Etwa wenn die erstellten Prognosen dazu führen, dass eine Wohngegend häufiger unter polizeilicher Beobachtung steht und Personen mit höherer Wahrscheinlichkeit als verdächtig registriert werden, nur weil sie sich an einem errechneten Kriminalitätshotspot aufhalten. Durch das Gefühl des Überwachtwerdens kann zudem das eigene Sicherheitsgefühl sinken, da insgesamt auf ein erhöhtes Kriminalitätsrisiko geschlossen wird.

Grundsätzlich besteht hinsichtlich der Schaffung von Transparenz und Nachvollziehbarkeit hier die Herausforderung, dass neue Ängste geweckt werden, wenn Bürger:innen beispielsweise erfahren, dass ihre Wohngegend ein erhöhtes Kriminalitätsrisiko aufweist. Bürger:innen müssen solche Informationen einordnen und verstehen können. Das ist insbesondere vor dem Hintergrund eine Herausforderung, dass Predictive-Policing-Systeme Wahrscheinlichkeiten errechnen, die mitunter schwierig zu interpretieren und zu vermitteln sind. Dabei muss auch mitgedacht werden, dass Transparenz gesellschaftliche Debatten und Prozesse in Gang bringen kann. Zum Beispiel, wenn eine schlechte Prognose für einen Stadtteil ohne Einordnung veröffentlicht und damit zu einer Abwertung des Gebiets führen würde.

Nicht zu unterschätzen ist der Zeitaufwand, der notwendig ist, um Nachvollziehbarkeit zu gewährleisten und aufrechtzuerhalten. Da es zunächst keine direkte Interaktion zwischen Polizeibehörden und betroffenen Bürger:innen

gibt, müssen spezielle Formate gefunden und aktive Kommunikationsarbeit betrieben werden. Außerdem müssen Informationen regelmäßig aktualisiert werden, da fortwährend an den Systemen gearbeitet wird und sich deren Funktionsweisen möglicherweise ändern.

Eine weitere Schwierigkeit ist, dass insbesondere die Frage der Wirkung von Predictive Policing für die Akzeptanz in der Bevölkerung wichtig ist und deswegen auch darüber informiert werden sollte. Hier stellt es eine grundsätzliche Herausforderung dar, dass die Effektivität von prognosebasierter Polizeiarbeit schwer messbar und ihr potenzieller Nutzen schwer kommunizierbar ist.

Auch bei Predictive Policing kommt das Risiko des „Gaming“ hinzu. Je nachdem welche Informationen über das ADM-System wann bereitgestellt werden, könnte die größere Nachvollziehbarkeit von potenziellen Tätern ausgenutzt werden. Kriminelle Aktivitäten würden sich entsprechend in Bereiche verlagern, die nicht unter besonderer polizeilicher Aufmerksamkeit stehen.

Wege für mehr Nachvollziehbarkeit bei Predictive Policing

Um Predictive Policing für die Bürger:innen einer Stadt oder Region, in der die entsprechenden Systeme eingesetzt werden, nachvollziehbar zu machen, ist es in einem ersten Schritt wichtig darüber zu informieren, dass diese Verfahren eingesetzt werden. Im Sinne des Gemeinwohls ist es aber auch von zentraler Bedeutung zu erklären, wie das Predictive-Policing-System funktioniert und warum es genutzt wird: Was ist dessen Ziel und Zweck und wofür wird es explizit nicht genutzt? Welche Daten werden genutzt? Wer hat das System entwickelt? Wer ist verantwortlich? Da für Bürger:innen auch die Frage der Wirkung bedeutsam ist, sich hier aber Probleme hinsichtlich der Messbarkeit ergeben, ist es auch wichtig, dass zukünftig belastbare Evaluationen vorangetrieben und entsprechende Nachweise im Sinn einer Qualitätstransparenz vorgelegt werden.³⁵ Sollten Funktionsweisen und Einsatzzwecke ausgeweitet werden, muss dies ebenfalls transparent gemacht werden. Wenn beispielsweise personenbezogenen Analysen zum Einsatz kommen, müssen zusätzlich individuelle Informationen über die erstellten Prognosen zur Verfügung gestellt werden.³⁶

Da sich (bei ortsbezogenen Systemen) die Prognosen auf Städte und Regionen beziehen und Bürger als Betroffene nicht direkt mit dem ADM-System interagieren, müssen Formate gefunden werden, die möglichst viele Personen erreichen. Dafür ist v. a. eine umfassende Presse- und Öffentlichkeitsarbeit

³⁵ Unter anderem zum Thema Wirksamkeit und Notwendigkeit von Evaluation: Egbert, Simon/Krassmann, Susanne (2019): Predictive Policing. Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die polizeiliche Praxis. Hamburg: 54.

³⁶ So wurde bspw. in Hamburg ein neues Polizeigesetz diskutiert, das personenbezogene Anwendungen ermöglichen würde. Siehe Bröckling, Marie (2019): Juristinnen kritisieren „Palantir-Paragraf“ im geplanten Polizeigesetz. In: Netzpolitik.org. 24.09.2019. <https://netzpolitik.org/2019/hamburg-juristinnen-kritisieren-palantir-paragraf-im-geplanten-polizeigesetz/> (Stand: 15.11.2019)

zentral. Wenn aktive Kommunikationsarbeit bislang nicht zur Arbeitsweise von Polizeibehörden gehörte, muss hier ein Umdenken stattfinden. Denkbare Formate sind z. B. Webseiten, Flyer, Apps, aber auch Informationsveranstaltungen in den betroffenen Nachbarschaften. Im Rahmen des Predictive-Policing-Projekts SKALA des Landeskriminalamts Nordrhein-Westfalen (LKA NRW) wird beispielsweise gemeinsam mit Fraunhofer FOKUS (Fraunhofer-Institut für offene Kommunikationssysteme) an einer App gearbeitet, über die sich die Öffentlichkeit informieren kann. Über Push-Nachrichten wird mitgeteilt, ob die Wahrscheinlichkeit erhöht ist, dass in der eigenen Wohngegend verstärkt mit Einbrüchen zu rechnen ist. Die App soll außerdem erklären, wie die Prognosen grundsätzlich erstellt werden, mit welchen Daten gearbeitet wird (Kriminalitäts- und Soziostrukturdaten), was grundsätzlich eine Prognose ist und was die Wahrscheinlichkeiten bedeuten. Darüber hinaus wurden in Nordrhein-Westfalen auch die Projektberichte über SKALA veröffentlicht.

