

Alexander Brödner

# Automatisierte (Un)Gleichheit durch Künstliche Intelligenz

## Inwiefern befördert der Einsatz von Algorithmen bei Entscheidungsprozessen Diskriminierung?

### 1. DIDAKTISCHE ZIELSETZUNG/DIDAKTISCHE ERLÄUTERUNGEN

Algorithmen beeinflussen unser Leben mittlerweile in vielfältiger Art und Weise. Oftmals unbemerkt sind sie in sogenannten Algorithmischen Entscheidungssystemen (englisch: Automated decision-making, kurz: ADM) am Werk. Solche ADM-Systeme übernehmen in Unternehmen beispielsweise eine Vorauswahl von Bewerberinnen und Bewerbern bei der Besetzung offener Stellen, unterstützen in Banken die Einschätzung der Kreditwürdigkeit von Antragstellenden oder steuern die Anzeige individualisierter Preise im Online-Handel. Auch in der öffentlichen Verwaltung gewinnen ADM-Systeme an Bedeutung. Es ist dabei ein Irrglaube, dass Entscheidungen durch Algorithmen immer neutral und objektiv sind. Im Gegenteil: Sie können zur Folge haben, dass Menschen in ihrer gesellschaftlichen Teilhabe beeinträchtigt, ihre Grundrechte beschnitten werden oder ihnen der Zugang zu öffentlichen Gütern und Dienstleistungen verwehrt wird.

Ziel der Unterrichtseinheit (UE) ist es, bei den Schülerinnen und Schülern (SuS) ein kritisches Bewusstsein für die Möglichkeit von Diskriminierung durch den Einsatz von ADM-Systemen zu fördern. Dafür findet zunächst eine Hinführung zum Begriff der Diskriminierung anhand von Plakaten einer Kampagne der Antidis-

kriminierungsstelle des Bundes statt (M 1a). Dem folgt eine Einführung in das Allgemeine Gleichbehandlungsgesetz (AGG) anhand eines Erklärfilms (M 1b). Das AGG – umgangssprachlich auch als Antidiskriminierungsgesetz bekannt – ist seit 2006 deutsches Bundesgesetz. Es soll „Benachteiligungen aus Gründen der Rasse oder wegen der ethnischen Herkunft, des Geschlechts, der Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Identität“ verhindern und beseitigen. Als Differenzierungsangebot für schnelle Lernende bietet sich hier eine verknüpfende Reflexion über die Wirkung der Plakate und die konzeptionelle Entwicklung neuer Plakate an. Das Ergebnis von M 1 ist eine vorläufige Formulierung des Problems der Diskriminierung, die durch M 2 erweitert werden soll. Hierfür wird anhand eines Artikels aus einem philosophischen Handbuch die Unterscheidung von direkter und indirekter Diskriminierung erarbeitet. Mit M 3 erleben die SuS verschiedene Formen von Entscheidungsprozessen anhand des Beispiels der Kreditvergabe hautnah. Sie diskutieren und reflektieren die Möglichkeiten von Diskriminierung in solchen Entscheidungsprozessen. M 4 verlangt schließlich eine Beurteilung von Diskriminierung durch algorithmengestützte automatisierte Entscheidungsprozesse anhand eines realen Fallbeispiels. Als Differenzierungsangebot für schnelle Lernende

bietet sich hier eine intensive Beschäftigung mit dem AGG an. M 5 vermittelt grundlegende Kenntnisse der Ursachen von Diskriminierung durch ADM-Systeme und fordert die SuS zur kritischen Auseinandersetzung auf. Anschließend soll die Möglichkeit von diskriminierungsfreien Entscheidungsprozessen durch Algorithmen diskutiert werden. Mit M 6 steht am Abschluss der UE die Erarbeitung eines Empfehlungsschreibens zur Verbesserung des AGG unter Berücksichtigung der gewonnenen Kenntnisse. Dies lässt sich auch als Lernerfolgskontrolle mit ggf. angepassten Anforderungen nutzen. ■

### UNSER AUTOR

Foto: Privat



Dr. phil. Alexander Brödner unterrichtet Mathematik, Physik und Philosophie/Ethik an einem Berliner Gymnasium.

### 2. LÖSUNGSVORSCHLÄGE

#### Zu M 1a und M 1b

*Allgemeine Formulierung des Problems der Diskriminierung: Ungerechtfertigte Benachteiligung von Menschen aufgrund von bestimmten Eigenschaften dieser Menschen*

*Sechs Gründe von Benachteiligung (Diskriminierungsmerkmale), vor denen das AGG schützt: Ethnische Zugehörigkeit und Herkunft (Rassismus), Geschlecht und Geschlechtsidentität, Religion/Weltanschauung, Beeinträchtigung/Erkrankung, Alter, sexuelle Identität*  
*Anwendungsbereiche, in denen das AGG gilt: Arbeitsleben und*

Alltagsgeschäfte (Geschäfte des täglichen Lebens wie Einkäufe, Restaurant-, Diskotheken- und Friseurbesuche, Bahn- und Busfahrten)

### Zu M 2

*Direkte Diskriminierung:* Dabei besteht eine direkte Verbindung zwischen der individuellen Einstellung von Personen und deren Diskriminierungsverhalten. Eine Person urteilt beispielsweise explizit herabwürdigend über Angehörige einer bestimmten Gruppe und schließt diese Menschen vom Zugang zu ihrem Unternehmen aus.

