

Sachdokumentation:

Signatur: DS 4654

Permalink: www.sachdokumentation.ch/bestand/ds/4654



Nutzungsbestimmungen

Dieses elektronische Dokument wird vom Schweizerischen Sozialarchiv zur Verfügung gestellt. Es kann in der angebotenen Form für den Eigengebrauch reproduziert und genutzt werden (private Verwendung, inkl. Lehre und Forschung). Für das Einhalten der urheberrechtlichen Bestimmungen ist der/die Nutzer/in verantwortlich. Jede Verwendung muss mit einem Quellennachweis versehen sein.

Zitierweise für graue Literatur

Elektronische Broschüren und Flugschriften (DS) aus den Dossiers der Sachdokumentation des Sozialarchivs werden gemäss den üblichen Zitierrichtlinien für wissenschaftliche Literatur wenn möglich einzeln zitiert. Es ist jedoch sinnvoll, die verwendeten thematischen Dossiers ebenfalls zu zitieren. Anzugeben sind demnach die Signatur des einzelnen Dokuments sowie das zugehörige Dossier.

/ Positionspapier

Schutz vor algorithmischer Diskriminierung

September 2023

Automatisierte Entscheidungen dringen zunehmend in unser tägliches Leben ein – zum Beispiel, wenn Algorithmen unsere Bewerbungen bewerten, Steuererklärungen automatisch bearbeiten, Krankheiten mittels sogenannter Künstlicher Intelligenz diagnostizieren, Verbrechen vorhersagen, das Rückfallrisiko von Straftäter*innen bewerten oder die Arbeitsmarktintegrationschancen von Geflüchteten prognostizieren sollen. Diese algorithmischen Systeme erstellen Prognosen, geben Empfehlungen oder fällen Entscheidungen (automatisierte Entscheidungsfindungssysteme, kurz: ADM) oder sie generieren Inhalte, die Entscheidungen über Menschen betreffen. Dabei können sie zu Diskriminierung führen.

In der Schweiz wird Diskriminierung im Rahmen des Diskriminierungsverbots der Bundesverfassung ([Art. 8 Abs. 2 BV](#)) als Ungleichbehandlung von Personen aufgrund eines geschützten Merkmals ohne Rechtfertigung durch einen sachlichen Grund definiert. Der aktuelle gesetzliche Rahmen in der Schweiz bietet jedoch keinen ausreichenden und wirksamen Schutz vor Diskriminierung, die durch den Einsatz von algorithmischen Systemen entstehen kann. In diesem Positionspapier stellen wir die Herausforderungen der algorithmischen Diskriminierung sowie unsere Forderungen für einen besseren Schutz vor Diskriminierung durch Algorithmen in der Schweiz vor.

Inhalt

/ Wie können Algorithmen diskriminieren?	2
/ Relevanz und Beispiele algorithmischer Diskriminierung	2
/ Ursachen algorithmischer Diskriminierung	4
/ Besondere Eigenschaften algorithmischer Diskriminierung	6
/ Warum reicht der jetzige Rechtsrahmen zum Schutz vor algorithmischer Diskriminierung nicht aus?	7
/ Wie kann der Schutz vor Diskriminierung gestärkt werden?	8

Wie können Algorithmen diskriminieren?

