



MUN-SH 2025
Menschenrechtsrat

Maschinelles Lernen als Diskriminierungs- verstärker

Zusammenfassung

Maschinelles Lernen (ML) basiert auf Algorithmen, die sich selbst verbessern und anpassen können. Diese Anpassungen ermöglichen es, Daten zu interpretieren. Allerdings kann ML bestehende Vorurteile und Diskriminierung verstärken, insbesondere durch fehlerhaftes Design des Systems oder unausgewogene Trainingsdaten. Wenn bestimmte Bevölkerungsgruppen in den Daten nicht ausreichend vertreten sind, übernimmt das ML-System diese Ungleichheiten und verstärkt sie durch Feedback-Schleifen. Auch die Zielsetzung der ML-Anwendung und ungleicher Zugang zu Technologien können Diskriminierung fördern. Das Risiko für solche Verzerrungen wird zusätzlich durch die Unterrepräsentation bestimmter Gruppen bei der Entwicklung von ML-Systemen erhöht.

ML-Systeme werden in zahlreichen Anwendungsbereichen eingesetzt, auch von staatlichen Institutionen. Die Formen der dadurch entstehenden oder verstärkten Diskriminierung sind vielfältig und reichen von ungleichen Behandlungsergebnissen bis hin zu sozialen Benachteiligungen.

Sowohl die Menschenrechte als auch der Internationale Pakt über bürgerliche und politische Rechte verbieten Diskriminierung. Staaten sind verpflichtet, ihre Bürger*innen vor diskriminierenden Praktiken zu schützen, und können dies durch Richtlinien, Gesetze sowie die Förderung von Transparenz und Überprüfungsmechanismen in ML-Systemen tun. Dies umfasst die Kontrolle von Trainingsdaten auf Ungleichheiten und die Analyse von Systemergebnissen auf diskriminierende Tendenzen. Bildungsinitiativen können zudem das Bewusstsein für Diskriminierung durch ML erhöhen.

Verschiedene UN-Institutionen haben Netzwerke von Expert*innen und Projekte ins Leben gerufen, um die internationale Zusammenarbeit in diesem Bereich zu fördern. Der Menschenrechtsrat (MRR) befasst sich ebenfalls regelmäßig mit diesem Thema, fordert Berichte an und verabschiedet Resolutionen, die stärkere Regulierung, internationale Kooperation und die Förderung digitaler Bildung fordern.

Dennoch bleibt viel zu tun. Die Staatengemeinschaft muss Lösungen finden, die alle Bevölkerungsgruppen mit einbeziehen, klare Verantwortlichkeiten festlegen, Richtlinien vorgeben und ein effektives Vorgehen gegen Diskriminierung ermöglichen.

Punkte zur Diskussion

- Welche Maßnahmen können Staaten treffen, um sicherzustellen, dass die Entwicklung und der Einsatz von maschinellem Lernen (ML) nicht zur Verstärkung von Diskriminierung führen?
- Wie stark sollten Staaten die Nutzung und Entwicklung von ML in Hinsicht auf mögliche Diskriminierung regulieren und in welchen Bereichen und welche verbindlichen und unverbindlichen Richtlinien sollten gelten?
- Wie können Staaten dafür sorgen, dass bei der Entwicklung und dem Einsatz von ML größtmögliche Transparenz herrscht und die Verantwortung bestimmter Menschen, Institutionen oder Unternehmen für Entscheidungen von KI eindeutig geklärt ist?
- Welche Maßnahmen müssen getroffen werden, um sicherzustellen, dass marginalisierte Bevölkerungsgruppen in allen Bereichen, die ML betreffen, von der Entwicklung über die Möglichkeit zur Nutzung bis hin zur Regulierung von ML stärker teilhaben können?
- Welche Maßnahmen können eingesetzt werden, um Menschen Mittel zu geben, gegen Diskriminierung durch ML vorzugehen, sodass anhaltende Diskriminierung beendet wird und Betroffene entschädigt werden?
- Sollte der Einsatz von ML-Systemen in sensiblen Bereichen wie Strafverfolgung aufgrund potenzieller Diskriminierungsrisiken eingeschränkt oder ganz verboten werden?
- Durch welche Maßnahmen kann ein bewusster und verantwortungsvoller Umgang mit ML in der Bevölkerung erreicht werden? Welche Rolle spielen Bildungsmaßnahmen und Aufklärungskampagnen dabei?