Es ist außerdem wichtig, die Verantwortung für die Schaffung von Transparenz und Nachvollziehbarkeit klar festzulegen. Im Zuge von SKALA wurden für die externe Öffentlichkeitsarbeit eigene Konzepte entwickelt und konkrete Ansprechpartner:innen berufen. Es kommt auch darauf an, offene Kommunikationskanäle zu schaffen, an die sich Bürger:innen mit Fragen wenden können, und Mechanismen für Beschwerden einzurichten. Eine Möglichkeit könnte es sein, eine Position ähnlich der des Polizeibeauftragten zu schaffen, die als direkte Ansprechperson für Bürger:innen in Belangen rund um Predictive Policing fungiert. Aber auch die Polizeibeamt:innen auf der Straße sollten in der Lage sein, das Predictive-Policing-System im Dienst zu erklären und zu versuchen, mit Bürger:innen darüber ins Gespräch zu kommen.

Dass sich die Frage, wie Nachvollziehbarkeit ermöglicht werden kann, bereits auf Ebene des eingesetzten Modells für die statistischen Prognosen entscheidet, macht ebenfalls das konkrete Vorgehen des LKA NRW deutlich. Bei der Entwicklung des Predictive-Policing-Projekts SKALA hat die Behörde verschiedene Modellansätze getestet und sich aus Gründen der besseren Nachvollziehbarkeit bewusst für die Verwendung des Modells entschieden, deren Berechnungen im Detail nachvollziehbar sind (sog. Entscheidungsbäume).³⁷

³⁷ Landeskriminalamt NRW (2018): Abschlussbericht Projekt Skala. Düsseldorf: 51ff. https://lka.polizei.nrw/sites/default/files/2018-06/180208_Abschlussbericht_SKALA.pdf (Stand: 18.11.2019)

GESUNDHEITS-APPS – TRANSPARENZ GEGEN RISIKEN UND NEBENWIRKUNGEN

Im Gesundheitsbereich ist der Einsatz von algorithmischen Systemen nichts Neues.³⁸ Mittlerweile existiert darüber hinaus eine Vielzahl von mobilen Anwendungen, die auch Personen ohne medizinischen Hintergrund datenbasiert bei der Prävention, Diagnose, Therapie oder der Überwachung von Krankheiten unterstützen. Beispielsweise können Apps mithilfe von Algorithmen durch die Erfragung von Symptomen und Laborwerten automatisiert Verdachtsdiagnosen erstellen. Fast jeder zweite Deutsche nutzte bereits 2017 eine Gesundheits-App, Tendenz stark steigend.³⁹

Nachvollziehbarkeit zwischen Arzt/Ärztin und Patient:in – die Ausgangslage

In Deutschland sind Ärzt:innen an umfangreiche Informations- und Aufklärungspflichten gebunden, die für Transparenz und Nachvollziehbarkeit in der Behandlung von Patient:innen sorgen sollen. Sie sind zum einen seit Inkrafttreten des Patientenrechtegesetzes im Februar 2013 im Bürgerlichen Gesetzbuch (BGB) geregelt. Zum anderen gehört die Aufklärung von Patient:innen zur Berufspflicht von Ärzt:innen, wie sie in den Berufsordnungen der Ärztekammern niedergelegt ist. Patient:innen sollen laut Gesetz in die Lage versetzt werden, sich der wesentlichen Umstände einer Behandlung bewusst zu sein (§ 630c BGB). Dazu gehört es, dass die Information in verständlicher Form und umfassend erfolgt, d. h. die Diagnose, beabsichtigte Therapie und die voraussichtliche gesundheitliche Entwicklung erläutert werden.⁴⁰ Die Verständlichkeit bezieht sich gemäß des Patientenrechtegesetzes auf die inhaltliche und sprachliche Verständlichkeit, was z. B. bedeutet, dass nicht zu viele Fachbegriffe verwendet werden und u. U. eine leicht verständliche Sprache gewählt werden muss. Ein konkreter ärztlicher Eingriff setzt ein Aufklärungsgespräch voraus, in dem die Art der Maßnahme, ihre Durchführung, zu erwartende Folgen, Risiken, Notwendigkeit, Eignung, Erfolgsaussichten und alternative Behandlungsmethoden dargelegt werden. Zudem muss solch ein Aufklärungsgespräch mündlich, persönlich und rechtzeitig vor einem Eingriff erfolgen. Wie umfangreich und intensiv die Information und Aufklärung zu erfolgen hat, entscheidet sich im Einzelfall. Es gibt die Möglichkeit, Aufklärungsbögen und Informationsblätter auszuhändigen. Sie ersetzen allerdings nicht das mündliche Gespräch, in dem die/der Ärztin/Arzt sicherstellen kann, dass Patient:innen alles verstanden haben und keine Fragen mehr haben. Das Leitbild ärztlicher Informations- und

38 Für eine Übersicht über Anwendungsbereiche von Algorithmen im Gesundheitsbereich, siehe Friele, Minou/Jannes, Christiane/Jannes, Marc/Woopen, Christiane (2018): Algorithmen in der digitalen Gesundheitsversorgung. Eine interdisziplinäre Analyse. <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/algorithmen-in-der-digitalen-gesundheitsversorgung/> (Stand: 14.11.2019)

39 Für eine Einführung in das Thema der Situation in Deutschland und einen Überblick über gesellschaftliche Chancen und Risiken von Gesundheitsanwendungen, siehe Klingel, Anita (2019): Gesund dank Algorithmus. Chancen und Herausforderungen von Gesundheits-Apps für Patient:innen. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Impulspapier_Gesund-Dank-Algorithmen.pdf (Stand: 13.11.2019)

40 Unabhängige Patientenberatung Deutschland (UPD) (2019): Ärztliche Informations- und Aufklärungspflichten. 20.08.2019. <https://www.patientenberatung.de/de/recht/patientenrechte/aerztliche-informations-und-aufklaerungspflichten> (Stand: 15.11.2019)

Aufklärungspflichten ist das der/des selbstbestimmen Patient:in, die/der selbst über den Fortgang der Behandlung entscheiden kann. Transparenz und Kommunikation zwischen Ärztin/Arzt und Patient:in wird als wichtig erachtet, um Vertrauen herzustellen, mit dem Ziel, die Therapie und Heilung einer/eines Patient:in voranzubringen.