*Indirekte Diskriminierung:* Dabei basiert Diskriminierung nicht auf individuellen Vorurteilen. Diskriminierung geschieht hier aufgrund von bestimmten strukturellen Zusammenhängen oder Mustern in der Gesellschaft als ganzer. Beispielsweise werden Angehörige einer bestimmten Gruppe von ADM-Systemen als nicht kreditwürdig eingestuft.

*Erweiterte Formulierung des Problems der Diskriminierung:* Ungerechtfertigte Benachteiligung von Menschen aufgrund von bestimmten Eigenschaften dieser Menschen. Dabei kann diese Benachteiligung aufgrund von absichtlichen Handlungen einzelner Personen entstehen oder auch durch strukturelle Zusammenhänge in der Gesellschaft als ganzer.

### Zu M 3

*Kriterien, die für die Entscheidung über Kreditwürdigkeit tatsächlich verwendet werden (für die 5 Fragen):*

Alter und Geschlecht, Familienstand und Kinder, Wohnsitz, Verträge (Telefon, Internet etc.), Vorstrafen, Einkommen und Ausgaben, Höhe der Kreditraten, berufliche Qualifikation, Besitz von Kreditkarten, Häufigkeit des Wechsels des Girokontos, Anzahl offener oder laufender Zahlungsverpflichtungen, Arbeitsvertrag, Sicherheiten (Bürgschaften, Immobilien etc.)

### Zu M 4

Die Frau aus Hessen wird benachteiligt, vermutlich aufgrund von Geschlecht und Alter. Dabei handelt es sich nicht um eine absichtliche Handlung einer einzelnen Person, sondern um indirekte Diskriminierung aufgrund bestimmter gesellschaftlicher Strukturen. Hier vermutlich aufgrund der Daten, die der Entscheidung des ADM-Systems zugrunde liegen und bestehende Diskriminierungsmuster in der Gesellschaft abbilden.

### Zu M 5

Kritische Fragen zu den einzelnen Bereichen:

*Daten:* Sind die Daten repräsentativ für die Gruppen, auf die das ADM-System einen Effekt haben wird? Wer wählt die Daten aus und warum wählt die Person diese Daten (z. B. Qualität der Daten, Verfügbarkeit oder Kosten)? Sind die Daten aktuell und ausreichend? Wer prüft das? Welche Fragen können mit den Daten beantwortet werden? Wer hat die Daten annotiert und nach welchen Kriterien?

*Ziel:* Welcher Zweck soll durch den Einsatz eines ADM-Systems erfüllt werden? Welche Annahmen über ADM-Systeme stecken hinter der Einschätzung, dass dies die beste Lösung für das identifizierte Problem ist? Welche (kritischen) Fragen werden gestellt, welche Alternativen in Betracht gezogen? Wer definiert das Ziel des Einsatzes, das Problem und die Lösung? Welche Annahmen hat diese Person? Welche Perspektiven auf die Gesellschaft hat die Entscheidungsperson?

*Modell:* Auf welchen Annahmen über das Problem, die Realität oder das Ziel basiert das Modell? Wer übersetzt das Problem in ein Modell und wie? Welchen Ausschnitt der Realität bildet das Modell ab? Ist die Vereinfachung dem Problem angemessen? Ist das Modell geprüft worden und nach welchen Qualitätsstandards? Fand die

Prüfung unternehmensintern oder durch eine unabhängige Stelle statt?

*Anwendung:* Wer wendet das ADM-System an und wofür? Sind Anwendende hinreichend geschult? Kennen sie die Möglichkeiten und Grenzen? Welche Kompetenzen und welches Wissen haben die Anwendenden? Haben Anwendende die Möglichkeit oder Freiheit, sich anders zu entscheiden als vom System empfohlen? Ist dann eine abweichende Entscheidung im Arbeitsalltag realistisch, z. B. angesichts knapper Zeit oder Personalressourcen? Welche Konsequenzen ergeben sich bei einem Widerspruch für Anwendende?

Werden diese kritischen Fragen in der Praxis bedacht, besteht die Möglichkeit, dass Algorithmen diskriminierungsfrei arbeiten.

### Zu M 6

Mögliche Verbesserungsvorschläge für das AGG:

1. Die Aufklärungsarbeit und Sensibilisierung bezüglich algorithmensbasierter Diskriminierung fördern.
2. Prüfen, ob Anwendungsbereiche und Diskriminierungsmerkmale im AGG mit Blick auf algorithmensbasierte Diskriminierung ausgeweitet werden sollten (ggf. insbesondere mit Rücksicht auf Proxy-Variablen, die stellvertretend für geschützte Kategorien nach AGG für automatisierte Entscheidungen herangezogen werden könnten, aber selbst nicht unter das AGG fallen).
3. Ein Verbandsklagerecht im AGG einführen.
4. Für mehr Transparenz und klare Verantwortlichkeiten bei Verdachtsfällen von Diskriminierung durch ADM-Systeme sorgen (zum Beispiel, indem Auskunftspflichten gegenüber Antidiskriminierungsstellen und Betroffenen gestärkt und Einspruchsmöglichkeiten für Betroffene eingeführt werden).