Hintergrund und Grundsätzliches

Maschinelles Lernen (ML) wird oft als leistungsstarke Technologie gepriesen, die in der Lage ist, komplexe Probleme zu lösen, von personalisierten Empfehlungen bis hin zur medizinischen Diagnose. Doch diese Technologie birgt auch Gefahren, insbesondere wenn es um die Verstärkung von Diskriminierung geht. **Algorithmen**, die mit großen Datenmengen trainiert werden, spiegeln oft die Verzerrungen wider, die in den zugrunde liegenden Daten vorhanden sind. Wenn diese Verzerrungen unreflektiert übernommen werden, können sie soziale Ungerechtigkeiten verschärfen und bestehende Diskriminierung – beispielsweise aufgrund von Geschlecht, Hautfarbe oder sozialem Status – weiter verstärken. Dieser Text beleuchtet die Mechanismen, durch die maschinelles Lernen zur Diskriminierung beitragen kann, und diskutiert, wie diese Risiken erkannt und minimiert werden können.

Zunächst sollten ein paar Konzepte zum Thema maschinelles Lernen geklärt werden. Dazu hilft es, den Begriff mit einem anderen zu vergleichen, den man deutlich häufiger hört: Künstliche Intelligenz (KI). Während es beim maschinellen Lernen schlicht um lernfähige Algorithmen geht, ist der Begriff der KI nur sehr vage definiert. Grob geht es bei KI darum, dass eine Maschine die Fähigkeit entwickelt, in einer menschenähnlichen Weise zu lernen, zu verstehen, zu interpretieren und Entscheidungen zu treffen. Maschinelles Lernen bietet die Grundlage hierfür. Oft werden die beiden Begriffe aber auch synonym verwendet.

Ein **Algorithmus** ist eine festgelegte Folge von Anweisungen oder Regeln, die ein Computer befolgt, um ein spezifisches Problem zu lösen oder eine bestimmte Aufgabe zu erfüllen.

Verarbeitungsschichten

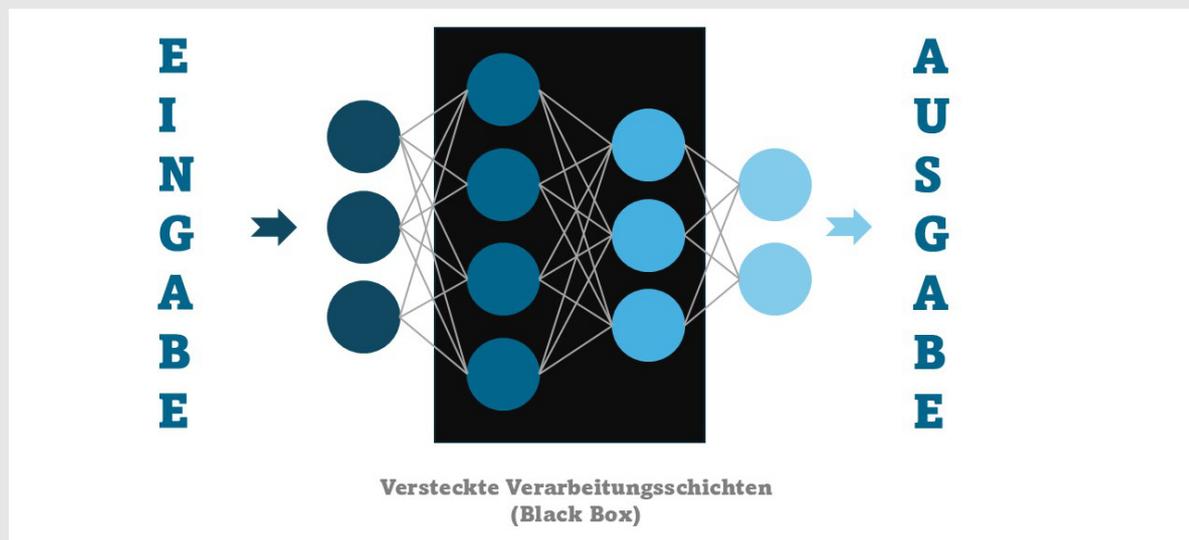
in einem maschinellen Lernsystem sind die verschiedenen Ebenen, durch die Daten fließen, während sie analysiert und bearbeitet werden. Jede Schicht führt spezifische Berechnungen oder Transformationen durch, um die Rohdaten schrittweise in eine verwertbare Form umzuwandeln.