/ Relevanz und Beispiele algorithmischer Diskriminierung

Der **Zugang zu Sozialleistungen** kann durch Algorithmen diskriminierend gestaltet werden. Zum Beispiel stufte ein [Algorithmus](#) in den Niederlanden alleinerziehende Mütter als besonders gefährdet für Sozialhilfebetrug ein. Untersuchungen ergaben, dass bestimmte Merkmale – Elternschaft, Frau, jung, nicht fließend Niederländisch sprechend oder mit Schwierigkeiten bei der Arbeitssuche – den Betrugsrisiko-Score einer Person erhöhten. Ebenfalls in den Niederlanden wurde im Jahr 2019 bekannt, dass die Steuerbehörden einen [selbstlernenden Algorithmus](#) zur Erstellung von Risikoprofilen verwendet hatten, um Betrug beim Kinderbetreuungsgeld zu erkennen. Die Behörden bestrafte Familien wegen blossen Verdachts auf Betrug, der auf den Risikoindikatoren des Systems beruhte. Zehntausende von Familien – oft mit geringem Einkommen oder Angehörige «ethnischer Minderheiten» – wurden fälschlicherweise aufgefordert, über Jahre erhaltene Kindergelder zurückzuzahlen. Die dadurch entstehenden massiven Schulden trieben einen Grossteil von ihnen in die Armut, mehr als eintausend Kinder wurden in Pflegefamilien untergebracht. Die Datenschutzbehörde kam zum Schluss, dass die Verarbeitung der Daten durch das eingesetzte System diskriminierend war.

Im **Bildungsbereich** können etwa Algorithmen, die für Prüfungsnoten eingesetzt werden, zu einem **ungleichen Zugang zu Bildung** führen. In Grossbritannien fielen während der Covid-19-Pandemie die Prüfungen, deren Resultate den Zugang zu Universitäten bestimmen, aus. Stattdessen liess die britische Regierung die Resultate anhand eines [Algorithmus berechnen](#). Das System bediente sich an historischen Prüfungsergebnissen der Schulen, was dazu führte, dass viele überdurchschnittlich gute Schüler*innen aus schlechteren Schulen zu tiefen Noten erhielten, während mittelmässige Schüler*innen aus besseren Schulen (insbesondere Privatschulen) eher bessere Noten erhielten. Somit benachteiligte der Algorithmus Schüler*innen aus einkommensschwachen Schichten.

Im **Arbeitsbereich** kann der Einsatz von Algorithmen zum Beispiel zu einem **ungleichen Zugang zu Stellenangeboten** führen. Eine [Recherche](#) von AlgorithmWatch hat gezeigt, dass auf Facebook Stellenausschreibungen nach Gender-Stereotypen angezeigt werden. Die Anzeige für Lastwagenfahrer*innen wird eher Männern angezeigt, die Anzeige für Kinderbetreuer*innen eher den Frauen. Eine andere [Recherche](#) stellte fest, dass der Algorithmus von Facebook ältere Stellensuchende diskriminierte, da er es ermöglichte, Stellenanzeigen auf bestimmte Altersgruppen auszurichten, wodurch ältere Personen die Stellenanzeige nicht sahen und somit vom Bewerbungsprozess ausgeschlossen wurden. Algorithmen können auch im Bewerbungs- und Einstellungsverfahren diskriminierende Folgen haben. Eines der bekanntesten [Beispiele](#) dafür ist eine Bewerbungssoftware, die Amazon entwickelt haben soll, um Einstellungsprozesse effizienter zu machen. Noch in der Testphase stellte sich aber Berichten zufolge heraus, dass ihr Empfehlungsalgorithmus Frauen diskriminierte, indem ihre Lebensläufe aussortiert wurden. Versuche, die Benachteiligung von Frauen abzustellen, blieben erfolglos.

Im **Gesundheitsbereich** können algorithmische Systeme unter anderem zu einem **ungleichen Zugang zur medizinischen Versorgung** führen. In den Vereinigten Staaten verwendeten Gesundheitsdienstleister einen [Algorithmus](#), um Patient*innen zu identifizieren, die von zusätzlicher Betreuung profitieren würden. Dieser benachteiligte Schwarze Patient*innen, was dazu führte, dass weniger von ihnen die zusätzliche Betreuung erhielten, die sie benötigten.