### 3. Unterrichtspraktischer Steckbrief

Philosophische Teildisziplin	Sozialphilosophie, Politische Philosophie, Ethik, Angewandte Ethik
Thema der Sequenz	Diskriminierung, insbesondere durch algorithmische Entscheidungssysteme
Zielgruppe	ab Jahrgangsstufe 10/11
Schlüsselbegriffe	Diskriminierung, Künstliche Intelligenz, Algorithmen, Automatisierte Entscheidungsprozesse, Allgemeines Gleichbehandlungsgesetz
Kompetenzen	Die SuS können <ul style="list-style-type: none"> <li>■ den Gehalt und die Wirkung visueller Medien deuten und reflektieren;</li> <li>■ zentrale Begriffe eines philosophischen Textes anhand von Beispielen erläutern;</li> <li>■ Kernaussagen aus philosophischen Texten schematisch darstellen;</li> <li>■ verschiedene Begriffe von Diskriminierung in komplexen Zusammenhängen anwenden;</li> <li>■ verschiedene Formen von Diskriminierung anhand von Fallbeispielen erkennen;</li> <li>■ ein begründetes Urteil zu philosophischen Fragen unter Berücksichtigung erworbener Kenntnisse vertreten;</li> <li>■ die Perspektive anderer Personen einnehmen und reflektieren;</li> <li>■ sich mit eigenen Beiträgen problembewusst und lösungsorientiert in eine Diskussion einbringen;</li> <li>■ im Dialog die eigene Position reflektieren und sie ggf. revidieren.</li> </ul>
Zeitaufwand	ca. 6–8 Unterrichtsstunden à 45 Minuten
Lernerfolgskontrolle	Empfehlungsschreiben zur Erweiterung des Allgemeinen Gleichbehandlungsgesetzes ( <b>M 6</b> )
Weiterführung / Anschlussmöglichkeiten	Thematisierung von Diskriminierung in anderen Bereichen unabhängig von automatisierten Entscheidungsprozessen; Thematisierung eines Begriffs von Gerechtigkeit (z. B. ausgehend von Rawls)



Arbeitsblätter bequem online anpassen!

Mehr dazu auf Seite 3.

MATERIALIEN/MEDIEN	inhaltlich/didaktisch/methodisch
<b>M   1a</b> Diskriminierung ist verboten!	Hinführung zum Thema Diskriminierung durch Deuten von Plakaten der Antidiskriminierungsstelle des Bundes
<b>M   1b</b> Das Allgemeine Gleichbehandlungsgesetz (AGG)	Einführung in das AGG anhand eines Erklärfilms (Dauer: 3:14 Minuten) ggf. Reflexion der Wirkung der Plakate und konzeptionelle Entwicklung neuer Plakate
<b>M   2</b> Diskriminierung aus philosophischer Perspektive	Erweiterung der Perspektive auf das Problem der Diskriminierung durch Erarbeitung der Unterscheidung von direkter und indirekter Diskriminierung
<b>M   3</b> Diskriminierungsfreie Entscheidungsverfahren: Beispiel Studienkredit	Diskussion verschiedener Formen von Entscheidungsprozessen anhand des Beispiels der Kreditvergabe Reflexion auf die Möglichkeit von Diskriminierung in Entscheidungsprozessen
<b>M   4</b> Diskriminierung durch automatisierte Entscheidung?	Beurteilung von Diskriminierung durch ADM-Systeme anhand eines Fallbeispiels ggf. Diskussion der Notwendigkeit einer Erweiterung des AGG
<b>M   5</b> Ursachen für Diskriminierung bei automatisierter Entscheidung	kritische Auseinandersetzung mit den Ursachen von Diskriminierung durch ADM-Systeme und Diskussion der Möglichkeit von diskriminierungsfreien Entscheidungsprozessen durch Algorithmen
<b>M   6</b> Erweiterung des Allgemeinen Gleichbehandlungsgesetzes (AGG)	Erarbeitung eines Empfehlungsschreibens zur Verbesserung des AGG unter Berücksichtigung der gewonnenen Kenntnisse

**M|1a** Diskriminierung ist verboten!

Die abgebildeten Plakate gehören zu einer Kampagne der Antidiskriminierungsstelle des Bundes, die über bestimmte Inhalte des sogenannten Allgemeinen Gleichbehandlungsgesetzes (AGG) aufklären soll. Das Gesetz ist auch unter dem Begriff „Antidiskriminierungsgesetz“ bekannt. Es trat in Deutschland im Jahr 2006 in Kraft.



**AUFGABEN**

- 1 Wählt in Gruppenarbeit eines der beiden Plakate aus. Beschreibt euch gegenseitig zunächst rein deskriptiv, was auf dem Plakat zu sehen ist.
- 2 Erläutert euch gegenseitig das Problem, auf das euer Plakat – aus Sicht der Antidiskriminierungsstelle des Bundes – hinweisen will.
- 3 Entwickelt gemeinsam eine allgemeine Formulierung des Problems, das in den Plakaten beispielhaft veranschaulicht wird. Notiert diese in möglichst prägnanter und kurzer Form.