Ein **Parameter** ist eine anpassbare Größe oder Variable innerhalb eines Algorithmus oder Modells, die dessen Verhalten und Ausgaben beeinflusst.

ML-Algorithmen interpretieren Eingabewerte, wie etwa Fotos, mithilfe mehrerer **Verarbeitungsschichten**. Dabei werden die Daten analysiert und gewichtet, bis ein Ergebnis entsteht, das zum Beispiel als die Erkennung eines Hundes oder einer Katze gedeutet werden kann. Wenn der Algorithmus nun beispielsweise mit einer großen Menge an Hunde- und Katzenfotos trainiert wird, passt er seine **Parameter** immer weiter an, bis er einen Großteil der Fotos erkennt. KI kann durch dieses System Gesichter und Objekte erkennen (analytische KI), Texte generieren (generative KI) und Vorhersagen treffen.



Quelle: DMUN (Johanna Lührs)



Risikofaktoren von ML zur Verstärkung von Diskriminierung

Eine **Feedback-Schleife** ist ein Prozess, bei dem die Ergebnisse einer Handlung oder Berechnung wieder zurück in das System fließen und als Grundlage für die nächste Runde von Handlungen oder Berechnungen dienen. Dabei können sich bestimmte Muster verstärken oder abschwächen.

Predictive Policing ist eine Methode der Kriminalprävention, bei der mittels Datenanalyse und Algorithmen Vorhersagen über wahrscheinliche Tatorte und -zeiten getroffen werden, um Polizeipräsenz gezielt einzusetzen.

Wie stehen nun also ML und Diskriminierung in Verbindung?

Unter Diskriminierung versteht man die ungerechtfertigte Benachteiligung von Personen aufgrund von geschützten Merkmalen. Solche geschützten Merkmale sind zum Beispiel Alter, Geschlecht, ethnische Herkunft, Religion, sexuelle Orientierung oder Behinderung.

ML-Systeme können Diskriminierung übernehmen und durch sogenannte **Feedback-Schleifen** verstärken. Ein Beispiel ist '**predictive policing**': ML-Systeme sagen voraus, in welchen Gegenden Straftaten wahrscheinlich sind. In diesen Gebieten wird die Polizeipräsenz erhöht, wodurch mehr Straftaten registriert werden. Diese Daten fließen erneut in die Vorhersagen ein, wodurch die Gegend als zunehmend gefährlich eingestuft wird.

Schon beim Designen eines ML-Systems können Entscheidungen getroffen werden, die eine unverhältnismäßige Ungleichbehandlung nach sich ziehen. Häufig sind bestimmte Bevölkerungsgruppen,

etwa Frauen, in der Entwicklung von ML-Systemen immer noch stark unterrepräsentiert. Zu Diskriminierung kann es zum Beispiel kommen, wenn voreingenommenen Kriterien für die Definition einer erfolgreichen Entscheidung des Systems festgelegt werden.

