Deutsches Institut
für Menschenrechte

Algorithmische Entscheidungssysteme

Menschenrechtliche Vorgaben und Entwicklungen auf internationaler Ebene

Information

Künstliche Intelligenz (KI) kann die Verarbeitung riesiger Datenmengen enorm erleichtern, doch birgt ihr Einsatz auch menschenrechtliche Risiken. Die vorliegende Information bietet einen Überblick darüber, wie verschiedene UN-Menschenrechtsgremien menschenrechtliche Fragen bei der Anwendung von KI beurteilen und welche Verpflichtungen sich für Staaten daraus ergeben.

Durch die rasanten technischen Entwicklungen in den vergangenen Jahren und die immer vielfältigeren Einsatzgebiete von Künstlicher Intelligenz (KI) ist gesetzgeberisches Eingreifen immer dringender geworden, um die Gefahren und Risiken beim Einsatz von KI zu regulieren. So veröffentlichte die EU-Kommission im April 2021 ihren Verordnungsentwurf für die Regulierung von KI, nachdem sie im Februar 2020 bereits ein Weißbuch für ein europäisches Konzept zum Umgang mit den Chancen und Risiken von KI veröffentlich hatte.[1] In Deutschland verabschiedete die Enquete-Kommission Künstliche Intelligenz des Deutschen Bundestages im Oktober 2020 ihren Abschlussbericht, in dem sie sich mit den Herausforderungen und Chancen von KI für Gesellschaft, Staat und Wirtschaft auseinandersetzt. Im Dezember 2020 beschloss die Bundesregierung die Fortschreibung ihrer KI-Strategie, mit der sie die Erforschung, Entwicklung und Anwendung von KI in Deutschland fördern will.[2] Im Folgenden werden die menschenrechtlich relevanten Kernpunkte der Debatte um KI erläutert.

Was sind algorithmische Entscheidungssysteme?

KI, Algorithmen oder algorithmische Systeme werden oft synonym benutzt. KI im Sinne einer eigenständig und „menschengleich“ denkenden und handelnden Maschine gibt es bisher nicht und wird es auch in absehbarer Zeit nicht geben.[3] Was es gibt, sind von Menschen entwickelte algorithmische Entscheidungssysteme,[4] die Daten analysieren und Lösungen für bestimmte Probleme oder Bereiche entwickeln und vorschlagen.[5] Dies geschieht entweder anhand eines durch die programmierende Person vorgegebenen Lösungsweges oder das System „lernt selbst“ aus den vorhandenen Daten und entwickelt daraus algorithmenbasierte Lösungswege. Je komplexer der selbstlernende Mechanismus wird, desto schwieriger oder unmöglich ist es für Menschen nachzuvollziehen, wie das System zu seinen Schlussfolgerungen gekommen ist.

Die Datenethikkommission differenzierte 2019 drei Varianten von Entscheidungen, in denen Algorithmen eine mehr oder weniger große Rolle spielen:[6] algorithmen**basierte** Entscheidungen sind „menschliche Entscheidungen, die sich auf algorithmisch berechnete (Teil-)Informationen stützen“; algorithmen**getriebene** Entscheidungen sind „menschliche Entscheidungen, die durch Ergebnisse algorithmischer Systeme in einer Weise geprägt werden, dass der tatsächliche Entscheidungsspielraum und damit die Selbstbestimmung des Menschen eingeschränkt werden“;

algorithmen**determinierte** Entscheidungen „führen automatisiert zu Konsequenzen, sodass im Einzelfall keine menschliche Entscheidung mehr vorgesehen ist“.

Menschenrechtliche Relevanz

Nicht alle Entscheidungen, die mithilfe von oder durch Algorithmen getroffen werden, haben unmittelbar menschenrechtliche Relevanz. Wenn beispielsweise Klimadaten verarbeitet werden oder ein algorithmenbasiertes Kontrollsystem an einem Fließband fehlerhaft produzierte Teile erkennt und automatisch aussortiert, ist zwar ein algorithmisches Entscheidungssystem am Werk – Menschen sind dadurch aber nicht betroffen oder nur insofern, als die eingesetzte Technik Fehler vielleicht besser erkennt als Menschen und dadurch Arbeitsplätze durch Automatisierung wegfallen oder sich verändern.

Menschenrechtlich relevant wird der Einsatz von algorithmischen Entscheidungssystemen erst dann, wenn sie Menschen in ihren Rechten betreffen.[7] Das ist dann der Fall, wenn personenbezogene Daten verarbeitet werden und ein algorithmisches Entscheidungssystem auf Grundlage dieser Daten beispielsweise darüber entscheidet, ob jemand eine staatliche Leistung erhalten oder in einem Bewerbungsverfahren für den Öffentlichen Dienst zugelassen werden soll.

