«Es geht nicht darum, Technologie zu verteufeln»

Angela Müller ist Geschäftsleiterin der zivilgesellschaftlichen Organisation AlgorithmWatch CH, die sich mit den Auswirkungen von Algorithmen auf Menschenrechte, Demokratie und Nachhaltigkeit befasst. Im Interview plädiert sie für weniger Symptombekämpfung und mehr Gestaltungswillen hinsichtlich der technologischen Entwicklung.

Interview: Isabel Knobel

Angela Müller, was ist ein Algorithmus und wo wird er eingesetzt?
Ein Algorithmus ist eine Handlungsanleitung, wie ein Kochrezept: Wenn das Wasser kocht, dann gebe ich die Spaghetti hinein. Ein Algorithmus braucht viele Zutaten, in diesem Beispiel Spaghetti, Wasser und Salz. Diese Zutaten respektive Daten werden erhoben und mittels des Algorithmus analysiert, verarbeitet und interpretiert. Am Ende wird dann ein Output, ein Resultat, ausgespuckt. Das sind dann eben keine gekochten Spaghetti wie im Beispiel, sondern Inhalte wie ein Text, ein Bild, eine Prognose, eine Empfehlung oder eine Entscheidung. Algorithmische Systeme begegnen uns überall da, wo Automatisierung stattfindet. Sie zeigen uns beispielsweise Inhalte oder Werbungen an, die zu unseren Profilen passen. Sie können Job-Bewerbungen vorsortieren, unsere Leistung am Arbeitsplatz messen oder unsere Kreditwürdigkeit prüfen. In der Polizeiarbeit werden sie genutzt, um Verbrechen wie Wohnungseinbrüche oder Gefährder*innen im Bereich der häuslichen Gewalt vorherzusagen. Es gibt zahlreiche weitere Einsatzbereiche und Anwendungsfälle, manchmal sehr simple, manchmal komplexere. Diese Systeme prägen schon heute unseren Alltag.

Sie haben verschiedene Arten von Outputs aus algorithmischen Verfahren erwähnt. Worin unterscheiden sie sich?
Es gibt algorithmische Systeme, die Prognosen machen, Empfehlungen abgeben oder Entscheidungen treffen, manchmal vollautomatisiert, oft teilautomatisiert. Sie sind prädikativ, das heisst, sie sagen auf Basis der Daten etwas vorher. Dann gibt es auch sogenannte generative Systeme, die Inhalte wie Texte, Bilder, Videos oder Stimmen erzeugen. Sie beruhen auf Wahrscheinlichkeitsrechnung. Wenn beispielsweise ein grosses Sprachmodell einen Satz generiert, wählt es als nächstes Wort immer das Wort, das am wahrscheinlichsten auf das letzte Wort folgt. Das hat nichts mit Copy-Paste aus einer vorhandenen Quelle und auch nichts mit der Wahrheit zu tun. AlgorithmWatch CH führte beispielsweise vor den Eidgenössischen Wahlen 2023 eine grosse Recherche dazu durch, wie KI-Chatbots auf Suchmaschinen zum Teil komplett falsche, irreführende Antworten gaben.[1] Das kann die öffentliche Meinungsbildung in einer Demokratie durchaus beeinflussen.

Foto: David Bächtold, CC BY 4.0

Angela Müller

Sie sprechen die Gefahren im Zusammenhang mit Algorithmen an. Welche sind aus Ihrer Sicht die grössten?
Es ist oft eine Blackbox, wie und wo algorithmische Systeme eingesetzt werden. Gleichzeitig haben wir zahlreiche Beweise, dass sie vielfältige Auswirkungen auf Menschen und ihre Grundrechte, auf Meinungsbildung und demokratische Prozesse haben. Da müssen wir genau hinschauen. Wenn wir Entscheidungen automatisieren, müssen wir weiterhin Verantwortung zuweisen können. Algorithmen sind nie neutral und objektiv. Sie werden immer von Menschen gemacht, aus einem bestimmten Interesse und zu einem bestimmten Zweck. Dabei werden zahlreiche Entscheidungen getroffen, etwa welche Daten als Grundlage dienen, welche Kriterien als relevant erachtet und wie sie programmiert werden. All diese Entscheidungen sind Quellen von möglichen Ungerechtigkeiten und letztlich Diskriminierung. Nehmen wir den Status der Frauen als Beispiel: Werden historische Daten als Grundlage verwendet, bilden sie die Vergangenheit und die dann herrschenden gesellschaftlichen Strukturen und Machtverhältnisse ab. Das System erkennt in diesen Daten bestimmte Muster und übernimmt sie

«Nicht alle Menschen werden im Internet gleich abgebildet: Es ist geprägt von gewissen Stimmen und Perspektiven auf die Welt.»