Die wohl größte Anfälligkeit für Diskriminierung liegt allerdings in den Trainingsdaten eines ML-Systems. Wenn diese unvollständig, nicht repräsentativ sind oder historisch bedingt vorurteilsbehaftet sind, wirkt sich das stark auf das ML-System aus. Der Chatbot Tay von Microsoft wurde beispielsweise durch die Interaktion mit Nutzer*innen auf Twitter trainiert. So übernahm er Hass gegenüber Frauen und People of Color. Ein ML-System von Amazon sollte Bewerbungen bewerten und hat aufgrund der bei Amazon vorherrschenden Geschlechterverteilung Männer stark bevorzugt. Des Weiteren führen Worte wie "Muslim", "schwul" oder "Jude" bei der Erkennung beleidigender Sprache oft fälschlicherweise dazu, dass ein Text als beleidigend eingestuft wird. Auch staatlich eingesetzte ML-Systeme neigen nicht selten aufgrund mangelhafter Trainingsdaten zu Diskriminierung. In den Niederlanden etwa wurde eine Reihe an Menschen, die größtenteils einen Migrationshintergrund hatten, aufgrund der Einschätzung eines ML-Systems fälschlicherweise der Steuerhinterziehung beschuldigt. Das wirft auch die Frage auf, ob Staaten in sensiblen Bereichen wie der Strafverfolgung vollständig verzichten sollten.

Kommen wir von den Risiken in der Entwicklung von ML-Systemen nun zur Nutzung. Zunächst kann Diskriminierung schlicht durch die Zielsetzung des Einsatzes der ML-Systeme erfolgen. Ersichtlich ist das beispielsweise bei **autonomen Waffensystemen**. Die Kriterien, die von solchen Systemen genutzt werden, um Ziele zu wählen, beinhalten wahrscheinlich immer wieder auch Alter, Geschlecht oder ethnische Abstammung. Doch selbst, wenn ein ML-System weder diskriminierend agiert noch eingesetzt wird, kann es zu Diskriminierung kommen, und zwar durch ungleiche Zugriffsmöglichkeiten. Beispielsweise haben Frauen weltweit noch immer deutlich schlechterer Zugang zum Internet als Männer. Wenn beispielsweise Dienstleistungen zunehmend durch KI bereitgestellt werden, kann Benachteiligung durch ungleichen Zugang zu diesen Systemen verstärkt werden. Die Staatengemeinschaft sollte also gemeinsam dafür sorgen, dass marginalisierte Gruppen in Entwicklung, Nutzung und Entscheidungsprozessen zur Regulierung von ML mit einbezogen werden.

Autonome Waffensysteme sind militärische Systeme, die durch Künstliche Intelligenz und Automatisierung in der Lage sind, ohne menschliche Eingriffe Ziele auszuwählen und anzugreifen.



Ein Potential für Diskriminierung entsteht offensichtlich dann, wenn geschützte Merkmale Teil der Eingabe für ML-Systeme sind, also wenn beispielsweise einem ML-System, das über Bewerbungen entscheidet, Daten zu Geschlecht oder Herkunft einer Person zur Verfügung gestellt werden. Doch Vorsicht gilt auch vor sogenannten **Proxies**. Das sind scheinbar neutrale Informationen, die stark mit einer geschützten Charakteristik korrelieren, wie beispielsweise bestimmte Namen mit ethnischer Herkunft.

Hinzu kommt, dass ML-Systeme als ‚**Black Box**‘ gelten. Das bedeutet, es ist fast unmöglich nachzuvollziehen, wie das System zu seinen Entscheidungen kommt. Dies erschwert die Zuweisung von Verantwortung für fehlerhafte oder diskriminierende Entscheidungen.

Es wird deutlich, dass ML-Systeme häufig keineswegs neutral agieren. Dennoch werden sie von vielen Menschen als objektiv und zuverlässig wahrgenommen, wodurch das Risiko der Diskriminierung durch solche Systeme oft unterschätzt und nicht ausreichend beachtet wird.