**M|1b** Das Allgemeine Gleichbehandlungsgesetz (AGG)

Auf ihrer Website hat die Antidiskriminierungsstelle des Bundes einen Erklärfilm zum AGG veröffentlicht:

**Erklärfilm zum AGG (3:14 Minuten)**

[https://www.antidiskriminierungsstelle.de/SharedDocs/videos/DE/Gebaerdensprache/agg\\_erklaerfilm.html](https://www.antidiskriminierungsstelle.de/SharedDocs/videos/DE/Gebaerdensprache/agg_erklaerfilm.html)



**AUFGABEN**

- 1 Schaut euch den Erklärfilm an. Ergänzt eure allgemeine Formulierung des Problems der Diskriminierung aus **M 1a** um:
  - a) die sechs Gründe von Benachteiligung (Diskriminierungsmerkmale), vor denen das AGG schützt;
  - b) die beiden Anwendungsbereiche, in denen das AGG gilt.
- 2 **Zusatzauftrag:**
  - a) Diskutiert, inwieweit die beiden Plakate den Zweck erfüllen, über Inhalte des AGG aufzuklären. Notiert Gründe dafür und dagegen.
  - b) Entwickelt gemeinsam die konzeptionelle Idee für ein drittes Plakat zum AGG.



**M|2** Diskriminierung aus philosophischer Perspektive

In philosophischen, rechtlichen, sozialwissenschaftlichen, politischen und alltagssprachlichen Zusammenhängen ist der Begriff der Diskriminierung im Deutschen negativ konnotiert. Er dient hier zur Bezeichnung von Unterscheidungen, die als ungerechtfertigt betrachtet werden, weil sie moralisch bedeutungslose bzw. im jeweiligen Kontext irrelevante Unterschiede zwischen Menschen zum Anlass für die ungleiche Verteilung von Lasten und Vorteilen oder für die herablassende Behandlung von Personen nehmen. Klassische Beispiele von Diskriminierung sind die Benachteiligung oder Missachtung von Personen aufgrund ihrer Hautfarbe, ihrer Muttersprache und ethnischen Zugehörigkeit, ihrer Herkunft, ihres Geschlechts oder ihrer Religionszugehörigkeit. Insofern Merkmale wie die genannten hier als Grund für die (vergleichsweise) schlechte Behandlung von Personen fungieren, sind diskriminierende Handlungen mehr oder weniger direkt mit einer bewussten oder unbewussten Abwertung der jeweiligen Eigenschaften verbunden. Die entsprechenden (Un)Werturteile kommen in Form von stereotypen Vorurteilen über die Gruppe der Merkmalsträger zum Ausdruck, die sowohl zur Motivation der Benachteiligung als auch zu ihrer nachträglichen Rechtfertigung dienen können. Da sie häufig den Verweis auf das diskriminierende Differenzierungskriterium ersetzen, ist Diskriminierung oft nur schwer nachweisbar. [...]

Während die Forschung zur Diskriminierung sich lange Zeit vornehmlich auf das diskriminierende Individuum konzentrierte

und von einer direkten Verbindung von vorurteilsbehafteter individueller Einstellung und subjektivem Diskriminierungsverhalten ausging (direkte/intentionale Diskriminierung), konzentrierte man sich ab Mitte der 1960er Jahre weit stärker auf sozio-strukturelle und historische Bedingungsfaktoren von Diskriminierung. Entsprechende Erklärungsansätze [...] betrachten Diskriminierung nicht als ein auf individuellen Vorurteilen basierendes Fehl- und Abwehrverhalten. Sie verstehen sie als objektives, institutionelles Phänomen, das in sozialen Hierarchien manifestiert ist und weitgehend unabhängig von den Absichten und Einstellungen individueller Akteure indirekt und im Verborgenen über die Verteilungsinstitutionen der Gesellschaft wirkt (indirekte/strukturelle/mittelbare Diskriminierung). [...] Diskriminierung ist demnach nicht notwendigerweise mit individuellen Schädigungsabsichten, intendierten Missachtungen von Personen oder entsprechenden kollektiven Unwerturteilen verbunden. Sie wirkt vielmehr absichtslos durch die Institutionen der Gesellschaft. Weil diese von vergangenen Diskriminierungen geprägt [...] und daher den Bedürfnissen und Interessen der unterschiedlichen gesellschaftlichen Gruppen nicht gleichermaßen gut angepasst sind, wirken sie zum Nachteil der Angehörigen sozialer Minderheiten.

Quelle: Susanne Boshammer: Diskriminierung. In: Handbuch der Politischen Philosophie und Sozialphilosophie. Hrsg. v. Stefan Gosepath, Wilfried Hinsch u. Beate Rössler, in Zusammenarb. m. Robin Celikates und Wulf Kellerwessel. Bd. 1: A–M. Berlin: De Gruyter 2008. S. 232-238, hier S. 232ff.

**AUFGABEN**

Findet neue Partnerinnen oder Partner, mit denen ihr eure bisher entwickelte Formulierung zum Problem der Diskriminierung aus philosophischer Perspektive ergänzt. Geht dazu wie folgt vor:

- 1 Lest den obigen Handbucheintrag zur Diskriminierung.
  - a) Erstellt ein Begriffsnetz, das den Zusammenhang zwischen den drei Begriffen Diskriminierung, direkte Diskriminierung und indirekte Diskriminierung darstellt.
  - b) Erläutert direkte und indirekte Diskriminierung anhand eines Beispiels.
- 2 Entwickelt mithilfe der Unterscheidung zwischen direkter und indirekter Diskriminierung eine erweiterte Formulierung des Problems der Diskriminierung.
- 3 Kehrt zurück zu eurem ursprünglichen Partner oder eurer ursprünglichen Partnerin und vergleicht eure Vorschläge für eine erweiterte Formulierung des Problems der Diskriminierung. Verfasst auf Grundlage von beiden Vorschlägen eine gemeinsame Fassung. (Einigt euch nun auch im Plenum auf eine gemeinsame Fassung.)