«De facto treffen auch Systeme Entscheidungen, die eigentlich nur eine Prognose machen sollten.»

in seinen Prognosen, die eigentlich für die Zukunft sind. Geht man davon aus, dass Frauen früher eher nicht berufstätig oder beruflich weniger erfolgreich waren und das System dies als Fakt annimmt, dann kann sich das reproduzieren. Gerade im Fall von Frauen passiert das sehr oft.

Kennen Sie konkrete Fälle, wo Frauen von algorithmischen Systemen diskriminiert wurden?

Bekannte und eindrückliche Beispiele kommen aus der Arbeitswelt. Bei verschiedenen Rekrutierungssystemen zeigte sich, dass sie eher Männer als Frauen empfehlen, unter anderem weil die bisherige Belegschaft männlicher war.[2] Selbst wenn das Geschlecht nicht explizit im Lebenslauf steht, können Algorithmen über sogenannte Proxy-Kriterien Korrelationen ableiten, z. B. über Hobbys, die eher mit dem weiblichen Geschlecht assoziiert werden. Auch beim Einsatz von Systemen, die Sozialhilfebetrug aufdecken sollen, kommt es häufig zu Diskriminierungen. Bei einem Fall in Rotterdam wurden alleinerziehende Mütter als Hochrisikogruppe für Sozialhilfebetrug klassifiziert.[3] Und nicht zuletzt sind Social-Media-Plattformen anfällig für Diskriminierungen. In einem Experiment fanden wir heraus, dass bei Stellenanzeigen auf Facebook Gender-Stereotypen verwendet werden. So wird eine Anzeige für Lastwagenfahrerinnen und Lastwagenfahrer eher Männern angezeigt, während die Stelle für Kinderbetreuende eher Frauen angezeigt wird.[4] Diskriminierungen durch algorithmische Systeme sind zudem oft intersektional. In der Gesichtserkennung werden Systeme viel öfter mit Gesichtern von weissen Männern trainiert, was zur Folge hat, dass weibliche Gesichter und People of Color schlechter erkannt werden.

Ist also die mangelnde Datengrundlage das grösste Problem?

Diskriminierung durch Algorithmen kann viele verschiedene Ursachen haben. Die Datengrundlage ist eine davon. Daten für algorithmische Systeme kommen oft aus dem Internet. Dieses ist aber keine Blumenwiese. Nicht alle Menschen werden im Internet gleich abgebildet: Es ist geprägt von gewissen Stimmen und Perspektiven auf die Welt. Andere Stimmen und Diskurse, die vielleicht auch wichtig wären, finden dagegen nie Eingang in die Datengrundlage und so sind die Daten auch nicht frei von Ungerechtigkeiten, Stereotypen und Verschwörungstheorien. Neben der Datengrundlage ist wichtig, wie algorithmische Systeme entwickelt und eingesetzt werden. Wir sehen, dass die Nutzenden – sei es nun die Polizistin, die eine Gefährdungsmeldung macht, oder der Recruiter, der eine Bewerbung anschaut – die Tendenz haben, der Prognose oder der Empfehlung des Systems zu folgen. Wenn dieses grün oder rot anzeigt, werden die meisten Menschen entsprechend entscheiden. Dieses übermässige Vertrauen in die Automatisierung nennt man *automation bias.* De facto treffen also auch Systeme Entscheidungen, die eigentlich nur eine Prognose machen sollten. Das kann auch eine Quelle der Diskriminierung sein, die weniger mit der Technik als mit den Menschen zu tun hat, die sie benutzen.

Ein Mensch ist unter Umständen auch biased, kann blinde Flecken haben, voreingenommen sein oder sogar diskriminierend handeln. Treffen Algorithmen im Vergleich vielleicht nicht doch bessere, verlässlichere Entscheidungen?

Es geht nicht darum, zu behaupten, dass der Mensch nie diskriminiert und das algorithmische System immer – beides stimmt nicht. Es gibt jedoch wichtige Unterschiede zwischen Menschen und Maschinen. Der erste ist das Ausmass: Nehmen wir das Beispiel einer Richterin im Gerichtssaal. Sie kann auch voreingenommen und nicht immer konsistent sein, aber sie wird nur einen oder wenige Entscheide pro Tag treffen. Ein algorithmisches System kann dagegen in kurzer Zeit sehr viele Entscheide treffen. Der zweite Unterschied ist die Nachvollziehbarkeit von Verantwortlichkeit: Wenn die Richterin einen Entscheid trifft, wird das in einem Rechtsstaat klar dokumentiert. Als verurteilte Person hat man Rechtsmittel zur Verfügung und kann Entscheide anfechten. Bei einem Blackbox-Algorithmus hingegen ist es sehr schwierig, sich zu wehren. Oft weiss man gar nicht, dass ein Entscheid algorithmisch gefällt wurde oder wie das System zu seinem Entscheid kam. Drittens verkennt der Vergleich von Mensch und Maschine, wovon wir reden: In einem Gerichtssaal darf man als angeklagte Person seine Position darlegen und kann einem Menschen gegenüber argumentieren. Wenn man dagegen nur die fertige Prognose eines Computers sieht, hat man nicht die Möglichkeit, zu argumentieren und für seine Rechte einzustehen. Das ist ebenfalls ein wesentlicher Aspekt unseres Rechtsstaats. Es geht also nicht nur darum, ob der Output eines Systems besser oder schlechter ist als der Output eines Menschen. Es geht immer auch um den Prozess dahinter, der legitim sein muss.