Rechtliche Grundlagen und bisherige Beschlüsse

Die Allgemeine Erklärung der Menschenrechte (AEMR) von 1948 bietet eine erste Grundlage, um sich mit der rechtliche Lage zum Thema Diskriminierung zu beschäftigen. Artikel 2 der AEMR beinhaltet ein akzessorisches Diskriminierungsverbot. Das bedeutet, dass es die AEMR untersagt, Menschen bezüglich der in der AEMR festgeschriebenen Rechte aufgrund geschützter Merkmale unterschiedlich zu behandeln. In den UN-Leitprinzipien für Wirtschaft und Menschenrechte, die 2011 vom MRR verabschiedet wurden, ist festgehalten, dass die Verantwortung von Staaten, vor Menschenrechtsverletzungen zu schützen, sich auch auf Menschenrechtsverletzungen von Unternehmen innerhalb des Staates beziehen. Die Staaten sind verpflichtet, effektive Schritte zur Überwachung und Verhinderung dieser zu unternehmen.

Black Box: Ein System, dessen innere Funktionsweise und Entscheidungsfindung schwer oder gar nicht nachvollziehbar ist.

Ein **Proxy** ist eine Variable, die einen starken Zusammenhang zu einer anderen Variablen hat. In der Informatik wird dies gezielt verwendet, um auf schwer messbare Eigenschaften zu schließen. Ungewollte starke Zusammenhänge zwischen Variablen können aber Probleme verursachen.



Ein allgemeines Diskriminierungsverbot wurde in Artikel 26 des **Internationalen Pakts über bürgerliche und politische Rechte** 1966 von der Generalversammlung verabschiedet. Außerdem existieren viele Spezialabkommen, in denen die Verpflichtungen der Staaten, Diskriminierung zu bekämpfen, konkretisiert werden. Artikel 6 der UN-Rassendiskriminierungskonvention beinhaltet beispielsweise ein Recht auf Schutz, Rechtsmittel und Entschädigung bei Diskriminierung. Das bezieht sich auch auf Fälle, in denen Diskriminierung ein Ergebnis der Nutzung künstlicher Intelligenz ist, und zwar unabhängig davon, ob diese Diskriminierung beabsichtigt ist. In vielen Fällen ist es Personen bisher jedoch nicht möglich, dieses Recht in Anspruch zu nehmen, somit steht die Staatengemeinschaft vor der Aufgabe, die genannten Rechte effektiv durchzusetzen.

Bisherige Maßnahmen

Die Zielsetzung ist also, Diskriminierung durch maschinelles Lernen so effektiv wie möglich zu verhindern. Dazu gibt es verschiedene Handlungsebenen und Wege des Umgangs. Einerseits gibt es technologische Möglichkeiten, die Verstärkung von Diskriminierung durch maschinelles Lernen etwa durch **'data blending'** zu verringern. Trainingsdaten sollten außerdem darauf untersucht werden, ob sie soziologisch oder historisch voreingenommen sind und das spätere Anwendungsgebiet abdecken. Abschließend sollten Systeme auf diskriminierende Ergebnisse getestet werden.

Technische Ansätze alleine reichen aber nicht, um die vorhandenen Probleme anzugehen. Eine große Rolle spielen strukturelle Aspekte. Unternehmen sollten dafür sorgen, dass größtmögliche Transparenz herrscht.

Da das ML-System selbst eine Black Box ist, gilt das für Trainingsdaten, Parameter, Nutzung und Zielsetzung des ML-Systems. Des Weiteren sollten Verantwortlichkeiten innerhalb von Arbeitsgruppen klar festgelegt sein.

Um Zugänglichkeit zur Nutzung zu gewährleisten, sollte darauf geachtet werden, dass auch marginalisierte Gruppen die Möglichkeit zu Zugang und Interaktion mit den Systemen haben.

Staaten können durch verschiedene Maßnahmen einen diskriminierungsfreien Einsatz von ML-Systemen fördern. Ziel ist hierbei die Si-

Internationaler Pakt über bürgerliche und politische Rechte: Ein völkerrechtliches Abkommen, das 1966 von der Generalversammlung der Vereinten Nationen verabschiedet wurde. Der IPB-PR sichert Grundrechte wie Meinungsfreiheit, Recht auf Leben und Gleichberechtigung.