Im Bereich **vorausschauende Polizeiarbeit (predictive policing) und Justiz** kann algorithmische Diskriminierung ebenfalls besonders schwerwiegende Folgen für Betroffene haben. Ein von Strafverfolgungsbehörden in den USA benutzter [Algorithmus](#), der mit überproportional vielen Daten von Männern mit hellen Hauttönen trainiert wurde, erkannte Gesichter mit dunkleren Hauttönen schlechter, was zu falsch positiven Treffern führte. Ein weiteres [Beispiel](#) ist das unverhältnismässige «overpolicing» bestimmter Gruppen bei der Verwendung von Algorithmen zur vorausschauenden Polizeiarbeit. Im Strafvollzug wurde in den USA der COMPAS-Algorithmus verwendet, um das Rückfallrisiko von Straftäter*innen vorherzusagen. Eine [Recherche](#) hat gezeigt, dass Schwarze doppelt so häufig fälschlicherweise als potenziell rückfällig eingestuft wurden. In der Schweiz kommt ein System mit ähnlichem Verwendungszweck zum Einsatz, das gemäss einer [Studie](#) wissenschaftlich weniger gut geprüft ist.

Auch im **Migrationsbereich** können algorithmische Systeme diskriminierende Auswirkungen haben. In den Niederlanden soll ein [Profiling-Algorithmus](#) zur Berechnung der Risikobewertung von Antragstellern, die ein Visum für einen kurzfristigen Aufenthalt in den Niederlanden und im Schengen-Raum beantragen, verwendet worden sein. Das System hätte anhand von Variablen wie Nationalität, Geschlecht und Alter ein Profil von Millionen von Visumantragsteller*innen erstellt. Antragsteller*innen, die als «hohes Risiko» eingestuft werden, werden automatisch in eine «Intensivprüfung» aufgenommen, die umfangreiche Untersuchungen und Verzögerungen nach sich ziehen kann. Ähnlich wurde auch in Grossbritannien zwischen 2015 und 2020 ein [Algorithmus](#) zur Prüfung von Visumanträgen eingesetzt. Das System sollte die Staatsangehörigkeit von Antragsteller*innen identifizieren und ergab für Personen mit «verdächtiger Staatsangehörigkeit» eine höhere Risikobewertung, die als Folge einer strengeren Kontrolle durch die Behörden unterzogen wurden.

Diese Liste von Beispielen ist nicht erschöpfend und zeigt die Vielfalt der betroffenen Bereiche. Der Einsatz von algorithmischen Systemen nimmt rasant zu. **Wo, von wem und wozu sie eingesetzt werden, ist aber sowohl den Betroffenen als auch der Gesellschaft weitgehend unbekannt** – sie können dazu nur spekulieren. Was aber naheliegt: Je mehr Algorithmen eingesetzt werden, desto mehr Diskriminierungsvorfälle können entstehen.

/ Ursachen algorithmischer Diskriminierung

Algorithmische Systeme sind weder neutral noch objektiv. In einer Gesellschaft bereits existierende strukturelle Diskriminierungsmuster werden sich in ihnen wiederfinden. Doch wie genau kann es zu algorithmischer Diskriminierung kommen?

Algorithmen verwenden eine bestimmte Datengrundlage. Die Realität lässt sich über Daten nie eins zu eins darstellen. **Daten** sind nicht neutral und spiegeln (Macht-)Verhältnisse der Gesellschaft wider. Sie 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 können immer nur diese widerspiegeln. Werden sie benutzt, um zum Beispiel Voraussagen für die Zukunft zu treffen, beruht dies auf der Annahme, dass die Zukunft der Vergangenheit ähnelt. Vor allem bei «lernenden» Algorithmen beeinflussen die verwendeten Daten den Output der Algorithmen stark. Wenn etwa in einem Unternehmen weniger Frauen in Führungspositionen vertreten sind und ein Rekrutierungsalgorithmus mit den Daten der bestehenden Belegschaft trainiert wird, kann das System diese Situation reproduzieren. Dies kann auch dann geschehen, wenn die Geschlechtsidentität nicht explizit in den Daten erfasst ist: Das System kann etwa Korrelationen zu Vornamen oder Hobbies finden, die stellvertretend für die Geschlechtsidentität «weiblich» stehen können. Gleichzeitig ist wichtig zu betonen, dass Daten von Menschen aufbereitet, bereinigt und annotiert werden, um sie für lernende Algorithmen verwendbar zu machen. Menschen entscheiden auch, welche Daten wann und wie erhoben werden. Auch hier fließen also Entscheidungen mit ein – etwa darüber, wieviel Ressourcen investiert werden (oder eben nicht), um sicherzustellen, dass Datensätze repräsentativ sind, um das Diskriminierungsrisiko zu verringern.