Dabei können algorithmische Entscheidungssysteme positive als auch negative menschenrechtliche Relevanz entfalten. Im medizinischen Bereich, beispielsweise der radiologischen Diagnostik, können sie unterstützend wirken und so zur Verwirklichung des Rechts auf Gesundheit beitragen. Rechte von Menschen sind auch betroffen, wenn mithilfe eines algorithmischen Entscheidungssystems eine Vorhersage über menschliches Handeln getroffen wird, beispielsweise darüber, ob jemand zu einer riskanten Fahrweise im Straßenverkehr neigt oder eine Rechnung bezahlen wird. Solche Vorhersagen aufgrund von Wahrscheinlichkeiten können dann die Entscheidung, ob eine Person beispielsweise einen Kredit bekommt, wesentlich beeinflussen.

Entscheidungen, die mithilfe von Algorithmen getroffen werden, können aber nicht nur Auswirkungen auf individueller Ebene haben, sondern auch gesamtgesellschaftlich. So sortieren soziale Medien mithilfe von algorithmischen Entscheidungssystemen Nachrichten vor und zeigen den einzelnen Nutzer_innen bevorzugt solche, die das algorithmische Entscheidungssystem als relevant für diese erkannt hat. Das hat zur Folge, dass Nutzer_innen nur noch solche Nachrichten angezeigt bekommen, die denen ähneln, die sie sich vorher angesehen haben. So werden andere gesellschaftliche Diskurse zunehmend ausgeblendet. Stattdessen wird die Weltsicht der einzelnen Nutzer_innen durch die angezeigten Nachrichten bestätigt und es entsteht das Zerrbild, dass „alle“ genau so denken wie sie selbst. Eine Auseinandersetzung mit unterschiedlichen Meinungen findet so immer weniger statt.

Einige Menschenrechte, wie das Recht auf Privatsphäre und das Diskriminierungsverbot, sind durch die Anwendung algorithmischer Entscheidungssysteme besonders gefährdet; ihre Verletzung zieht meistens eine Verletzung auch anderer Menschenrechte nach sich.[8] Daneben ist auch das Recht auf Zugang zu Abhilfe von besonderer Bedeutung. Denn für Betroffene wird es schwierig, gegen eine Entscheidung Einspruch zu erheben, wenn diese algorithmengetrieben oder algorithmendeterminiert getroffen wurde. Falls sie überhaupt Kenntnis davon erlangen, dass sie in ihren Rechten verletzt wurden, bleibt es für sie fast immer unmöglich, Einblick in die Entscheidungsgrundlage des algorithmischen Entscheidungssystems zu erhalten.

Das Recht auf Privatsphäre hat im digitalen Zeitalter eine immense Bedeutung und Hebelwirkung. Wird es angemessen geschützt, ermöglicht es politische, wirtschaftliche, soziale und kulturelle Teilhabe und damit die Ausübung der mit dieser Teilhabe verbundenen Menschenrechte. Umgekehrt kann die Aushöhlung der Privatsphäre zur Verletzung und Einschränkung anderer Menschenrechte führen. So kann beispielsweise der Einsatz automatischer Gesichtserkennungssysteme zur Aushöhlung der Versammlungsfreiheit führen, wenn Menschen aus Angst davor, identifiziert und

verfolgt zu werden, nicht an einer Demonstration teilnehmen.[9] Oder das Recht auf ein faires Verfahren kann beeinträchtigt werden, wenn die Identifizierung von Straftaten oder Straftäter_innen an algorithmische Entscheidungssysteme ausgelagert wird, die falsch-positive Ergebnisse liefern oder diskriminierende Entscheidungen treffen.[10] Im Rahmen von Grenzkontrollen, insbesondere an Flughäfen, kann der Einsatz automatischer Gesichtserkennungssysteme diskriminierend wirken, wenn diese Systeme Menschen aufgrund ihrer Hautfarbe oder ihres Geschlechts nicht oder weniger gut erkennen.[11] Denn durch die Nicht- oder Falscherkennung werden sie von automatisierten Grenzkontrollen ausgeschlossen, womit negative Stereotype wiederholt und verfestigt werden. Aber auch wirtschaftliche und soziale Rechte können betroffen sein: Wenn Staaten bei der Verwaltung der Daseinsvorsorge, etwa bei der Bewilligung von Sozialleistungen, immer mehr auf von Technologie-Unternehmen entwickelte algorithmische Entscheidungssysteme vertrauen, die nicht mehr durchschaubar und hinterfragbar sind.[12]