Die richtige Art und den Umfang der Information zu finden ist in der Praxis eine sehr individuelle Angelegenheit, die eine Menge Fingerspitzengefühl verlangt und oft eine Frage des Ermessens und der eigenen Erfahrung ist. Für die Zufriedenheit von Patient:innen ist eine partnerschaftliche Kommunikation vonseiten der Ärzteschaft wichtig, das bedeutet z. B., dass sie sich für die persönliche Sicht der Patient:innen auf das Problem interessiert oder sie ermutigt, selber Fragen zu stellen. Da Informationspflichten und -erwartungen in der Praxis oftmals auf Mediziner:innen treffen, die unter großem zeitlichen Druck stehen, ist fraglich, inwieweit sie diese Erwartungen und Pflichten erfüllen.

Gesundheits-Apps: neue Ansprüche an Nachvollziehbarkeit?

Durch die Einführung von Gesundheitsanwendungen werden Urteile und Entscheidungen, die zuvor Ärzt:innen vorbehalten waren, auf algorithmische Systeme übertragen. Sie ermöglichen es Nutzer:innen zwar, jederzeit individuelle Informationen und Einschätzungen rund um ihre Gesundheit zu erhalten. Im Unterschied zur Versorgung durch eine/n Ärztin/Arzt, fehlt hier aber die direkte Kommunikation mit medizinischem Fachpersonal und damit zunächst der Zugang, um Sachverhalte genauer erklärt und eingeordnet zu bekommen. Darüber hinaus können Patient:innen, die Gesundheits-Apps nutzen, häufig auch nicht einschätzen, wie das zugrunde liegende algorithmische System zu Therapieempfehlungen kommt oder wie zuverlässig Diagnosen auf Grundlage der verwendeten Daten sind. Die Gefahr besteht, dass Gesundheits-Apps nicht richtig genutzt werden und z. B. unpassenden Therapieempfehlungen gefolgt oder die Diagnostik falsch interpretiert wird. Bei Gesundheitsanwendungen sind Transparenz und Nachvollziehbarkeit daher wichtig, um eine richtige und angemessene Anwendung der App durch die Nutzer:innen sicherzustellen.

Es ist aber nicht nur wichtig, die Aussagen und Einschätzungen einer Gesundheits-App inhaltlich einordnen zu können. Für die Nutzung kommt es auch darauf an, die Qualität einer App richtig bewerten zu können. Während die Ausübung des Arztberufs an verschiedenste Standards gebunden ist (z. B. Medizinstudium, regelmäßige Fortbildungen etc.), die die Qualität im Gesundheitssektor grundsätzlich sicherstellen, ist die Qualität von Gesundheitsanwendungen für Nutzer:innen bislang nur schwer einzuschätzen. Das betrifft auch insbesondere die Frage der Datensicherheit und des Datenschutzes: Mit der Verwendung von Gesundheitsanwendungen ist das Risiko verbunden, dass sehr persönliche Daten beispielweise ungesichert ins Netz gelangen oder für andere, von den Nutzer:innen nicht beabsichtigte Zwecke genutzt werden. Hier muss dafür gesorgt, dass Nutzer:innen die Kontrolle darüber behalten, wofür ihre Daten genutzt werden und wofür

nicht. Fehlen Transparenz und Nachvollziehbarkeit, lassen sich sinnvolle und nützliche Anwendungen nur schwierig von schlechten Apps unterscheiden. Nutzer:innen können dann keine informierte Entscheidung mehr treffen und wählen am Ende möglicherweise eine Anwendung, die ihnen und ihrer Gesundheit potenziell sogar schadet.

Transparenz und Nachvollziehbarkeit bei algorithmenbasierten Gesundheits-Apps erlauben eine kritische Auseinandersetzung mit ADM-Systemen, die tief in die persönliche Lebensführung eingreifen, und lassen die Grenzen der Systeme erkennen. Insgesamt können Transparenz und Nachvollziehbarkeit so dazu beitragen, dass hilfreiche Gesundheitsanwendungen breiter genutzt werden und damit die Gesundheitsversorgung unabhängig von ärztlicher Betreuung gestärkt wird. Mangelt es hingegen an Transparenz und Nachvollziehbarkeit, kann es Nutzer:innen abschrecken, eine App zu verwenden und die Chancen, die mobile Gesundheitsanwendungen bieten, bleiben ungenutzt.

Eine Herausforderung ist dabei, dass Gesundheitsanwendungen, anders als Ärzte, den Kontext und die Bedürfnisse der Patient:innen bzw. der Nutzer:innen nicht kennen. Ärzt:innen wissen, in welcher physischen und psychischen Situation sich jemand befindet, und können aufgrund ihres Fachwissens und ihrer Erfahrung abschätzen, wann ein guter Moment z. B. für die Mitteilung einer Diagnose ist und welche Informationen zusätzlich oder möglicherweise zu einem anderen Zeitpunkt noch einmal vermitteln werden müssen. Wie Patient:innen haben auch Nutzer:innen von Gesundheits-Apps unterschiedliche Bedarfe und Erwartungen, die ADM-Systeme aber nicht zwangsläufig kennen. Während Ärzt:innen zudem je nach Situation entscheiden, wie sie mit ihrem Gegenüber kommunizieren, müssen bei Gesundheits-Apps bereits vorab Festlegungen getroffen werden und eine Personalisierung ist nur begrenzt möglich. Zusätzlich muss sichergestellt werden, dass Nutzer:innen den Empfehlungen von Gesundheits-Apps kein blindes Vertrauen entgegenbringen, denn am Ende ersetzen sie schließlich keine ärztliche Versorgung und Beratung.