Ein algorithmisches **Modell** beschreibt nur einen Ausschnitt der Wirklichkeit: Die komplexe Realität 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 Algorithmus gelöst werden soll. Durch diese Entscheidungen bringen sie ihre Annahmen, Perspektiven und Voreingenommenheiten mit in die Systeme hinein – und diese werden von den Systemen in ihrem Output reproduziert, was zu Diskriminierung führen kann. Wenn etwa geeignete Bewerber*innen automatisiert anhand ihrer Lebensläufe ausgewählt werden sollen, müssen erst die Kriterien festgelegt werden, an denen sich «geeignet» festmachen lässt. Diese Kriterien können jedoch direkt oder stellvertretend für diskriminierungsrechtlich geschützte Merkmale stehen. So wurde beispielsweise vom österreichischen Arbeitsmarktservice ein [Algorithmus](#) verwendet, um die Wahrscheinlichkeit zu berechnen, dass Arbeitslose wieder in den Arbeitsmarkt eingegliedert werden. Dieser vergab Minuspunkte, wenn Arbeitslose Fürsorgepflicht hatten – aber nur dann, wenn es sich um Frauen handelte. Für Männer mit betreuungspflichtigen Kindern gab es dafür keinen Abzug – sicher deshalb, weil es in vielen Fällen so ist, dass Frauen die Betreuung übernehmen und sich somit das Kinderhaben bei Männern weniger auf die Arbeitssituation auswirkt. Wenn nun aber diese Tatsache aus der Vergangenheit in ein Modell eingebaut wird, das Entscheidungen über die Zukunft trifft, reproduziert dies nicht nur bestehende, sondern führt auch zu massiven neuen Formen der Ungerechtigkeit.

Während es wichtig ist, die Verzerrung («Bias») in den Trainingsdaten und Modellen zu thematisieren, darf dies nicht davon ablenken, dass Diskriminierung auch andere Ursachen haben kann. Selbst ein auf Diskriminierung und Voreingenommenheit geprüft System kann diskriminierend eingesetzt werden beziehungsweise diskriminierende Folgen haben. Wenn zum Beispiel eine Gesichtserkennungssoftware entwickelt wird, die alle Menschen sehr gut erkennt, jedoch vor allem benutzt wird, um «People of Color» gezielt zu kontrollieren, dann kann die Anwendung des ADM-Systems rassistische Auswirkungen haben.

Menschen entscheiden auch über den **Zweck** und das **Ziel**, für das ein algorithmisches System verwendet werden soll. Zunächst wird ein Problem identifiziert, das gelöst werden soll. Das kann zum Beispiel die Aufgabe sein, eine geeignete Person aus einem Bewerbungspool auszuwählen. So kann entschieden werden, dass, anstelle eines Menschen, ein ADM-System für den Recruiting-Prozess zum Einsatz kommt. Des Weiteren wird entschieden, welche Ziele das System verfolgen soll. Ziele können zum Beispiel die Kostensenkung und die Effizienzsteigerung sein. Wenn diese als Ziele festgelegt werden, aber etwa Gleichbehandlung oder Gerechtigkeitsaspekte nicht, kann dies zu Diskriminierung führen. Dabei hätten Algorithmen durchaus das **Potenzial**, den Einfluss von subjektiver Voreingenommenheit zu reduzieren und dadurch **Diskriminierungen zu reduzieren**. Das bedingt aber, dass dies als explizites Ziel definiert wird.