Datenanalyse ohne gesellschaftlichen Kontext

Die Aufzählung dieser verschiedenen Beispiele macht bereits eines deutlich: In der öffentlichen Debatte überwiegen die Beispiele negativer Konsequenzen aus dem Einsatz algorithmischer Entscheidungssysteme. Grund hierfür ist die Funktionsweise dieser Systeme. Sie basieren auf mathematischen Regeln, nach denen eine Entscheidung getroffen wird, sowie Algorithmen, die dieses mathematische Modell anwendbar machen.[13] Um das Modell anzuwenden, braucht es schließlich riesige Datenmengen und eine umfassende Analyse dieser Daten. Beides, Datensätze und Datenanalyse, können Fehlerquellen sein und zu Risiken für die Wahrnehmung der Menschenrechte führen.[14] Denn diesen Daten, den „Trainingssätzen“, fehlt es am gesellschaftlichen Kontext: Ein algorithmisches Entscheidungssystem identifiziert einen Menschen auf einem Bild nicht, weil es weiß, was ein Mensch ist, sondern weil es mit großen Mengen an Bildern darauf trainiert wurde, einen Menschen zu erkennen. Zeigen diese „Trainingssätze“ nur bestimmte Menschen, beispielsweise nur Frauen oder nur Personen bestimmter Herkunft, können algorithmische Entscheidungssysteme andere Menschen auch nicht als solche erkennen. Sie sind daher fehleranfällig: Ihre Entscheidungen basieren auf den zugrundeliegenden, realitäts- und damit wertbehafteten Trainingsdaten.[15]

Dabei kann Automatisierung, also die Auslagerung einzelner Schritte im Entscheidungsprozess an eine Maschine, auch Diskriminierung entgegenwirken, wenn Entscheidungen auf objektiven Kriterien beruhen.[16] Da jedoch die Datensätze, die zur Automatisierung von Entscheidungsschritten herangezogen werden, aus der gesellschaftlichen Realität gewonnen werden, enthalten sie genau die Vorurteilsstrukturen, die in der Geselschaft zu Diskriminierung führen. Ein bekanntes Beispiel dafür sind Bewerbungsverfahren, die auf die Daten vergangener Bewerbungsverfahren zurückgreifen, um die besten Kandidat_innen auszuwählen. Wurden in der Vergangenheit vorzugsweise Männer eingestellt, könnte das algorithmische Entscheidungssystem in seiner Analyse der Daten folgern, dass Männer bessere Kandidaten als Frauen sind.[17] Probleme können aus den Datensätzen und der Datenanalyse auch dadurch entstehen, dass anonymisierte Daten de-anonymisiert werden oder die Datenanalyse Korrelationen entdeckt und verarbeitet, welche keine kausalen Zusammenhänge haben.[18]

Automatisierung kann zudem Transparenz, Untersuchbarkeit und Erklärbarkeit einer Entscheidung verhindern.[19] Denn selbstlernende algorithmische Entscheidungssysteme sind nicht (mehr) nur auf die von (menschlichen) Programmierer_innen entworfenen Entscheidungsregeln angewiesen, sondern in der Lage, eigene neue Entscheidungsfindungsregeln aufzustellen und anzuwenden.[20] Diese Anpassungs- und Lernfähigkeit hat erheblichen Einfluss darauf, ob die getroffene Entscheidung transparent und nachvollziehbar ist und führt damit zur Frage, wer genau für eine Entscheidung verantwortlich ist.[21]

Staatenpflichten beim Einsatz algorithmischer Entscheidungssysteme

Staaten sind verpflichtet, die Menschenrechte zu achten, zu schützen und zu gewährleisten. Die Achtungspflicht bedeutet, dass der Staat Menschenrechte nicht verletzen darf, wenn er in der Interaktion mit seinen Bürger_innen algorithmische Entscheidungssysteme anwendet. Wenn ein Staat beispielsweise bei der Bearbeitung und Bewilligung von Sozialleistungen auf algorithmische Entscheidungssysteme zurückgreift, muss er sicherstellen, dass er dabei nicht diskriminiert oder unverhältnismäßig in die Privatsphäre eingreift.