Wege zu mehr Transparenz und Nachvollziehbarkeit bei Gesundheits-Apps

Im Falle von algorithmenbasierten Gesundheits-Apps sind die Betroffenen zugleich die Nutzer:innen des algorithmischen Systems. Im Vergleich zu den zwei anderen Anwendungsbereichen ergeben sich daraus andere Anforderungen an Nachvollziehbarkeit und andere Möglichkeiten, Nachvollziehbarkeit zu schaffen, da es eine Benutzeroberfläche gibt, die für die Informationsvermittlung genutzt werden kann.

Um die Funktionalität und Qualität der Anwendung verstehen und bewerten zu können, sollten Nutzer:innen zum einen erfahren, wie das zugrunde liegende ADM-System die Ergebnisse und Empfehlungen der App generiert. Das schließt zum einen Informationen über die allgemeine Funktionsweise der App ein, z. B. auf welcher Datenbasis das System trainiert und wie die Anwendung entwickelt wurde. Da Nachvollziehbarkeit im Falle von

Gesundheitsanwendungen insbesondere wichtig ist, um eine angemessene Anwendung sicherzustellen, sollten Nutzer:innen auch darüber informiert werden, was sie von der App erwarten können und was nicht. Grenzen können aufgezeigt werden, indem Wahrscheinlichkeiten klar kommuniziert werden und deutlich gemacht wird, dass es sich um Prognosen oder Empfehlungen und nicht um gesicherte Vorhersagen handelt. Zudem könnte auf Fehlerraten hingewiesen oder explizit empfohlen werden, ärztlichen Rat hinzuzuziehen. Wichtig sind zum anderen Auskünfte, auf welche Daten die App bei der Nutzung zurückgreift (z. B. nur die durch Selbsteingabe generierten oder andere?) und welche Daten für die Erstellung einer Diagnose o. Ä. besonders relevant waren, damit Nutzer:innen einschätzen können, wie die individuellen Ergebnisse zustande gekommen sind. Eine in Deutschland verbreitete App zur Symptomanalyse zählt die möglichen Ursachen der Symptome auf und ergänzt sie jeweils durch eine nähere Beschreibung der Ursachen, eine Empfehlung über nächste Schritte und eine Visualisierung der Wahrscheinlichkeiten sowie welche der eingegebenen Symptome für die Analyse relevant waren.

Um Vertrauen in die App zu generieren, ist es darüber hinaus wichtig, in verständlicher Form mitzuteilen, wie die persönlichen Daten einer/eines Nutzer:in behandelt werden. Zudem ist es sinnvoll, Auskünfte über die Einhaltung bestimmter Gütekriterien, z. B. im Rahmen von Zertifizierungsprozessen, zu erteilen und auch bestehende Evaluationen in verständlicher Form zur Verfügung zu stellen.

6. KEIN „ONE SIZE FITS ALL“ – EINE CHECKLISTE FÜR MEHR NACHVOLLZIEHBARKEIT VON ADM

Die öffentliche Debatte über Algorithmenregulierung geht noch selten über die abstrakte, nicht weiter ausdifferenzierten Forderung nach Transparenz und Nachvollziehbarkeit hinaus. An den drei Praxisbeispielen in Abschnitt 5 wird klar, dass die Ausgangslage und damit die Anforderungen an Transparenz und Nachvollziehbarkeit in verschiedenen Bereichen der Gesellschaft sehr unterschiedlich sind. Diese Vielfalt macht es notwendig, für jeden Anwendungsbereich spezifisch zu justieren, was wie nachvollziehbar gemacht werden muss. Die folgende Checkliste gibt als Fragenkatalog einen Überblick, welche Aspekte besonders berücksichtigt werden müssen, wenn ADM-Prozesse für Betroffene nachvollziehbar gemacht werden sollen.

- Wie hoch ist die Teilhaberelevanz?

Zunächst sollte die Höhe der Teilhaberelevanz einer durch ADM getroffenen Entscheidung einen Einfluss auf das Maß an Transparenz und Nachvollziehbarkeit haben. In vielen alltäglichen Situationen, in denen ADM-Systeme zum Einsatz kommen (z.B. Navigation), sind die Anforderungen an Transparenz und Nachvollziehbarkeit deutlich geringer anzusetzen als in teilhaberelevanten Bereichen. Zum Beispiel überall dort, wo sich Menschen einem ADM-System nicht entziehen können und/oder in besonderem Maße Potenzial für Diskriminierung und Risiken für Leib und Leben bestehen – wie etwa in den exemplarisch vorgestellten Anwendungsbereichen, in denen es um Zugang zum Arbeitsmarkt, um die bestmögliche Gesundheitsversorgung oder die Ausübung von Staatsgewalt geht.⁴¹

- Was soll mit Transparenz erreicht werden?

In den dargestellten Fallbeispielen wurde deutlich, dass vielfältige Chancen mit Nachvollziehbarkeit verbunden sind. Damit Nachvollziehbarkeit nicht als reiner Selbstzweck endet, sondern einen tatsächlichen Nutzen bringt, sollten Stellen, die ADM einsetzen, für sich konkretisieren, welche Funktionen und Ziele Transparenz erreichen soll. Da in den verschiedenen Bereichen verschiedene Rechte und Interessen der Betroffenen berührt werden, sind auch Ziele und somit Form und Umfang der geforderten Transparenz

⁴¹ Zur Bestimmung der Teilhaberelevanz siehe Vieth, Kilian/Wagner, Ben (2017): Teilhabe, ausgerechnet. Wie algorithmische Prozesse Teilhabechancen beeinflussen können. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Teilhabe_ausgerechnet.pdf (Stand: 14.11.2019). Das Gutachten der Datenethikkommission schlägt eine entsprechende Überprüfung der Systemkritikalität vor, siehe Datenethikkommission (2019). Gutachten der Datenethikkommission. https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_DE.pdf?__blob=publicationFile&v=2 (Stand: 15.11.2019)

jeweils stark unterschiedlich.⁴² Beispielsweise ist es denkbar, dass Vertrauen in Predictive-Policing-Systeme der Polizei vor allem voraussetzt, dass Bürger:innen genau wissen, welche Daten in das System einfließen, wofür sie eingesetzt werden und wofür nicht. Bei Gesundheits-Apps geht es v. a. darum, das richtige Maß an Vertrauen in die Qualität der automatisiert generierten Ergebnisse zu schaffen, um eine richtige Anwendung sicherzustellen. Hier gilt es, genau zu untersuchen, um welche Ziele es in den Sektoren jeweils geht und welche Form von Transparenz benötigt wird, um das jeweilige Ziel zu erreichen.