**M|3** Diskriminierungsfreie Entscheidungsverfahren: Beispiel Studienkredit

Stellt euch folgende Situation vor: Ihr habt das Abitur in der Tasche und wurdet für ein Studium an der renommierten Universität in Oxford (England) angenommen. Es gibt nur ein Problem: Für das Studium an der Oxford-University müsst ihr circa 10.000 Euro Studiengebühren pro Jahr bezahlen. Um das Studium finanzieren zu können, braucht ihr einen Kredit von einer Bank. Doch die Bank gibt nicht jeder Person einen Kredit. Nur an Personen, deren Kreditwürdigkeit erwiesen ist, verleiht sie Geld.

Quelle: Autorentext

**Info-Box: Kreditwürdigkeit und die Schufa**

Einen Kredit erhält nur, wer kreditwürdig ist. Vor Abschluss eines Kreditvertrages müssen Antragsstellende gegenüber dem Kreditinstitut deshalb einen Nachweis ihrer Bonität erbringen, also glaubhaft belegen, dass sie ihren Zahlungsverpflichtungen zu den vereinbarten Fristen und in der vereinbarten Höhe nachkommen können. Um dies zu überprüfen, haben Banken verschiedene Möglichkeiten. Am häufigsten holen sie eine sogenannte Schufa-Auskunft ein. Die Schufa ist ein privates Unternehmen, das Daten von Verbraucherinnen und Verbrauchern sammelt und auf dieser Grundlage für jede Person einen Score-Wert erstellt. Dieser Wert zeigt an, für wie kreditwürdig die Schufa diese Person hält. Ein niedriger Schufa-Score bedeutet eine geringe Kreditwürdigkeit. Wer von der Schufa eine solche Einstufung erhält, hat deutlich schlechtere Chancen auf einen Kredit.

Quelle: Autorentext

**AUFGABEN**

- 1 Teilt die gesamte Lerngruppe in zwei gleich große Gruppen auf.

**Gruppe A**

a) Arbeitet zunächst in 3er-Gruppen. Stellt euch vor, ihr seid Mitarbeiterinnen oder Mitarbeiter der Bank, bei der die im Einführungstext beschriebene Person einen Kredit beantragt. Eure Aufgabe ist es, in einem persönlichen Interview herauszufinden, ob die Person kreditwürdig ist. Entwickelt dafür 5 Fragen, die verschiedene Kriterien überprüfen und deren Beantwortung durch die Person Aufschluss darüber gibt, ob sie den Kredit bekommen sollte oder nicht. Die Fragen sollten jeweils in wenigen mündlichen Sätzen zu beantworten sein.

b) Tauscht euch mit den anderen Gruppen aus und einigt euch auf die 5 Fragen, die für ein diskriminierungsfreies Entscheidungsverfahren am besten geeignet sind.

**Gruppe B**

a) Arbeitet zunächst in 3er-Gruppen. Stellt euch vor, ihr seid Mitarbeiterinnen oder Mitarbeiter der Bank, bei der die im Einführungstext beschriebene Person einen Kredit beantragt. Eure Aufgabe ist es, durch einen computergestützten Fragebogen herauszufinden, ob die Person kreditwürdig ist. Entwickelt dafür 5 Fragen, die verschiedene Kriterien überprüfen und deren Beantwortung durch die Person Aufschluss darüber gibt, ob sie den Kredit bekommen sollte oder nicht. Da der Fragebogen automatisch von einem Computer ausgewertet wird, müssen die Fragen so gestaltet sein, dass darauf entweder mit Ja oder Nein, mit einer Zahl oder mit einem einzelnen Wort geantwortet werden kann.

b) Tauscht euch mit den anderen Gruppen aus und einigt euch auf die 5 Fragen, die für ein diskriminierungsfreies Entscheidungsverfahren am besten geeignet sind.

- 2 Findet euch in 2er-Gruppen mit jeweils einer Person aus Gruppe A und einer Person aus Gruppe B zusammen. Eine Person von euch stellt sich vor, dass sie einen Kredit beantragt. Die andere Person ist Mitarbeiterin oder Mitarbeiter der Bank. Führt jeweils das Interview durch bzw. lasst den Fragebogen ausfüllen. Danach tauscht ihr die Rollen. Reflektiert anschließend eure Erfahrungen.
- 3 Diskutiert im Plenum, welches der beiden Entscheidungsverfahren das bessere ist.

**M|4** Diskriminierung durch automatisierte Entscheidung?

Automatisierte Entscheidungen [durch Algorithmen] durchdringen zunehmend unseren Alltag: zum Beispiel, wenn unsere Kreditwürdigkeit beim Online-Einkauf beurteilt wird, wenn uns über sogenannte „targeted ads“ (personalisierte Werbung) Stellenanzeigen auf Basis von Online-Persönlichkeitsprofilen angezeigt werden oder wenn Sozialleistungen berechnet werden. Diese Systeme des automatisierten Entscheidens (ADM-Systeme) können zu Diskriminierungen führen. [...]

Gerade in Pandemiezeiten boomt das Onlineshopping. Insbesondere diejenigen, die auf Rechnung zahlen möchten, werden oft vorab und meist unbemerkt in Echtzeit einer Online-Bonitätsprüfung unterzogen. Für diese automatisierten Credit-Scoring-Verfahren beauftragen Unternehmen in der Regel private Auskunftsteien wie die Schufa [...].