Der Staat muss darüber hinaus Menschenrechte gewährleisten, also ermöglichen, dass Menschen ihre Menschenrechte ausüben. Algorithmische Entscheidungssysteme können hierbei unterstützend wirken. Um wirkungsvolle politische Maßnahmen ergreifen zu können, ist die Erhebung und Analyse von Daten oft unerlässlich. Daher rufen auch Menschenrechtsgremien Staaten immer wieder dazu auf, aussagekräftige Daten zu sammeln und politische Maßnahmen auf Grundlage der gewonnenen Erkenntnisse zu ergreifen. Selbstlernende algorithmische Entscheidungssysteme können dabei helfen, große Datenmengen zu analysieren oder auch Strukturen und Zusammenhänge zu entdecken, nach denen bisher vielleicht nicht gesucht wurde.[22] Ebenso können algorithmische Entscheidungssysteme genutzt werden, um aus einer Gruppe von Personen jene zu identifizieren, die Anspruch auf besondere Unterstützungsleistungen haben, beispielsweise um einem Studienabbruch von Student_innen entgegenzuwirken.[23]

Schließlich ist der Staat verpflichtet, seine Bürger_innen auch vor Menschenrechtsverstößen durch Dritte schützen, zum Beispiel vor Diskriminierung, die entstehen kann, wenn Unternehmen algorithmische Entscheidungssysteme anwenden. So dürfen Arbeitgeber_innen bei Einstellungen beispielweise nicht nach Geschlecht oder Herkunft diskriminieren. Dieses Verbot gilt natürlich auch, wenn bei der Auswahl algorithmische Entscheidungssysteme genutzt werden. Staaten sind also in der Pflicht, regulierend einzugreifen, um solche Menschenrechtsverstöße zu verhindern.

Normative Vorgaben und Entwicklungen auf internationaler Ebene

Die Menschenrechte sind in einer Reihe von Verträgen der Vereinten Nationen auf internationaler Ebene verankert. Dazu gehören der Zivil- und der Sozialpakt, aber auch gruppenspezifische und thematische Menschenrechtsverträge wie die UN-Behindertenrechtskonvention oder die Anti-Rassismus-Konvention, die beide insbesondere in Hinblick auf Diskriminierungsverbote und Teilhaberechte von Bedeutung sind.

Die Menschenrechtsverträge kennen keine speziellen „digitalen Rechte“. Den Grund- und Menschenrechten liegt aber immer ein dynamisches Verständnis zugrunde, ihre Interpretation orientiert sich also an der jeweiligen Lebenswirklichkeit, mit der sie konfrontiert werden, sowie den technologischen und gesellschaftlichen Weiterentwicklungen, welche sich im Laufe der Zeit ergeben. Alle auf internationaler Ebene geschützten Menschenrechte entfalten ihre Wirkung deshalb offline wie online[24] und somit auch mit Blick auf die Anwendung algorithmischer Entscheidungssysteme. Auch der UN-Menschenrechtsrat weist explizit darauf hin, dass Menschenrechte auch im digitalen Raum zu achten, zu schützen und zu gewährleisten sind.[25] Verschiedene UN-Gremien haben sich inzwischen mit den Auswirkungen algorithmischer Entscheidungssysteme auf die Menschenrechte befasst.

2019 griff die **UN-Generalversammlung** in ihrer Resolution zum Recht auf Privatsphäre im digitalen Zeitalter das Thema auf. Die Resolution hebt die Notwendigkeit hervor, algorithmische Entscheidungssysteme auf der Grundlage von Menschenrechten zu gestalten, zu evaluieren und zu regulieren und ruft Unternehmen dazu auf, die Achtung des Rechts auf Privatsphäre und anderer Menschenrechte in der Gestaltung, dem Betrieb, der Bewertung und der Regulierung von algorithmischen automatisierten Entscheidungsfindungssystemen sicherzustellen, sowie für die Behebung der Menschenrechtsverstöße zu sorgen, die sie verursacht oder zu denen sie beigetragen haben.[26]

In seiner 41. Sitzung (Juli 2019) beauftragte der **UN-Menschenrechtsrat** sein Beratungsgremium (advisory committee), zur 47. Sitzung (Juni 2021)

eine Studie über die möglichen Auswirkungen, Chancen und Herausforderungen neuer und neu entstehender digitaler Technologien im Hinblick auf die Förderung und den Schutz der Menschenrechte vorzulegen.[27] Anschließend nahm der UN-Menschenrechtsrat in seiner 42. Sitzung in seiner Resolution zum Recht auf Privatsphäre im digitalen Zeitalter den Zusammenhang zwischen algorithmischen Entscheidungssystemen und Diskriminierung in den Blick. Darin erkennen die Mitgliedsstaaten an, dass die Gestaltung, Entwicklung und Einführung neuer und neu entstehender Technologien menschenrechtsbasiert reguliert werden muss, um das Recht auf Privatsphäre angemessen zu schützen. Zu dieser Regulierung gehöre „eine sichere, gesicherte und qualitativ hochwertige Dateninfrastruktur und durch die Entwicklung von auf den Menschen ausgerichteten Prüfungsmechanismen sowie von Rechtsbehelfsmechanismen".[28] Auf Grundlage der Resolution fand im Mai 2020 ein eintägiges Expertenseminar statt[29], der in Auftrag gegebene thematische Bericht der UN-Hochkommissarin für Menschenrechte liegt noch nicht vor.[30]