- Welchen Kontakt hat die/der Betroffene zum algorithmischen System?

Bürger:innen, Bewerber:innen und Nutzer:innen von Gesundheits-Apps kommen alle in ganz unterschiedlicher Weise mit ADM-Systemen in Kontakt. Während die Gesundheits-App direkt durch die Betroffenen genutzt wird und somit direkt mit ihnen kommuniziert werden kann, ist dies bei Recruiting-Systemen (z. B. bei Nutzung von Onlinebewerbungstools) nur eingeschränkt und bei ortsbezogenem Predictive Policing gar nicht der Fall. Das bedeutet, dass insbesondere bei ADM-Prozessen ohne direkte Interaktion mit den Betroffenen über spezielle Formate nachgedacht werden muss, um Transparenz und Nachvollziehbarkeit zu schaffen.

- Wird eine Entscheidung über eine ganze Gruppe getroffen oder spezifisch über ein Individuum?

Auch sind algorithmische Entscheidungen in stark unterschiedlichem Maße spezifisch für die jeweiligen Betroffenen: Während bei Gesundheits-Apps und Recruiting-Systemen für jede Einzelperson spezifische Entscheidungen getroffen werden, handelt es sich bei ortsbezogenem Predictive Policing um ganze Gruppen an Individuen, die alle gleichermaßen betroffen sind. Die Informationen zur Herstellung von Nachvollziehbarkeit sollten demnach immer entsprechend spezifisch gehalten werden. Während bei einem Bewerbungsverfahren die individuelle Sprachkenntnis für den Erfolg ausschlaggebend sein und dies den Bewerber:innen mitgeteilt werden kann, könnte bei ortsbezogenem Predictive Policing z. B. die Einkommensstruktur eines Stadtteils die wichtigste Rolle spielen, worüber alle Bewohner:innen gleichermaßen aufzuklären wären.

- Wie viel Zeit, Wissen und Bereitschaft bringen Betroffene mit, um sich mit dem ADM-Prozess auseinanderzusetzen?

Wenn es darum geht, algorithmische Entscheidungen für diejenigen nachvollziehbar zu machen, die von ihnen betroffen sind, gilt es auch, genau in den Blick zu nehmen, wer diese Personen sind. Wie viel Wissen, Zeit und Bereitschaft bringen sie mit, um sich mit einem ADM-Prozess

42 Van Drunen, M.Z./Helberger, N./Bastian, M (2019): Know your algorithm: what media organizations need to explain to their users about news personalization. In: International Data Privacy Law, ipz011. <https://academic.oup.com/idpl/advance-article-abstract/doi/10.1093/idpl/ipz011/5544759?redirectedFrom=fulltext> (Stand: 15.11.2019)

auseinanderzusetzen? Insbesondere da die Gefahr besteht, dass Informationsangebote ins Leere laufen und keinen Mehrwert für Nutzer:innen oder Betroffene bringen (siehe Abschnitt 4), ist der Fokus auf die Bedürfnisse der Adressaten, an die sich die Erläuterungen richten sollen, von zentraler Bedeutung. In Situationen, in denen für jemanden persönlich viel auf dem Spiel steht oder Ergebnisse von ADM nicht den eigenen Erwartungen entsprechen, wie z. B. im Rahmen eines Bewerbungsverfahrens, haben Betroffene in der Regel ein größeres Interesse und mitunter den Wunsch, sich vertieft mit den Hintergründen eines ADM-Prozesses zu beschäftigen. Insgesamt muss es Betroffenen aber immer so einfach wie möglich gemacht werden, wichtige Informationen schnell erfassen zu können.

- Welche regulatorischen Anforderungen gibt es?

Außerdem muss immer geprüft werden, welche regulatorischen Transparenzanforderungen und Vorgaben bestehen, die den Umfang und Inhalt von Informationsangeboten in konkreten Anwendungsbereichen bestimmen. Der Text der Datenschutz-Grundverordnung (DSGVO) fordert bspw. „ausagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen“. Zudem werden Ansprüche formuliert, wie Informationen zur Verfügung gestellt werden müssen, nämlich in „verständlicher und leicht zugänglicher Form“ und „in einer klaren einfachen Sprache“.⁴³

Eine Checkliste für mehr Nachvollziehbarkeit von algorithmischen Systemen

- ? → Wie hoch ist die Teilhaberelevanz?

- ? → Was soll mit Transparenz erreicht werden?

- ? → Welchen Kontakt hat die/der Betroffene zum algorithmischen System?

- ? → Wird eine Entscheidung über eine ganze Gruppe getroffen oder spezifisch über ein Individuum?

- ? → Wie viel Zeit, Wissen und Bereitschaft bringen Betroffene mit, um sich mit dem ADM-Prozess auseinanderzusetzen?

- ? → Welche regulatorischen Anforderungen gibt es?

Quelle: eigene Darstellung (Leonie Beining & Carla Hustedt)
 Stiftung
 Neue
 Verantwortung | BertelsmannStiftung

⁴³ Siehe DSGVO; Artikel 12 bis 15. Absatz 7 bringt sogar die Verwendung von Bildsymbolen ins Spiel, um den Anforderungen nachzukommen. Der - nicht bindende, aber für die Auslegung der DSGVO möglicherweise wichtige - Erwägungsgrund 71 sieht vor, dass Betroffene Anspruch auf eine Erläuterung einer Entscheidung haben, vgl. <https://dsgvo-gesetz.de/>.