Bisweilen kommt es zu problematischen Bewertungen. 2021 etwa wandte sich eine Frau in Hessen an eine Antidiskriminierungsberatungsstelle, nachdem bei einer Online-Bestellung ihr Kauf auf Rechnung abgelehnt worden war. [...] Die betroffene Frau [war] überrascht, da ihr zufolge für sie keine negativen Schufa-Auskünfte vorlagen und sie über ein gutes Einkommen verfügt. Sie berichtete bei der Beratungsstelle, dass sie bei dem Kreditprüfungsunternehmen angerufen habe, um nach dem Grund für die Ablehnung zu fragen. Dort habe ein Mitarbeiter am Telefon die Vermutung geäußert, dass das System wahrscheinlich auf Basis einer Kombination aus Alter und Geschlecht diese Entscheidung getroffen habe. Die Entscheidung selbst konnte die Frau nicht anfechten. Durch diesen Vorgang fühlte sie sich diskriminiert.

Ein weitverbreiteter Trugschluss lautet, dass automatisierte Entscheidungen objektiv sind und daher nicht diskriminierend wirken können. Allerdings sind Algorithmen und sogenannte Künstliche Intelligenz niemals neutral oder objektiv. Zum einen entscheiden Menschen, welche Schritte durch einen Algorithmus ausgeführt werden und nach welchen Kriterien automatisierte Entscheidungen getroffen werden. Zum anderen beruhen automatisierte Entscheidungen oft auf Datensätzen. Diese Datensätze bilden mitunter bestehende Diskriminierungsmuster in der Gesellschaft ab. Sowohl die menschlichen Entscheidungen darüber, nach welchen Kriterien ein ADM-System Entscheidungen trifft, als auch Datensätze mit einem Bias<sup>1</sup> können zu diskriminierenden Entscheidungen durch ADM-Systeme führen.

1 Der Begriff Bias (engl. für Vorurteil, Neigung, Einseitigkeit, Verzerrung, Voreingenommenheit u. a.) verweist in Bezug auf Algorithmen und Datensätze auf fehlerhafte und unfaire Entscheidungen, die zu einer systematischen Ungleichbehandlung führen können.

Quelle: Sarah Michot, Anne Mollen, Anna Lena Schiller, Jessica Wulf: Digital Autonomy Hub: Algorithmenbasierte Diskriminierung – Warum Antidiskriminierungsgesetze jetzt angepasst werden müssen. AW AlgorithmWatch gGmbH und Gesellschaft für Informatik e.V. (Hrsg.): POLICY BRIEF #5, Februar 2022. S. 3f. [https://algorithmwatch.org/de/wp-content/uploads/2022/02/DAH\\_Policy\\_Brief\\_5.pdf](https://algorithmwatch.org/de/wp-content/uploads/2022/02/DAH_Policy_Brief_5.pdf) [03.08.2023] (verändert)



**AUFGABEN**

- 1 Findet euch in 2er-Gruppen zusammen. Eine Person übernimmt die Rolle der im Artikel genannten Frau aus Hessen, die andere Person interviewt die Betroffene als Journalistin oder Journalist zu ihrer im Text beschriebenen Erfahrung. Die Interviewte hat die Aufgabe, im Sinne des obigen Textes zu antworten. Bereitet euch selbstständig vor und führt das Interview durch.
- 2 Benennt die Instanz, die für die Entscheidung im beschriebenen Fall verantwortlich war.
- 3 Beurteilt auf Grundlage der in **M 2** erarbeiteten erweiterten Formulierung des Problems der Diskriminierung, ob es sich im Fall der Frau aus Hessen um Diskriminierung handelt.
- 4 Zusatzauftrag: Diskutiert auf Grundlage der §§ 1–3 des AGG, inwieweit das Gesetz Fälle von automatisierten Entscheidungen durch Algorithmen fassen kann.

**M|5 Ursachen für Diskriminierung bei automatisierter Entscheidung**

Im Fall der Frau aus Hessen (M 4) kam bei der Prüfung auf Kreditwürdigkeit ein System zum automatisierten Entscheiden zum Einsatz. Diese sogenannten ADM-Systeme (englisch: Automated decision-making) führen Entscheidungen mithilfe von Algorithmen herbei. Die Entscheidung wird sozusagen von einem Computer errechnet. ADM-Systeme werden z. B. auch verwendet, um aus einem Kreis von Bewerberinnen und Bewerbern um eine Wohnung oder einen Job diejenigen herauszufiltern, die nach bestimmten Kriterien am geeignetsten erscheinen. Dass die Verwendung von ADM-Systemen zu ungerechten oder sogar zu diskriminierenden Ergebnissen führen kann, hat folgende Gründe:

**Daten**

Die Realität lässt sich über Daten nicht eins zu eins darstellen. Daten sind nicht neutral und spiegeln (Macht-)Verhältnisse der Gesellschaft [und damit auch bestimmte Diskriminierungsmuster] wider. Menschen entscheiden, welche Daten wann und wie erhoben werden. Daten unterscheiden sich in ihrer Qualität zum Beispiel darin, ob sie aktuell oder veraltet sind und welche Gruppen in ihnen über- oder unterrepräsentiert werden. Daten sind in der Vergangenheit erfasst worden und spiegeln diese wider. Werden sie benutzt, um [...] die Zukunft vorherzusagen, beruht dies auf der Annahme, dass die Zukunft der Vergangenheit gleicht oder ähnelt.