In seiner Allgemeinen Empfehlung Nr. 36 zur Prävention und Bekämpfung von Racial Profiling durch Strafverfolgungsbehörden hat der **UN-Fachausschuss gegen rassistische Diskriminierung** auch das Thema Polizeiarbeit mit Hilfe von algorithmischen Entscheidungssystemen aufgegriffen.[31] Das kann sowohl die Aufklärung als auch die Vorhersage von Straftaten (predictive policing) betreffen. Der Ausschuss empfiehlt daher umfassende Maßnahmen, um Menschenrechte beim Einsatz von algorithmischen Entscheidungssystemen umfassend zu schützen.

Mehrere UN-Sonderberichterstatter haben sich explizit mit den Auswirkungen von algorithmischen Entscheidungssystemen auf die Menschenrechte beschäftigt. In seinem Bericht zu den Auswirkungen von Technologien künstlicher Intelligenz auf die Menschenrechte, insbesondere auf das Recht auf Meinungsfreiheit und freie Meinungsäußerung, auf Privatsphäre und Nichtdiskriminierung hebt der **UN-Sonderberichterstatter zur Meinungsfreiheit** hervor, dass Staaten (a) eine menschenrechtliche Auswirkungsanalyse (human rights impact assessment) durchführen sollten, bevor sie algorithmische Entscheidungssysteme einsetzen; und (b) sicherstellen, dass sie Regulierungsmaßnahmen ergreifen, damit Menschenrechte bei Design, Bereitstellung und Einsatz von algorithmischen Entscheidungssystemen durch Unternehmen im Vordergrund stehen.[32]

Der **UN-Sonderberichterstatter zu extremer Armut und Menschenrechte** hat sich 2019 mit den Auswirkungen des Einsatzes algorithmischer Entscheidungssysteme in Bezug auf sozialstaatliche Leistungen beschäftigt („digital welfare state dystopia"). Er zeigt anhand von vielen Beispielen aus allen Regionen der Welt, wie die Nutzung von algorithmischen Entscheidungssystemen für die Verwaltung von Sozialstaatsleistungen zu Reproduktion und Verstärkung von Diskriminierung und Ungleichheit geführt hat. Die Digitalisierung der Daseinsvorsorge birgt auch die Gefahr, dass – unter dem Anspruch, Abläufe effizienter zu gestalten und Kosten zu sparen – Bürger_innen nur noch von „maschinellen" Entscheidungen verwaltete Antragsteller_innen sind und nicht mehr Rechteinhaber_innen. Daher empfiehlt der UN-Sonderberichterstatter (a) die explizite Prüfung der verwendeten Datensätze auf Menschenrechtsstandards, (b) Transparenzpflichten, damit die Öffentlichkeit und Betroffene die eingesetzten algorithmischen Entscheidungssysteme verstehen, und (c) dass Staaten Daseinsvorsorge in einer Art betreiben, die „die Menschlichkeit und die Belange der weniger Wohlhabenden in jeder Gesellschaft in vollem Umfang berücksichtigt". Technologische Entwicklung, so der Sonderberichterstatter, hat durchaus das Potenzial, dazu beizutragen, „einen höheren Lebensstandard für die Schwachen und Benachteiligten zu gewährleisten".[33]

Anfang 2021 veröffentlichte der **UN-Sonderberichterstatter zum Recht auf Privatsphäre** Empfehlungen zum Schutz des Rechts auf Privatsphäre in der Entwicklung und dem Einsatz von algorithmischen Entscheidungssystemen. Darin stellt er acht grundsätzliche Datenschutzerwägungen für den Umgang mit algorithmischen Entscheidungssystemen vor: rechtliche Zuständigkeit, ethische und gesetzliche Datenverarbeitungsgrundlage, Datengrundlagen, Verantwortung und Aufsicht, Kontrolle, Transparenz und Erklärbarkeit, Rechte betroffener Personen und Schutzmaßnahmen. Auch er plädiert für eine umfassende menschenrechtliche Auswirkungsanalyse.[34]