Auch die **Anwendung** eines algorithmischen Systems kann Diskriminierung verursachen. Wenn sich eine Behörde oder ein Unternehmen entscheidet, ein System für eine bestimmte Aufgabe einzusetzen, erfolgt dies immer in einem bestimmten Kontext, der mit dem System interagieren wird. Dazu gehören beispielsweise Annahmen darüber, unter welchen Umständen Nutzende den Empfehlungen oder Entscheidungen des Algorithmus widersprechen würden, könnten oder dürften. Menschen, die sich von Algorithmen beraten lassen, können im Laufe der Zeit gegenüber dem System abhängiger und unkritischer werden. Das kann zur Folge haben, dass ein Instrument, das ursprünglich zur Entscheidungs*unterstützung* gedacht war, in der Praxis zu einem Instrument der Entscheidungs*findung* wird. Ein «human in the loop» ist zwar in diesen Fällen auf dem Papier weiterhin da, gibt aber die Entscheidungsaufsicht und -hoheit *de facto* ab. Deswegen ist es sehr wichtig, dass geeignete Prozesse eingeführt werden, um diesen «[Automation Bias](#)» zu vermeiden, Anwendende geschult werden und lernen, wie diese Systeme funktionieren, damit sie die Empfehlungen und Prognosen angemessen hinterfragen können.

Letztlich drohen Ungerechtigkeiten und Diskriminierung besonders dann, wenn Algorithmen in Bereichen eingesetzt werden, in denen bereits ein **Machtgefälle** besteht – sei es zwischen Stellenbewerber*innen und Unternehmen, Arbeitnehmenden und Arbeitgeberinnen, Verdächtigen und Polizei, Flüchtenden und Grenzschutzbehörden, Sozialhilfebezüger*innen und Behörden, Schüler*innen und Lehrpersonen oder Einzelperson und Social Media-Plattform. Es droht, dass insbesondere jene Seite davon profitiert, dass ein Algorithmus entwickelt und angewendet wird, die bereits über mehr Macht und Mittel verfügt – während die andere Seite den Entscheidungen und Outputs der Systeme – und damit potenziell auch ihren ungerechten Auswirkungen – verstärkt und oft unwissentlich ausgeliefert ist.

/ Besondere Eigenschaften algorithmischer Diskriminierung

Eine zentrale Herausforderung ist, dass Diskriminierung durch ein algorithmisches System oft **schwer erkennbar** ist – sowohl für Betroffene als auch für Anwender*innen. Dies kann der Tatsache geschuldet sein, dass Menschen nicht wissen, dass ein System verwendet wird, oder dass das System intransparent ist. Die Diskriminierung kann insbesondere auch **systematisch** erfolgen, im System selbst verankert sein und damit eine grosse Anzahl von Menschen betreffen (Skalierungseffekt). Dies macht es zudem für Einzelne schwieriger, die Diskriminierung zu erkennen und sich als davon betroffen zu identifizieren.

Durch **Rückkopplungsschleifen** können sich bestehende «Biases» verstärken. Wenn etwa ein ADM-System für vorausschauende Polizeiarbeit der Polizei vorschlägt, in bestimmten Vierteln mehr zu patrouillieren, wird die Polizei dort mehr Verbrechen aufdecken als anderswo (wo so vielleicht viele Verbrechen unentdeckt bleiben). Das wiederum verstärkt die Vorhersagen des Systems.

Darüber hinaus ergibt sich algorithmische Diskriminierung oft aus **Proxy-Variablen**, Variablen die stellvertretend für andere Variablen eingesetzt werden (zum Beispiel, wenn diese nicht direkt beobachtbar oder messbar sind). Mit Proxy-Variablen kann auf bestimmte – auch geschützte – Merkmale geschlossen werden, ohne dass sie direkt abgefragt oder erfasst werden. Aus der Angabe «30 Jahre Arbeitserfahrung» lässt sich beispielsweise ablesen, dass die Person mindestens Mitte 40 sein muss. Von der Postleitzahl lassen sich oft Annahmen über den sozioökonomischen Status oder einen Migrationshintergrund ableiten. So können Proxy-Variablen einen entscheidenden Einfluss auf die Entscheidungen eines ADM-Systems haben.