Herstellung eines ethischen Umgangs in allen Phasen des Lebenszyklus eines ML-Systems. Die UNESCO empfiehlt hierfür den Einsatz von möglichst divers besetzten Expert*innengruppen mit klaren Verantwortlichkeiten und Kompetenzen. Diese sollen nationale ML/KI-Strategien aufstellen, die immer wieder auf Effektivität geprüft werden sollten und über einen detaillierten Handlungsplan verfügen.

Staaten haben einerseits die Möglichkeit, mit Richtlinien Umgangsweisen vorzugeben. Aber auch explizite Verbote und Pflichten von Unternehmen und weiteren Akteur*innen können festgelegt werden. So können Unternehmen etwa Gutachten mit konkreten Kriterien zu diskriminierenden Tendenzen von ML-Systemen vorgeschrieben werden. Des Weiteren können Unternehmen zu Transparenz verpflichtet werden, was etwa die Verantwortlichkeiten oder die Trainings- und Nutzungsdaten von ML-Systemen angeht.

Auch Bildungsinitiativen, die einen bewussten Umgang von Nutzer*innen und Entwickler*innen mit der Problematik erzeugen, sind ein effektiver Ansatz.

Internationaler Pakt über bürgerliche und politische Rechte: Ein völkerrechtliches Abkommen, das 1966 von der Generalversammlung der Vereinten Nationen verabschiedet wurde. Der IPB-PR sichert Grundrechte wie Meinungsfreiheit, Recht auf Leben und Gleichberechtigung.

Der **Business Council for Ethics of AI** ist ein Expertengremium, das sich mit ethischen Fragestellungen im Bereich Künstliche Intelligenz und maschinelles Lernen auseinandersetzt. Ziel ist die Erarbeitung ethischer Leitlinien und Standards, die Unternehmen und Staaten helfen sollen, KI-Anwendungen verantwortungsbewusst und transparent zu entwickeln und anzuwenden.

Die Vereinten Nationen agieren im Kontext der gesamten Thematik als ein Raum, in dem Werte festgelegt, Ziele gesetzt, Diskussionen gestartet und Expert*innengruppen gebildet werden. Die UNESCO hat in verschiedenen Veröffentlichungen Prinzipien und Leitlinien für Unternehmen und Staaten festgelegt, sowie Prinzipien zur Nutzung von ML-Systemen durch die Organe der UN festgeschrieben. Sie hat Plattformen und Expert*innennetzwerke wie den **Business Council for Ethics of AI**, das Global AI Ethics and Governance Observatory und eine High Level Expert Group ins Leben gerufen. Der Menschenrechtsrat hat mehrere Resolutionen zum Thema "Neue und aufstrebende digitale Technologien und Menschenrechte" beschlossen, zuletzt Resolution 53/29 vom Juli 2023. Diese Resolution fordert einen menschenrechtsbasierten Ansatz für den Umgang mit digitalen Technologien wie Künstlicher Intelligenz, um Risiken wie Diskriminierung und Datenschutzverletzungen zu verhindern. Sie betont die Notwendigkeit stärkerer Regulierungen, internationaler Zusammenarbeit und digitaler Bildung, um Menschenrechte zu schützen und nachhaltige Entwicklung zu fördern. Zusätzlich beauftragte der MRR mehrfach Berichte von Expert*innen zu dem Thema. Auch das Ge-



neralsekretariat beschäftigt sich mit dem Thema, beispielsweise in der **Roadmap for Digital Cooperation**. Weitere Projekte der Vereinten Nationen zu dem Thema sind der **AI for Good Summit** und die **Plattform Globalpolicy.AI**.