7. EMPFEHLUNGEN

Damit aus den allgemeinen Forderungen nach Transparenz und Nachvollziehbarkeit wirksame Instrumente werden, die es Menschen ermöglichen, sich mit sie betreffenden algorithmischen Systemen auseinanderzusetzen, bleibt noch viel zu tun. Zukünftig wird es darum gehen, Ansätze weiter zu konkretisieren und sie anschließend auch umzusetzen. Dabei besteht für unterschiedliche Seiten Handlungsbedarf. Sowohl politische Entscheidungsträger:innen, Stakeholder aus der Wissenschaft und Gesellschaft als auch Anwender:innen teilhaberelevanter algorithmischer Systeme sollten verschiedene Aspekte im Blick behalten:

Frühzeitig und proaktiv kommunizieren

Für Anwender:innen von teilhaberelevanten algorithmischen Systemen gilt grundsätzlich, dass sie so früh wie möglich offen und proaktiv über den Einsatz von ADM informieren sollten. Eine zögerliche und passive Kommunikation von Zielen, Zwecken und Funktionsweisen algorithmischer Systeme ist mit einer gemeinwohlorientierten Gestaltung algorithmischer Entscheidungsprozesse nicht vereinbar. Wird zu spät kommuniziert, kann das zudem unnötiges Misstrauen und Ängste auslösen, die einem grundsätzlich nutzbringenden Einsatz von ADM im Wege stehen.

Verantwortlichkeiten klar benennen und Ressourcen bereitstellen

Die Schaffung von Nachvollziehbarkeit für Betroffene kann mitunter ressourcenintensiv sein. Das gilt insbesondere dann, wenn der Kontakt zu Betroffenen aktiv gesucht werden muss und spezielle Kommunikationsformate entwickelt werden müssen (wie z. B. bei Predictive Policing). Daher müssen Verantwortlichkeiten für Transparenz und Nachvollziehbarkeit klar benannt sowie entsprechende Mittel und Personal bereitgestellt werden. Wenn zukünftig Vorgaben zu Algorithmentransparenz umgesetzt werden müssen, darf dies nicht als reine Compliance-Angelegenheit begriffen werden, die von der Rechtsabteilung der Anbieter- und Anwenderunternehmen abgehandelt wird. Unternehmen, die ADM entwickeln und einsetzen, sollten die Schaffung von Transparenz und Nachvollziehbarkeit vielmehr als Chance für Akzeptanz und als möglichen Wettbewerbsvorteil begreifen.

Bedürfnisse von Betroffenen in den Blick nehmen

Bei der Entwicklung von Informations- und Kommunikationsformaten müssen die Bedürfnisse der Zielgruppe identifiziert werden. Ansätze wie z.B. Legal Design oder Projekte rund um Usable Privacy sollten in Betracht gezogen werden, um geeignete und effektive Informationsangebote für Betroffene zu schaffen.⁴⁴ Dabei ist es von großer Bedeutung, auch die

⁴⁴ Legal Design meint die Kombination von Design-Thinking-Methoden und Recht, um bspw. juristische Information wie Verträge verständlicher darzustellen oder aber um neue Dienstleistungen und Produkte im Bereich des Rechts zu entwickeln und dabei den Fokus auf die Bedürfnisse der Nutzer:innen zu legen. Für einen Überblick siehe <http://eitlab.eu/resources/>. Usable Privacy umfasst grundsätzlich benutzerfreundliche Verfahren zur Förderung von Privatsphäre, bspw. Projekte zur benutzerfreundlichen Umsetzung von Datenschutzerklärungen, siehe z. B. <https://www.usableprivacy.org/>.

Bedürfnisse besonderer Zielgruppen, wie Kindern oder Menschen mit Behinderungen, mitzudenken.

Nachvollziehbarkeit bei Anbietern von ADM-Systemen einfordern

Unternehmen und öffentliche Stellen, die den Einsatz von ADM-Systemen planen, sollten im Rahmen des Einkaufs entsprechender Dienstleistungen und der Vergabe von Aufträgen sicherstellen, dass sich der Aspekt Nachvollziehbarkeit für Betroffene in den Vergabestandards widerspiegelt. So wird ausgeschlossen, dass z. B. Geschäftsgeheimnisse von Softwareanbieter:innen die Schaffung von Nachvollziehbarkeit beeinträchtigen. Grundsätzlich muss gelten, dass Anwender:innen immer so einfache ADM-Systeme wie möglich einsetzen sollten, um die Schaffung von Nachvollziehbarkeit von vorneherein zu erleichtern.

Rechtliche Anforderungen an Nachvollziehbarkeit anpassen

Aufgabe der Politik ist es, rechtliche Vorgaben zu Transparenz- und Informationspflichten zukünftig zu konkretisieren. Hier gilt es zum einen, anwendungsbereichspezifische Anforderungen an Nachvollziehbarkeit zu beachten und keine „one size fits all“-Regelungen zu formulieren. Vorgaben müssen mit Blick auf die Umsetzung konkret genug sein, um zu verhindern, dass zu allgemeine Auflagen in „ADM-Erklärungen“ münden, die zu keinem Wissenszuwachs bei den Betroffenen führen, wie es beispielsweise bei gebräuchlichen Datenschutzerklärungen häufig der Fall ist. Wie es das britische Information Commissioner Office für Großbritannien vorschlägt, so könnte auch die Politik in Deutschland die Entwicklung von Leitlinien vorantreiben, die vorgeben, wie Akteur:innen, die algorithmische Entscheidungssysteme entwickeln und anwenden, Nachvollziehbarkeit für Betroffene sicherstellen sollten.⁴⁵ Insbesondere für die Konkretisierung etwaiger Kennzeichnungspflichten ist es wichtig, Vorgaben auf Basis verhaltenswissenschaftlicher Expertise und Erfahrungen aus anderen Bereichen wie Datenschutz oder Mensch-Maschine-Interaktion zu entwickeln und zu testen.⁴⁶