**Ziel**

Menschen entscheiden nicht nur über die verwendeten Daten, sie entscheiden auch über das Ziel der Verwendung eines ADM-Systems. Zunächst wird ein Problem identifiziert, das gelöst werden soll. Das kann zum Beispiel die Aufgabe sein, eine geeignete Person aus einem Bewerbungspool auszuwählen. Die Lösung könnte in einem von Menschen durchgeführten Recruiting-Prozess bestehen. Es kann jedoch auch entschieden werden, dass ein ADM-System für die Lösung des Problems geeignet scheint. Des Weiteren wird entschieden, welche Ziele das System verfolgen soll. Ziele können zum Beispiel Kostensenkung und Effizienzsteigerung sein. Ein Ziel kann aber auch sein, den Einfluss von Voreingenommenheit auf Entscheidungsprozesse zu reduzieren und Diskriminierung zu vermeiden.

**Modell**

[ADM-Systeme arbeiten mit Modellen.] Ein Modell beschreibt einen Ausschnitt der Wirklichkeit; die Realität in ihrer Komplexität wird reduziert. Eine Situation wird so vereinfacht, dass sie in ein mathematisches Problem übersetzt werden kann. Menschen entscheiden, welche Kriterien und Aspekte relevant für ein Problem sind, das mithilfe des Modells gelöst werden soll – beispielsweise durch Anwendung eines Algorithmus. Wenn zum Beispiel geeignete Bewerber\*innen automatisiert anhand ihrer Lebensläufe ausgewählt werden sollen und definiert wird, was „geeignet“ bedeutet, müssen die Kriterien festgelegt werden, an denen sich „geeignet“ festmachen lässt. Solche Kriterien könnten sein: Berufserfahrung, höchster Abschluss, Sprachkenntnisse usw. [Solche Kriterien können jedoch auch infrage gestellt werden.]

**Anwendung**

Die Entscheidung, ein ADM-System für eine Aufgabe einzusetzen, basiert auf Annahmen über die Möglichkeiten und Fähigkeiten dieser Systeme. Menschen tendieren dazu, ADM-Systeme zu überschätzen und sich auf Empfehlungen durch ADM-Systeme zu verlassen. Diese Tendenz wird in der englischsprachigen Fachliteratur „Automation Bias“ genannt. Sie wird umso stärker, je mehr Zeitdruck Anwendende haben oder je mehr negative Konsequenzen sie fürchten müssen, wenn sie der Empfehlung des ADM-Systems widersprechen oder nicht folgen. Deswegen ist es sehr wichtig, dass Anwendende geschult werden und lernen, wie diese Systeme funktionieren, damit sie die Empfehlungen angemessen hinterfragen können.

Quelle: Automatisierte Entscheidungssysteme und Diskriminierung Ursachen Verstehen, Fälle erkennen, Betroffene unterstützen. Ein Ratgeber für Antidiskriminierungsstellen. Mathias Spielkamp (Hrsg.). AW AlgorithmWatch gGmbH, Juni 2022. S. 12ff. [https://algorithmwatch.org/de/wp-content/uploads/2022/06/AutoCheck-Ratgeber\\_ADM\\_Diskriminierung\\_DE-AlgorithmWatch\\_Juni\\_2022.pdf](https://algorithmwatch.org/de/wp-content/uploads/2022/06/AutoCheck-Ratgeber_ADM_Diskriminierung_DE-AlgorithmWatch_Juni_2022.pdf) [03.08.2023] (verändert)

**AUFGABEN**

Arbeitet kritische Fragen heraus, die vor dem Einsatz von ADM-Systemen gestellt werden sollten:

- 1 Findet euch zu dritt in Expertengruppen für jeweils einen der im Text genannten Bereiche (Daten, Ziel, Modell, Anwendung) zusammen. Erarbeitet jeweils zwei kritische Fragen für euren Bereich.
- 2 Findet euch nun zu viert in einer neuen Gruppe so zusammen, dass ihr alle vier Bereiche abdeckt. Stellt euch gegenseitig die kritischen Fragen eurer Bereiche vor, notiert diese und diskutiert sie.
- 3 Diskutiert, inwiefern es möglich ist, dass ADM-Systeme diskriminierungsfreie Entscheidungen treffen und das besser tun, als Menschen es könnten. Notiert die Bedingungen, die dafür erfüllt sein müssten.



**M|6 Erweiterung des Allgemeinen Gleichbehandlungsgesetzes (AGG)**

Ein gesetzlicher Rahmen zum Schutz vor Diskriminierung wurde in Deutschland 2006 mit dem Allgemeinen Gleichbehandlungsgesetz eingeführt [...]. Auch bei einer Diskriminierung durch ADM-Systeme kann das Gesetz Anwendung finden. Allerdings wird Diskriminierung durch Algorithmen oft gar nicht erst erkannt, so dass die Betroffenen sich einer diskriminierenden Praktik nicht bewusst sind. Laut AGG können aber nur Betroffene selbst gegen die Diskriminierung juristisch vorgehen. Solange die Betroffenen nichts darüber wissen, bleibt die Diskriminierung also unangefochten. Teilweise erkennen Menschen eine Diskriminierung durch ADM-Systeme nicht, weil sie Algorithmen für neutral oder objektiv halten. Aber auch die Intransparenz von ADM-Systemen trägt dazu bei. Ein [...] Lösungsansatz bestünde darin, ein Verbandsklagerecht im AGG einzuführen, damit Antidiskriminierungsverbände und andere Organisationen stellvertretend für Benachteiligte in gerichtlichen Verhandlungen auftreten könnten.