Insgesamt wird deutlich, dass es zahlreiche Ansatzpunkte gibt und bereits vielfältige Projekte umgesetzt wurden. Dennoch besteht erhebliches Verbesserungspotenzial, und die Weltgemeinschaft steht erst am Anfang, die Risiken von maschinellem Lernen als Verstärker von Diskriminierung anzugehen. Um wirkungsvoll voranzukommen, ist eine koordinierte Zusammenarbeit auf allen Ebenen erforderlich. Dabei sollten konkrete Strategien erarbeitet werden, die alle Bevölkerungsgruppen in allen durch ML beeinflussten Bereichen mit einbeziehen, klare Verantwortlichkeiten festlegen und Richtlinien vorgeben, sodass ein verantwortungsbewusster Umgang mit ML und ein effektives Vorgehen gegen Diskriminierung ermöglicht werden können.

Roadmap for Digital Cooperation:

Initiative des Generalsekretariats der Vereinten Nationen, die den globalen Dialog und die Zusammenarbeit im Bereich der Digitalisierung und Künstlichen Intelligenz fördern soll.

AI for Good Summit:

Eine internationale Konferenz, die von der Internationalen Fernmeldeunion (ITU) der Vereinten Nationen veranstaltet wird. Ziel des Summits ist es, den positiven Einsatz von Künstlicher Intelligenz zur Lösung globaler Herausforderungen wie Gesundheit, Klimawandel und Bildung zu fördern.

Globalpolicy.AI:

eine Plattform der Vereinten Nationen, die Ressourcen, Forschung und politische Richtlinien zu künstlicher Intelligenz und digitaler Transformation weltweit bereitstellt.

Hinweise zur Recherche

Sie sollten darauf gefasst sein, dass Sie bei der Recherche auf Texte und auch Bilder stoßen könnten, die Diskriminierung oder Waffen enthalten. Melden Sie sich bei Gesprächsbedarf gerne bei den Vertrauenspersonen für Teilnehmende.

vertrauenspersonen@mun-sh.de

Sollten Sie nun inhaltliche Fragen zum Text haben, können Sie sich gerne per Mail unter j.luehrs@mun-sh.de an Johanna Lührs wenden. Jetzt ist es an der Zeit, dass Sie sich selbst über das Thema und die Position ihres Landes informieren. Damit Sie die Problematik einordnen können, sollten Sie zumindest oberflächlich verstanden haben, was ML ausmacht. In den hilfreichen Quellen finden Sie Möglichkeiten, sich darüber weiter zu informieren.

Ein allgemeiner Hinweis: Lassen Sie sich nicht dadurch verunsichern, dass sehr oft von KI und nicht von ML die Rede ist. KI ist schlicht ein Unterbereich des ML, die Informationen sind also einfach übertragbar.

Um sich über die bisherigen Entscheidungen des MRR zu informie-



ren, können Sie Resolution A/HRC/53/L.27/Rev.1. zur Hand ziehen, hierbei handelt es sich um die aktuellste Resolution des MRR zu dem Thema. Hier können Sie auch direkt ein Gefühl dafür bekommen, wie Resolutionen aufgebaut sind. Leider ist diese Resolution nur auf Englisch verfügbar, Tools wie DeepL können bei der Übersetzung helfen.

Für die Recherche zur Position Ihres eigenen Landes empfehlen sich beispielsweise Mitschriebe von Debatten zu einem vergleichbaren Thema, Zeitungsartikel, Pressemitteilungen und Ähnliches. Sie können sich auch darüber informieren, wie Ihr Land allgemein mit Diskriminierung umgeht und ob Firmen, die viel Geld mit ML verdienen, in Ihrem Land ansässig sind.

Falls Sie wirklich tief in das Thema einsteigen wollen, lohnt sich außerdem ein Blick in das in den Quellen verlinkte Dokument "Diskriminierungsrisiken durch Verwendung von Algorithmen". Das ist natürlich viel zu lang, um es komplett zu lesen, aber schauen Sie gerne im Inhaltsverzeichnis, ob ein spezifisches Thema Sie genauer interessiert.