Wissenschaftliche Interdisziplinarität und Austausch mit der Praxis vorantreiben

In ihrer KI-Strategie kündigt die Bundesregierung an, die Forschung zu Transparenz und Nachvollziehbarkeit voranzutreiben.⁴⁷ Dabei sollte auch die Forschung zu Nachvollziehbarkeit mit Fokus auf Betroffene von ADM intensiv gefördert werden. Hier muss beachtet werden, dass sich verschiedene Disziplinen bereits mit dem Aspekt der Nachvollziehbarkeit von automatisierten Entscheidungen und der Wirkung und Wirksamkeit von Transparenz und Erklärungen von algorithmischen Systemen beschäftigen. Stakeholder

45 Information Commissioner's Office (2019): Project explain. Interim report. <https://ico.org.uk/media/2615039/project-explain-20190603.pdf> (Stand: 18.11.2019)

46 Busch, Christoph (2018): Gutachten: Algorithmic Accountability: 59. <http://www.abida.de/sites/default/files/ABIDA%20Gutachten%20Algorithmic%20Accountability.pdf> (Stand: 14.11.2019)

47 Bundesregierung (2018): Strategie Künstliche Intelligenz der Bundesregierung: 14. https://www.bmbf.de/files/Nationale_KI-Strategie.pdf (Stand: 18.11.2019)

aus Wissenschaft und Forschungsförderung sollten daran arbeiten, dass der Austausch zwischen unterschiedlichen Forschungssträngen (v. a. Machine Learning, Mensch-Maschine-Interaktion, Psychologie, Recht) gestärkt wird, um Formate und Maßnahmen zu entwickeln, die für Betroffene tatsächlich zu mehr Nachvollziehbarkeit führen. Hier ist auch eine stärkere Zusammenarbeit zwischen Wissenschaft und Praxis gefragt, um wissenschaftliche Ansätze in die Praxis zu überführen und unter realen Bedingungen zu evaluieren.

Sozialethische Kompetenzen von bei Entwickler:innen stärken

ADM für Betroffene nachvollziehbar zu machen, setzt voraus, dass man sich bereits im Entwicklungsprozess konzeptionell mit algorithmischen Systemen auseinandersetzt und die Zusammenhänge, die das Modell abbildet, versteht und darstellen kann. Nur wenn die Ziele, der Einsatzkontext und die Auswirkungen eines ADM bereits bei der Entwicklung verstanden werden, kann dies auch für Betroffene begründet werden.⁴⁸ In entwickelnden Organisationen müssen Kompetenzen geschaffen werden, um Nachvollziehbarkeit nicht nur als technische Herausforderung zu verstehen, sondern auch gesellschaftliche Bedürfnisse hinsichtlich Nachvollziehbarkeit zu adressieren. Verantwortliche in der Wissenschaft und in der Bildungspolitik sollten dafür Sorge tragen, dass ADM-Entwickler:innen von morgen die notwendigen Kompetenzen im Rahmen aktualisierter Curricula, neuer sozio-informatischer Studiengänge und durch die Zusammenarbeit in interdisziplinären Teams erwerben können.

Algorithmic Literacy in der Bevölkerung fördern

Damit aus einem mehr an Informationen auch ein Wissenszuwachs bei Betroffenen folgt, sollten politische Akteur:innen Maßnahmen fördern, durch die das Wissen über Algorithmen in der ganzen Bevölkerung gestärkt werden kann („algorithmic literacy“). Die individuelle Auseinandersetzung mit ADM-Systemen setzt mitunter voraus, bereits für grundlegende Aspekte rund um ADM sensibilisiert zu sein. Hier können auch Akteur:innen der Zivilgesellschaft als Vermittler:innen von Wissen und Kompetenzen an die breite Bevölkerung einen wichtigen Beitrag leisten.⁴⁹

Ergänzende Ansätze im Umgang mit ADM verfolgen

Trotz der Bedeutung von Nachvollziehbarkeit für Betroffene für einen gemeinwohlorientierten Einsatz von ADM müssen politische Entscheidungsträger:innen immer im Blick behalten, dass dies nur eine Maßnahme im Umgang mit ADM ist. Zusätzlich müssen politische Akteur:innen einen geeigneten Rahmen für Kontrolle, Aufsicht und Haftung schaffen, auf den sich Betroffene verlassen können. Dabei sollten auch zivilgesellschaftliche Akteur:innen rechtlich und finanziell dazu befähigt werden,

48 Von Büna, Paul (2019): The future of AI may depend on historians and musicologists. <https://idalab.de/blog/data-science/the-future-of-ai-may-depend-on-historians-and-musicologists> (Stand: 15.11.2019)

49 Beining, Leonie (2017): Der Puls der Gesellschaft. Wie Daten und Algorithmen die Rahmenbedingungen für das Gemeinwohl verändern. <https://www.stiftung-nv.de/sites/default/files/datenundalgorithmen.pdf> (Stand: 14.11.2019)

als Watchdog-Organisationen zu fungieren oder ADM-Betroffene bei der Wahrnehmung ihrer Rechte zu unterstützen. Zudem müssen prozessuale Standards etabliert und implementiert werden, um algorithmische Systeme von vornherein gemeinwohlorientiert zu entwickeln und einzusetzen.⁵⁰ Eine darauf aufbauende Zertifizierung von algorithmischen Systemen wäre eine gute Möglichkeit, Betroffene in die Lage zu versetzen, wichtige Eigenschaften eines algorithmischen Systems auf den ersten Blick bewerten zu können, ohne sich notwendigerweise erst mit spezifischen Informationen auseinandersetzen zu müssen.

⁵⁰ <https://algorules.org/startseite/>

8. FAZIT

Algorithmische Entscheidungsprozesse auch für diejenigen transparent und nachvollziehbar zu machen, die von den Entscheidungen betroffen sind, ist wichtig, damit die Betroffenen die Funktionslogik dieser Systeme verstehen, sich kritisch mit ihnen auseinandersetzen, sie wenn nötig anfechten oder als Nutzer richtig anwenden können. Wie aber sollten Transparenz und Nachvollziehbarkeit von Algorithmen umgesetzt werden? Hier fehlt es bislang an konkreten Ideen und Vorschlägen. Es ist an der Zeit, allgemeine Forderungen nach Transparenz und Nachvollziehbarkeit zu präzisieren und sich mit Detailfragen zu beschäftigen. So bedarf es nicht nur je nach Anwendungsbereich und Nutzergruppen anderer Ansätze (siehe Abschnitt 4), sondern es bestehen auch grundsätzliche Herausforderungen (siehe Abschnitt 6), die stärker in den Blick genommen werden müssen.