Viele von Diskriminierung Betroffene scheuen den juristischen Aufwand und das damit verbundene Kostenrisiko. Die Antidiskriminierungsstellen berichten, dass sich Betroffene in vielen Fällen mit einer Entschuldigung zufriedengeben. Im Fall von algorithmenbasierter Diskriminierung ist dies allerdings fatal. Denn hier ist die Diskriminierung strukturell verankert und wird sich - wenn ein ADM-System nicht angepasst wird - zwangsläufig wiederholen.

Daher fordert Nina Bewig von der *Landesstelle für Gleichbehandlung – gegen Diskriminierung in Berlin* mehr Transparenz: „Algorithmenbasierte Diskriminierung lässt sich noch schwieriger nachweisen als analoge Diskriminierungsvorfälle. Daher ist ein gesetzlich verankertes Auskunftsrecht für Antidiskriminierungs- und Beratungsstellen - staatliche wie zivilgesellschaftliche - so enorm wichtig für die Rechtsstärkung von Betroffenen. Codes und technische Dokumentationen müssten verpflichtend offengelegt werden, um benachteiligende Outputs, die bei der Anwendung von ADM auftreten, zeitnah identifizieren zu können.“

Zudem ist das AGG auf die Anwendungsbereiche Arbeitsleben und Alltagsgeschäfte (wie Restaurantbesuche, Einkäufe oder Bankgeschäfte) begrenzt und schließt beispielsweise den Bildungs- oder Medienbereich aus. Daher wird immer wieder gefordert, die Anwendungsbereiche des AGG zu erweitern. Ebenfalls wird darüber diskutiert, ob Personen auf Grundlage weiterer Merkmale vor Diskriminierungen geschützt werden sollten. Dies ist besonders bei algorithmenbasierter Diskriminierung relevant. Denn sogenannte Proxy-Variablen können in ADM-Systemen stellvertretend für die nach AGG geschützten Merkmale als Entscheidungsgrundlage genutzt werden. So dürfte ein ADM-System, das in Bewerbungsprozessen genutzt wird, beispielsweise nicht Personen auf Grundlage ihres Alters ablehnen, da Alter nach AGG ein geschütztes Merkmal ist. Als Proxy-Variable könnte das System aber die Dauer der bisherigen Berufserfahrung nutzen, um dennoch ältere Menschen zu identifizieren und aus dem Bewerbungsprozess auszuschließen.

Quelle: Sarah Michot, Anne Mollen, Anna Lena Schiller, Jessica Wulf: Digital Autonomy Hub: Algorithmenbasierte Diskriminierung – Warum Antidiskriminierungsgesetze jetzt angepasst werden müssen. AW AlgorithmWatch gGmbH und Gesellschaft für Informatik e.V. (Hrsg.): POLICY BRIEF #5, Februar 2022. S. 4ff. [https://algorithmwatch.org/de/wp-content/uploads/2022/02/DAH\\_Policy\\_Brief\\_5.pdf](https://algorithmwatch.org/de/wp-content/uploads/2022/02/DAH_Policy_Brief_5.pdf) [03.08.2023] (verändert)

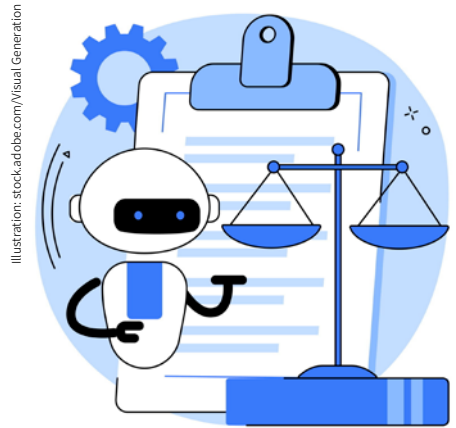


Illustration: stock.adobe.com/Visual Generation

**AUFGABE**

Im Koalitionsvertrag von 2021 hat die damals neue Bundesregierung in Aussicht gestellt, das AGG zu evaluieren, Schutzlücken zu schließen, den Rechtsschutz zu verbessern und den Anwendungsbereich anzupassen.

Lies den obigen Text und verfasse ein Empfehlungsschreiben an die Bundesregierung, das diskutiert, wie das AGG zu einem effektiven Instrument gegen algorithmenbasierte Diskriminierung weiterzuentwickeln ist. Dieses Empfehlungsschreiben soll folgendes enthalten:

- a) eine allgemein verständliche Erläuterung direkter und indirekter Diskriminierung mit jeweils einem Beispiel;
- b) eine Beurteilung der Anpassungsbedürftigkeit des AGG vor dem Hintergrund der Möglichkeit von direkter und indirekter Diskriminierung durch Algorithmische Entscheidungssysteme und deren Ursachen (mit Bezug zu zwei der vier Ursachenbereiche Daten, Ziel, Modell, Anwendung);
- c) drei konkrete und jeweils begründete Vorschläge zur Verbesserung des AGG.