Algorithmische Diskriminierung kann oft auch aus der **Kombination unterschiedlicher geschützter Merkmale** entstehen; beispielsweise, wenn Menschen gleichzeitig aufgrund ihrer Geschlechtsidentität und ihrer «Herkunft» diskriminiert werden. Es kann auch sein, dass sich der diskriminierende Effekt dadurch verstärkt und es ist dabei schwierig nachzuvollziehen, welcher Faktor darauf welchen Einfluss hat. Ausserdem können sogenannte «lernende» Systeme, die selbst Muster identifizieren, **neue Merkmale mit diskriminierenden Folgen** hervorbringen. Ein videobasiertes Gesichtserkennungssystem könnte beispielsweise Personen je nach Art ihrer Augenbewegungen oder Gestik unterschiedliche Ergebnisse (oder Bewertungen) zuweisen differenzieren und so beispielsweise bei Menschen mit Beeinträchtigungen zu unterschiedlichen Ergebnissen (oder Bewertungen) kommen. Letztlich ist es oft sehr schwierig, Informationen über die Systeme und ihre Funktionsweise zu erhalten (sogenannte «**Black Box**»-Problematik) sowie potenzielle diskriminierende Auswirkungen zu identifizieren, was dazu führt, dass **keine evidenzbasierte Debatte** stattfinden kann.

Warum reicht der jetzige Rechtsrahmen zum Schutz vor algorithmischer Diskriminierung nicht aus?

Der Gleichheitsgrundsatz («Alle Menschen sind vor dem Gesetz gleich») und das Diskriminierungsverbot sind in der Schweiz in der Bundesverfassung verankert ([Art. 8 Abs. 1 und Abs. 2 BV](#)). Die Bundesverfassung definiert die Diskriminierung als die «**Ungleichbehandlung von Personen aufgrund eines geschützten Merkmals, ohne Rechtfertigung durch einen sachlichen Grund**» ([Art. 8 Abs. 2 BV](#)). Eine Ungleichbehandlung ist gegeben, wenn eine Person oder eine Personengruppe eine im Vergleich mit einer anderen Person oder Personengruppen schlechtere Behandlung erfährt, obwohl sie sich in derselben oder in einer vergleichbaren Situation befindet. Die Bundesverfassung nennt in Art. 8 Abs. 2 BV biologische Merkmale («Rasse», Geschlecht, Alter, körperliche, geistige oder psychische Behinderung) wie auch kulturelle oder anderweitige Merkmale (Herkunft, Sprache, soziale Stellung, Lebensform, religiöse, weltanschauliche oder politische Überzeugung). Diese Aufzählung ist bewusst nicht abschliessend, da neue Gruppen, die systematischer Ausgrenzung ausgesetzt sind, erkannt werden und neue Abgrenzungsmechanismen entstehen können. Vor Diskriminierung geschützt sind generell stigmatisierte gesellschaftliche Gruppen. Zwei weitere Absätze des Artikels 8 der Bundesverfassung verlangen die Gleichstellung von Frauen und Männern in allen Lebenssituationen ([Art. 8 Abs. 3 BV](#)) und die Bekämpfung der Benachteiligung von Menschen mit Behinderungen ([Art. 8 Abs. 4 BV](#)).

Defizit 1: Das **Diskriminierungsverbot betrifft grundsätzlich nur staatliche Akteure**; ein allgemeines Diskriminierungsverbot für Private fehlt in der Schweiz. Aus dem Diskriminierungsverbot der Bundesverfassung folgt zwar eine grundrechtliche Schutzpflicht des Staates, Diskriminierung auch zwischen Privaten zu verhindern ([Art. 8 Abs. 2 i.V.m. Art. 35 Abs. 3 BV](#)). Jedoch existiert kein allgemeines Gesetz, das in genereller Weise Diskriminierung durch Private untersagt. **Algorithmen und ADM-Systeme verbreiten sich jedoch schnell in der gesamten Gesellschaft.** Es bedarf eines gesetzlichen Rahmens, der Mittel bereitstellt, um auch gegen Diskriminierung im Zusammenhang mit algorithmischen Systemen, die von privaten Akteuren verwendet werden, vorzugehen.