Klar ist, dass die Schaffung von Transparenz und Nachvollziehbarkeit nicht die alleinige Lösung für die vielfältigen gesellschaftlichen Herausforderungen und ethischen Fragen rund um algorithmische Entscheidungen ist. Algorithmische Entscheidungsprozesse transparent zu machen ist auch kein Freifahrtschein für den uneingeschränkten Einsatz von ADM-Systeme in unserer Gesellschaft. Wenn Algorithmen aber in Bereichen verwendet werden, in denen sie das gesellschaftliche Zusammenleben beeinflussen, ist Transparenz wichtig für einen gemeinwohlorientierten Einsatz. Um im digitalen Zeitalter mündig am gesellschaftlichen Zusammenleben teilhaben zu können, wird es zukünftig immer stärker darauf ankommen zu wissen, wo Algorithmen ins tägliche Leben eingreifen und es beeinflussen. Ansonsten besteht die Gefahr einer neuen „digitalen Spaltung“. Dann geht es nicht darum, wer Zugang zur digitalen Welt hat und wer nicht, sondern wer den Einfluss von Algorithmen versteht und wer sich blind auf sie verlässt und sich ihnen ein Stück weit ausliefert.⁵¹

⁵¹ Susarla, Anjana (2019): The new digital divide is between people who opt out of algorithms and people who don't. In: Quartz. 17.04.2019. <https://qz.com/1597901/> (Stand: 15.11.2019)

ÜBER DIE AUTORIN

Leonie Beining ist Projektleiterin im Projekt "Algorithmen fürs Gemeinwohl", das unterschiedliche gesellschaftliche Stakeholder zusammenbringt, um in einem intersektoralen Diskurs gesellschaftliche Chancen und Risiken algorithmischer Entscheidungssysteme zu identifizieren sowie Handlungsmöglichkeiten zu erarbeiten, die sicherstellen, dass algorithmische Entscheidungssysteme einen Beitrag zur Stärkung des Gemeinwohls leisten können. Zuvor war sie bereits als Projektmanagerin im Projekt „Gemeinwohl im digitalen Zeitalter“ tätig und entwickelte Strategien, wie sich zivilgesellschaftliche Akteure stärker in die Gestaltung der Digitalisierung einbringen können. Vor der SNV arbeitete sie in der Forschungsgruppe Globale Fragen der Stiftung Wissenschaft und Politik, wo sie sich u.a. mit dem Potential neuer Technologien für den Schutz von Kindern in bewaffneten Konflikten auseinandersetzte. Leonie studierte Politik- und Verwaltungswissenschaft an den Universitäten Konstanz und Potsdam, der Sciences Po Lille und der University of Bergen.

So erreichen Sie die Autorin:

Leonie Beining

Projektleiterin Algorithmen fürs Gemeinwohl

lbeining@stiftung-nv.de

+49 (0) 30 81 45 03 78 81

Twitter: [@LeonieBeining](https://twitter.com/LeonieBeining)

STIFTUNG NEUE VERANTWORTUNG

Neue Technologien verändern Gesellschaft. Dafür brauchen wir rechtzeitig politische Antworten. Die Stiftung Neue Verantwortung ist eine unabhängige Denkfabrik, in der konkrete Ideen für die aktuellen Herausforderungen des technologischen Wandels entstehen. Um Politik mit Vorschlägen zu unterstützen, führen unsere Expert:innen Wissen aus Wirtschaft, Wissenschaft, Verwaltung und Zivilgesellschaft zusammen und prüfen Ideen radikal.

BERTELSMANN STIFTUNG

Die Bertelsmann Stiftung setzt sich für eine gerechte Teilhabe aller am gesellschaftlichen Leben ein. Sie engagiert sich in den Bereichen Bildung, Demokratie, Gesellschaft, Gesundheit, Kultur und Wirtschaft. Durch ihr Engagement will sie alle Bürgerinnen und Bürger ermutigen, sich für das Gemeinwohl einzusetzen. Die 1977 von Reinhard Mohn gegründete, gemeinnützige Einrichtung hält die Mehrheit der Kapitalanteile der Bertelsmann SE & Co. KGaA. Die Bertelsmann Stiftung arbeitet operativ und ist unabhängig vom Unternehmen sowie parteipolitisch neutral.

ÜBER DAS PROJEKT

Das Kooperationsprojekt „Algorithmen fürs Gemeinwohl“ von Stiftung Neue Verantwortung und Bertelsmann Stiftung verfolgt eine zu den verbreiteten kritischen Ansätzen komplementäre Strategie, indem es das gesellschaftliche Potential von algorithmischen Systemen in einem Bottomup-Ansatz ins Zentrum der Projektarbeit stellt. Anhand einer Analyse von Anwendungsbereichen, in denen das Potential algorithmischer Entscheidungsfindung besonders groß ist, soll geklärt werden, wie algorithmische Systeme per Design gemeinwohlorientiert gestaltet werden können. Ziel ist es, Prinzipien zu formulieren, über die bereits in der konzeptionellen und technischen Entwicklungsphase sichergestellt wird, dass Algorithmen ihr gesellschaftliches Verbesserungspotenzial voll entfalten können, ohne negative Effekte, wie die systematische Benachteiligung einzelner Personen oder Gruppen, über Skaleneffekte auf weite gesellschaftliche Teile auszudehnen.

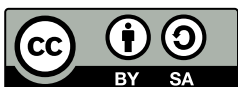
IMPRESSUM

Stiftung Neue Verantwortung e. V.
Center
Berliner Freiheit 2
10785 Berlin
www.stiftung-nv.de

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
www.bertelsmann-stiftung.de

Design: Jan Klöthe
www.jankloethe.de

Kostenloser Download: www.stiftung-nv.de



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