Elemente eines menschenrechtsbasierten Umgangs mit algorithmischen Entscheidungssystemen

Aus den oben genannten Berichten, Resolutionen und Allgemeinen Empfehlungen lassen sich folgende Schlussfolgerungen für Elemente eines menschenrechtsbasierten Ansatzes im Umgang mit algorithmischen Entscheidungssystemen ziehen:

Menschenrechtliche Auswirkungen prüfen: Behörden sollten vor der Entscheidung, ein algorithmisches Entscheidungssystem entwickeln zu lassen oder zu nutzen, die möglichen menschenrechtlichen Auswirkungen untersuchen (human rights impact assessment) und auch während des Betriebs regelmäßig die menschenrechtlichen Auswirkungen überprüfen.[35] Dies schließt sowohl eine Selbstbegutachtung als auch externe Begutachtung ein.[36] Dafür können auch zunächst Pilotprojekte unter der Aufsicht von unabhängigen und divers besetzten Gremien initiiert werden.[37]

Diskriminierungsrisiken ausschließen: Menschenrechtliche Auswirkungsanalysen sollten immer auch untersuchen, ob das algorithmische Entscheidungssystem diskriminierende Wirkung entfalten könnte oder entfaltet.[38] Damit muss ein Fokus auf die verwendeten Datensätze gelegt werden: Enthalten sie Daten, in dem eine Personengruppe bevorzugt wird? Ist das Entscheidungssystem so programmiert, dass es aus Datensätzen selbstlernend diskriminierende Näherungsvariablen (sogenannte Proxies) entwickelt und zur Lösungs- oder Entscheidungsfindung heranzieht und so gegebenenfalls diskriminierende Wirkung entfaltet?[39]

Partizipation sicherstellen: Zu einem menschenrechtsbasierten Umgang mit algorithmischen Entscheidungssystemen gehört auch die Beteiligung potenzieller Nutzer_innengruppen und Betroffener. Diese sollten im Entwicklungs- und Beschaffungsstadium, aber auch während der Nutzung algorithmischer Entscheidungssysteme durch öffentliche Beteiligungsformate einbezogen werden.[40]

Transparenz und Information: Mit der Anwendung von algorithmischen Entscheidungssystemen gehen auch umfassende Transparenz- und Informationspflichten einher. Information und Transparenz bedeuten nicht die komplette Offenlegung von Codes; mit solchen Informationen können die wenigsten Menschen etwas anfangen. Auch rechtlich kann eine komplette Offenlegung problematisch sein. Stattdessen sollten Nutzer_innen über die Existenz, den Zweck, die Beschaffenheit und die Wirkung des algorithmischen Entscheidungssystems informiert werden,[41] ebenso über Maßnahmen, die getroffen wurden, um die Verletzung von Menschenrechten zu verhindern.[42] Bürger_innen müssen wissen, dass eine für sie relevante Entscheidung durch oder mithilfe von algorithmischen Entscheidungssystemen getroffen wird. Dazu gehört auch, dass deutlich gemacht wird, wo und/oder wann Menschen im Entscheidungsprozess involviert sind, welcher Logik algorithmische Entscheidungssysteme folgen und welche Konsequenzen die Entscheidung im Einzelfall haben wird.[43] Weiterhin muss es zumindest für Fachleute möglich sein, die technische Funktionsweise und die Entscheidungen des algorithmischen Entscheidungssystems nachzuvollziehen[44] und so auch die Justiz bei der Prüfung, ob Menschenrechte durch das algorithmische Entscheidungssystem verletzt wurden, zu unterstützen.

Zugang zum Recht: Betroffene müssen Zugang zu ausreichendem, wirksamem, raschem und angemessenem Rechtsschutz bekommen können und algorithmisch getroffene Entscheidungen anfechten können. Vorfälle von Diskriminierung und andere negative menschenrechtliche Auswirkungen algorithmischer Entscheidungssysteme sollten umfassend untersucht und auch sanktioniert werden.[45] Zum Rechtsschutz gehört auch, dass Verantwortlichkeiten vorab geregelt werden und immer eine juristische oder natürliche Person verantwortlich sein muss.[46] Ansonsten können Widersprüche gegen Entscheidungen, die algorithmische Entscheidungssysteme getroffen haben, nicht bearbeitet werden und Missbräuche nicht den jeweiligen Verantwortlichen angelastet werden. Fehlender Rechtsschutz könnte dann auch zu Akzeptanzproblemen der verwendeten Technologie, aber insbesondere auch zu schwindendem Vertrauen gegenüber staatlichen Institutionen führen.