Quellen

Besonders hilfreiche Quellen

Video: Umfassende und anschauliche Erklärung dazu, was es mit maschinellem Lernen auf sich hat (deutsch)

Brown, Sara: Machine learning, explained. April 2022. (Hier ebenfalls eine Erklärung zu ML für die, die lieber lesen) (englisch).

Humanrights.ch: Rechtsschutz gegen Diskriminierung: UNO-Ebene. April 2020. (Übersicht über Grundlagen, die die Menschenrechte und andere internationale Verträge zum Thema Diskriminierung bieten) (deutsch)

Länderprofile mit Analysen zur Lage einiger weniger Länder (englisch)

Weitere Quellen

Orwat, Carsten: Diskriminierungsrisiken durch Verwendung von Algorithmen. September 2019. Diskriminierungsrisiken durch Verwendung Diskriminierungsrisiken durch Verwendung von Algorithmen. (Ausführlicher Bericht über Begriffsdefinitionen, Ursachen von Diskriminierungsrisiken, Beispiele und Handlungsoptionen) (deutsch).

Ashwini K.P.: Contemporary forms of racism, racial discrimination, xenophobia and related intolerance. Juni 2024. (Bericht über Diskriminierung und KI mit Empfehlungen an den MRR) (englisch).

Chen, Zhisheng, Ethics and discrimination in artificial intelligence-enabled recruitment practices. September 2023. (Studie zu Diskriminierung durch KI bei der Einstellung) (englisch).

High-Level Committee on Programmes: Principles for the Ethical Use of Artificial Intelligence in the United Nations System.

September 2022. (Prinzipien der UNESCO zur Nutzung von KI) (englisch).

UNESCO: Readiness assessment methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence. 2023. (Hilfe für Staaten, um einzuschätzen, wie gut sie dazu in der Lage sind, einen ethischen Umgang mit KI zu gewährleisten) (englisch)

UNESCO: Ethical impact assessment: a tool of the Recommendation on the Ethics of Artificial Intelligence. 2023. Ethical impact assessment: a tool of the Recommendation on the Ethics of Artificial Intelligence (Hinweise an Regierungsmitarbeiter, wie sichergestellt werden kann, dass KI den ethischen Standards der UNESCO entspricht) (englisch)

UNESCO: Recommendation on the Ethics of Artificial Intelligence. 2022. (Hinweise der UNESCO für einen ethischen Umgang mit KI) (englisch)

Generalversammlung: Internationaler Pakt über bürgerliche und politische Rechte. Resolution 2200A. 1966. (Hier befindet sich das Diskriminierungsverbot) (deutsch)

Generalversammlung: Universal Declaration of Human Rights. Resolution 217A. 1948. (Menschenrechte) (englisch)

Sekretariat des Hohen Kommissars der Vereinten Nationen für Menschenrechte: Human rights and technical standard-setting processes for new and emerging digital technologies: Report of the Office of the United Nations High Commissioner for Human Rights. Resolution A/HRC/53/42.

Generalsekretariat: Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation. A/74/821. Mai

2020. (Ziele des Generalsekretärs für die Nutzung digitaler Technologien) (englisch)

Menschenrechtsrat: New and emerging digital technologies and human rights. Resolution A/HRC/53/L.27/Rev.1. Juli 2023. (aktuellste Resolution zu digitalen Technologien und Menschenrechten)

Sekretariat des Hohen Kommissars der Vereinten Nationen für Menschenrechte: Guiding Principles on Business and Human Rights. 2011. (Welche Verantwortung tragen Staaten für Unternehmen?)

Menschenrechtsrat
Gremientext für das Thema
Maschinelles Lernen als Diskriminierungsverstärker
Autor*in: Johanna Lührs

Model United Nations Schleswig-Holstein 2025
mun-sh.de

Projektleitung MUN-SH 2025
Tom Bergmann – Carl-Jobst Hülsmann – Frederik Schissler
projektleitung@mun-sh.de

Ein Projekt des Deutsche Model United Nations (DMUN) e.V.
Birkenweg 1, 24235 Laboe