Defizit 2: Der bestehende Schutz gegen Diskriminierung reicht nicht aus, um den **besonderen Merkmalen der algorithmischen Diskriminierung** zu begegnen (u.a. Skalierungseffekt, Rückkopplungsschleifen, Diskriminierung durch Proxy-Variablen).

Defizit 3: Auch bei der **Durchsetzung** des Diskriminierungsverbots stellt algorithmische Diskriminierung uns vor neue Herausforderungen. Einerseits ist es **schwierig, die betroffenen Personen zu identifizieren** (die sich der Diskriminierung oft selbst nicht bewusst sind oder sein können, da sie nicht von der Anwendung des Systems wussten und Vergleichswerte fehlen oder weil die Diskriminierung systematisch erfolgt). Der derzeitige Rechtsrahmen, der auf einer klaren Identifizierung der diskriminierten Personen beruht, ist nur teilweise für diese Problematik geeignet. Andererseits verfügen die betroffenen Personen oft nicht über die notwendige Zeit und Mittel, um Rechtsmittel zu ergreifen und ihnen stehen praktische oder formale Hürden im Weg.

Wie kann der Schutz vor Diskriminierung gestärkt werden?

Damit der Schutz vor Diskriminierung den Herausforderungen algorithmischer Diskriminierung gerecht werden kann, fordern wir:

- / **Der Anwendungsbereich des Antidiskriminierungsrechts sollte ausgeweitet werden, um effektiv vor Diskriminierung durch algorithmische Systeme zu schützen.** Wenn algorithmische Systeme diskriminierende Auswirkungen haben, sind potenziell viele Menschen betroffen, denn ihre Vorhersagen oder Entscheidungen folgen alle demselben Muster. Die Risiken sind demnach nicht nur individuell, sondern auch strukturell. Die gesamtgesellschaftlichen Auswirkungen können entsprechend gross sein, wie die oben erwähnten Beispiele illustrieren.
- / **Die Diskriminierungsmerkmale sollten ausgeweitet werden, um vor Diskriminierung durch Proxy-Variablen und neuen Merkmalen zu schützen.** Denn bei der Anwendung algorithmischer Systeme können ungerechtfertigte Ungleichbehandlungen entstehen, die zu keiner erkennbaren diskriminierungsrechtlich geschützten Kategorie passen. Eine Ungleichbehandlung anhand vermeintlich neutraler Kategorien kann dennoch einen diskriminierenden Effekt hervorrufen, da die neutrale Kategorie (etwa die Postleitzahl) als Proxy stellvertretend für diskriminierungsrechtlich geschützte Kategorien (wie etwa der sozialen Stellung) stehen kann.
- / **Ein allgemeines Diskriminierungsverbot für Private sollte eingeführt werden.** In der Schweiz gibt es, von wenigen Ausnahmen abgesehen, keine Antidiskriminierungsgesetze, die sich auf private Akteure beziehen. Gerade im Bereich der algorithmischen Diskriminierung wiegt dies schwer, da Private in grosser Zahl algorithmische Systeme entwickeln und einsetzen, die relevante Auswirkungen auf das Leben der Menschen haben – wenn es etwa darum geht, Wohnungen zu vergeben, Versicherungsprämien zu berechnen, die Kreditwürdigkeit zu prüfen oder Arbeitsstellen zu vergeben. Für diesen Schutz der Diskriminierung durch Private könnte sich etwa ein allgemeines Gleichbehandlungsgesetz anbieten.
- / **Der Schutz vor Diskriminierung muss auch gewährleistet sein, wenn die Diskriminierung systematisch erfolgt und Betroffene nicht eindeutig identifizierbar sind.** Dasselbe gilt, wenn ein automatisiertes System zufällig oder aus der Systemlogik heraus eine Vorhersage oder Entscheidung mit diskriminierenden Auswirkungen auf Menschen trifft.
- / **Der Tatbestand der intersektionalen beziehungsweise auf mehreren Merkmalen beruhenden Diskriminierung sollte in das Antidiskriminierungsrecht aufgenommen werden,** damit dieses auch davor schützt, wenn ein Algorithmus sich durch eine Kombination geschützter Kategorien diskriminierend auf Menschen auswirkt.