Regulierung: Schließlich müssen Staaten regulierend eingreifen, damit die genannten Standards auch von Unternehmen eingehalten werden, die algorithmische Entscheidungssysteme (oft im staatlichen Auftrag) entwickeln und anwenden.[47] Dazu gehört auch die regelmäßige Evaluierung und Überarbeitung solcher Standards. Dies sollte auch die Gründung (oder Beauftragung vorhandener)

unabhängiger Aufsichtsbehörden umfassen, welche algorithmische Entscheidungssysteme auf die Einhaltung von Menschenrechtsstandards hin überprüfen können. Ergeben solche Prüfungen ein hohes Diskriminierungs- oder anderes menschenrechtliches Risiko, sollte auf den Einsatz der entsprechenden Systeme verzichtet werden.[48]

Algorithmische Entscheidungssysteme sind keine selbstständig denkenden Maschinen, sondern technische Hilfsmittel, die sehr große Datenmengen nach vorgegebenen oder selbstgelernten Algorithmen analysieren können. Sie sind damit in bestimmten Bereichen sehr leistungsfähig. Mit Blick auf die bisherige Praxis überwiegen aus menschenrechtlicher Perspektive Sorge bereitende Beispiele. Die menschenrechtlichen Auswirkungen von algorithmischen Entscheidungssystemen werden vor ihrem Einsatz nur unzureichend analysiert – wenn dies überhaupt getan wird. Dieser Befund gilt gleichermaßen für Staaten und für Unternehmen, die algorithmische Entscheidungssysteme entwickeln.

Staaten müssen dafür Sorge tragen, dass sie bei der Anwendung von algorithmischen Entscheidungssystemen Menschenrechte achten, schützen, fördern und gewährleisten. Dies kann auch bedeuten, dass der Einsatz algorithmischer Entscheidungssysteme in manchen Bereichen ausgeschlossen werden muss. Der Verordnungsentwurf der EU-Kommission zur Regulierung von KI geht durch die Klassifizierung mancher Anwendungen als „unannehmbares Risiko“ einen ersten Schritt in diese Richtung.

---

1 European Commission (2021): Proposal for a Regulation laying down harmonised rules on artificial intelligence. COM(2021) 206 final; European Commission (2020): Weißbuch zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final.

2 Bundesregierung (02.12.2020): Künstliche Intelligenz – „Die entscheidende Zukunftstechnologie des 21. Jahrhunderts“. https://www.bundesregierung.de/breg-de/aktuelles/fortschreibung-ki-strategie-1824340 (abgerufen am 24.05.2021); Deutscher Bundestag (28.10.2020): Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale. Drucksache 19/23700.

3 Access Now (2018): Human Rights in the Age of Artificial Intelligence, S. 8. https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf (abgerufen am 17.01.2020).

4 Für eine umfassende Darstellung siehe beispielsweise OECD (2019): Artificial Intelligence in Society. Paris: OECD Publishing.

5 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018): Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/73/348, Ziff. 4.

6 Datenethikkommission der Bundesregierung (2019): Gutachten der Datenethikkommission, S. 24. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html (abgerufen am 21.02.2020).

7 Zweig, Katharina A. (2019): Algorithmische Entscheidungen: Transparenz und Kontrolle. In: Konrad-Adenauer-Stiftung, Analysen & Argumente 2019 (338), S. 5.

8 UN, Special Raporteur on the right to privacy (2021): Report of the Special Raporteur on the right to privacy, UN Doc. A/HRC/46/37, Ziff. 2, 13.

9 Access Now (2018), a.a.O., S. 21.

10 Council of Europe (2018): Algorithms and Human Rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications, DGI (2017)12, S. 10.

11 UN, Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance (2020): Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, UN Doc. A/75/590, Ziff. 9.

12 UN, Special Rapporteur on extreme poverty and human rights (2019): Report of the Special Rapporteur on extreme poverty and human rights, UN Doc. A/74/493.

13 Algorithm Watch (2019): Atlas der Automatisierung. Automatisierte Entscheidungen und Teilhabe in Deutschland, S. 9. https://atlas.algorithmwatch.org/wp-content/uploads/2019/07/Atlas_der_Automatisierung_von_AlgorithmWatch.pdf (abgerufen am 17.01.2020).

14 Council of Europe (2018), a.a.O., S. 10.

15 Beispielsweise kann es passieren, dass ein Bilderkennungsprogramm Küchen automatisch mit Frauen assoziiert, weil die vorhandenen Trainingsdaten überwiegend Küchen und Frauen abbilden. Süddeutsche.de (27.08.2017): Sexismus im Algorithmus. https://www.sueddeutsche.de/digital/netzkolumne-vorurteile-im-algorithmus-1.3642503 (abgerufen am 21.05.2021).

16 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 6.

17 The Verge (10.10.2018): Amazon reportedly scraps internal AI recruiting tool that was biased against women. The secret program penalized applications that contained the word “women’s”. https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report (abegrufen am 21.05.2021).

18 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 7.

19 Ebd., Ziff. 6.

20 Council of Europe (2018), a.a.O., S. 6.

21 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 8.

22 The Guardian (20.02.2020): Powerful antibiotic discovered using machine learning for first time. https://www.theguardian.com/society/2020/feb/20/antibiotic-that-kills-drug-resistant-bacteria-discovered-through-ai (abgerufen am 24.05.2021).

23 Martini, Mario u.a. (2020): Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand. Gütersloh: Bertelsmann Stiftung, S. 22.
24 UN, General Assembly (2013): Resolution "The right to privacy in the digital age", UN Doc. A/RES/68/167, OP 3.
25 UN, Human Rights Council (2018): Resolution "The promotion, protection and enjoyment of human rights on the Internet", UN Doc. A/HRC/RES/38/7, OP 1.
26 UN, General Assembly (2018), Resolution "The right to privacy in the digital age", UN Doc. A/RES/73/179, PP 19, OP 7 (d).
27 UN, Human Rights Council (2019): Resolution "New and emerging digital technologies and human rights", UN Doc. A/HRC/RES/41/11, OP 1; die Studie wird voraussichtlich unter UN Docs. A/HRC/47/52 erscheinen, UN, Human Rights Council (2021): Agenda and annotations, UN Doc. A/HRC/47/1, Ziff. 82.
28 UN, Human Rights Council (2019): Resolution "The right to privacy in the digital age", UN Doc. A/HRC/RES/42/15, OP 5.
29 UN, Office of the High Commissioner for Human Rights: Expert Seminar on Artificial Intelligence and the Right to Privacy, 27-28 May 2020. Report of the proceedings. https://www.ohchr.org/Documents/Issues/DigitalAge/ExpertSeminarReport-RightPrivacy.pdf (abgerufen am 21.05.2021).
30 UN, Human Rights Council (2021): Agenda and annotations, UN Doc. A/HRC/47/1, Ziff. 46.
31 UN, Committee on the Elimination of Racial Discrimination (2020): General recommendation No. 36 (2020) on preventing and combating racial profiling by law enforcement officials, UN Doc. CERD/C/GC/36, Ziff. 21–36.
32 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 62–63.
33 UN, Special Rapporteur on extreme poverty and human rights (2019), a.a.O., Ziff. 50, 81–84.
34 UN, Special Raporteur on the right to privacy (2021), a.a.O., Ziff. 27–49, 50–61.
35 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 60; UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 53.
36 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 53.
37 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 59.
38 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 58; UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 52.
39 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 52.
40 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 60; UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 54.
41 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 50.
42 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 61.
43 UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 49.
44 Ebd., Ziff. 56.
45 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 65.
46 UN, Special Raporteur on the right to privacy (2021), a.a.O., Ziff. 9.
47 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 63–64; UN, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), a.a.O., Ziff. 63.
48 UN, Committee on the Elimination of Racial Discrimination (2020), a.a.O., Ziff. 62, 63.

Impressum

Information Nr. 35 | Juni 2021 | ISSN 2509-9493 (PDF)

HERAUSGEBER: Deutsches Institut für Menschenrechte
Zimmerstraße 26/27 | 10969 Berlin
Tel.: 030 259 359-0 | Fax: 030 259 359-59
info@institut-fuer-menschenrechte.de
www.institut-fuer-menschenrechte.de

AUTORIN: Rosa Öktem

LIZENZ:
Creative Commons (CC BY-NC-ND 4.0)
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

Das Institut

Das Deutsche Institut für Menschenrechte ist die unabhängige Nationale Menschenrechtsinstitution Deutschlands. Es ist gemäß den Pariser Prinzipien der Vereinten Nationen akkreditiert (A-Status). Zu den Aufgaben des Instituts gehören Politikberatung, Menschenrechtsbildung, Information und Dokumentation, anwendungsorientierte Forschung zu menschenrechtlichen Themen sowie die Zusammenarbeit mit internationalen Organisationen. Es wird vom Deutschen Bundestag finanziert. Das Institut ist zudem mit dem Monitoring der Umsetzung der UN-Behindertenrechtskonvention und der UN-Kinderrechtskonvention betraut worden und hat hierfür entsprechende Monitoring-Stellen eingerichtet.