- / **Unternehmen und Behörden, die algorithmische Systeme entwickeln, anbieten oder einsetzen, sollten dazu verpflichtet werden, regelmässig Folgenabschätzungen und standardisierte, unabhängige Audits durchzuführen**, wenn diese Systeme zum Einsatz kommen. Dies hat zum Ziel, Diskriminierungsrisiken und weitere Risiken für die Grundrechte der Betroffenen, die durch die Datengrundlage, den Einsatzzweck, das Modell oder den Anwendungskontext entstehen, zu erkennen und Massnahmen dagegen zu ergreifen. **Die Ergebnisse der Folgenabschätzungen und Audits sollten im Falle von Systemen, die von der öffentlichen Verwaltung eingesetzt werden, in öffentlich zugänglichen Verzeichnissen publiziert werden**, um öffentliche Transparenz zu schaffen und damit auch eine öffentliche Aufsicht (etwa durch Wissenschaft und Zivilgesellschaft) zu gewährleisten. Auch private Anwender sollten verpflichtet werden, besonders in sensitiven Bereichen eingesetzte Systeme sowie die Resultate von Folgenabschätzungen und Audits transparent zu machen.
- / **Es braucht bessere Möglichkeiten, um das Diskriminierungsverbot durchzusetzen.** Zuallererst müssen **praktische und formale Hürden abgebaut werden**, die Menschen davon abschrecken könnten, Rechtsmittel zu ergreifen. Dies bedeutet unter anderem, dass die **Beweislast auf Seiten der Betroffenen reduziert werden sollte**, da algorithmische Diskriminierung aufgrund intransparenter Systeme und weiterer oben erwähnter Faktoren oft nicht direkt erkennbar ist. Die Beweislast sollte in erster Linie bei der beschuldigten Partei liegen, die aufzeigen muss, dass es sich nicht um Diskriminierung handelt und dass sie alles in ihrer Macht Stehende unternommen hat, um diese zu verhindern (etwa durch Sorgfaltspflichten, Folgenabschätzungen und Audits). Weiter braucht es gegenüber öffentlichen und privaten Akteuren **kollektive Rechtsmittel**, um die für Betroffene anfallenden Hürden und Kosten weiter zu reduzieren und um der Tatsache gerecht zu werden, dass algorithmische Diskriminierung oft eine grosse Anzahl von Personen betrifft. Vor diesem Hintergrund ist auch ein **Verbandsklagerecht für zivilgesellschaftliche Organisationen** angezeigt. In Anbetracht der **«Black Box»-Problematik**, können Verbandsklagen eine zentrale Rolle spielen, um Fälle algorithmischer Diskriminierung aufzugreifen.
- / **Die Forschung und die evidenzbasierte Debatte rund um algorithmische Diskriminierung und ihre gesellschaftlichen Auswirkungen müssen gefördert werden.** Die Behörden sollten Kompetenzen aufbauen und für das Thema und die damit verbundenen Herausforderungen sensibilisiert werden. Zentral ist dabei, dass die Herausforderung algorithmischer Diskriminierung nicht als rein technisches Problem betrachtet wird, sondern dies einer interdisziplinären Herangehensweise bedarf. Schliesslich sollte der Einbezug der Betroffenen (auch bei der Entwicklung des Rechtsrahmens) gewährleistet werden.