Sollten wir den Einsatz von algorithmischen Systemen also besser stoppen?
Nein, es geht nicht darum, Technologie zu verteufeln oder gar zu verbieten. Algorithmische Systeme haben durchaus das Potenzial, unsere Gesellschaft besser und gerechter zu machen. Dafür muss Gerechtigkeit als explizites Ziel definiert werden, das man mit dem System verfolgt – im Gegensatz zu Systemen, die man rein zur Effizienzsteigerung einsetzt. Ein schönes Beispiel dafür ist ein Algorithmus für die Einteilung von Schülerinnen und Schülern in Klassen. Studien zeigen, dass soziale Durchmischung die Chancengerechtigkeit erhöht, und genau das soll mit diesem System erreicht werden.[5] Algorithmen sind allgemein gut darin, Muster zu erkennen und können so auch Diskriminierungsmuster erkennen. Wichtig ist aber, dass wir verstehen, dass die Technologie ein Werkzeug der Menschen ist und nicht plötzlich von allein die grossen Probleme löst und alles gerechter macht. Sie steht oft stellvertretend für gesellschaftliche Machtverhältnisse.

Wie können wir denn gut mit Technologie umgehen?
Ich finde es zentral, dass wir nicht nur Symptombekämpfung betreiben und uns darauf beschränken, uns gegen Intransparenz und Diskriminierung zu wehren. Wir sollten ambitionierter denken und uns zum Ziel setzen, Technologie so zu entwickeln und einzusetzen, dass sie von Beginn weg allen zugutekommt. Dazu müssen wir als Gesellschaft die technologische Entwicklung mitgestalten, statt uns von ihr gestalten zu lassen. Es ist wichtig, die Genderperspektive und andere Blickwinkel in die Ausgestaltung von Algorithmen einfliessen zu lassen, Betroffene einzubeziehen und Folgenabschätzungen zu machen, welche Auswirkungen ein System auf grundlegende Rechte hat. Dabei ist zentral, dass eine gesellschaftliche Debatte geführt wird. Ich höre oft von Menschen, die das lähmende Gefühl haben, dass sie von der technologischen Entwicklung wie von einer Naturgewalt überrollt werden. Dem müssen wir uns aber nicht fügen. Wir sollten anfangen, den Spiess umzudrehen und die Bedingungen für die Entwicklung und den Einsatz der Technologie zu formulieren, mit dem Ziel, dass sie der Gesellschaft zugutekommt.

Kann eine Gesellschaft das leisten?
Eine breite gesellschaftliche Debatte ist sehr wichtig für Veränderung. Wenn wir sie nicht führen, haben wir auch nicht das Gefühl, etwas bewegen zu können. Natürlich ist auch die Politik in der Pflicht, unsere Grundrechte und unsere Demokratie zu schützen. Das Thema hat in den letzten Jahren auf der politischen Agenda stark an Gewicht gewonnen. Der Bundesrat publiziert Ende 2024 eine Auslegeordnung zur KI-Regulierung. Leider steht er bisher nicht gerade auf dem Gaspedal, sondern wartet eher ab und beobachtet. Da dürften wir schon mutiger und entschlossener vorgehen.

Wie können Politikerinnen und Politiker Chancengerechtigkeit im Kontext von Algorithmen fördern?
Es gibt diverse Möglichkeiten, den Diskriminierungsschutz in der Schweiz zu verbessern, gerade auch im Privatsektor, in dem es bis dato nur sehr wenig Schutz vor Diskriminierung gibt. Wichtig ist zudem, dass Menschen sich einfacher gegen Diskriminierungen wehren können. Oft ist es schwierig, individuelle Diskriminierungen durch Algorithmen festzustellen, weil man meist gar nicht weiss, dass man von einem Algorithmus bewertet wurde oder wie dies genau geschehen ist. Deshalb wäre es sinnvoll, wenn man sich als Kollektiv zur Wehr setzen oder von einer Organisation vertreten lassen könnte, d.h. man müsste die Möglichkeit einer Verbandsklage einführen. Zudem könnten sich Politikerinnen und Politiker vermehrt für Forschungsförderung und speziell für interdisziplinäre Forschung einsetzen, die eine rein technische Forschung ergänzt. Nicht zuletzt sollten wir die ganzen globalen Wertschöpfungsketten hinter der Künstlichen Intelligenz durchleuchten und ihre Nachhaltigkeit kritisch beleuchten. Daten für die Entwicklung von algorithmischen Systemen werden oft im globalen Süden, in sogenannten *digital sweatshops,* von Menschen mit sehr schlechten Arbeitsbedingungen aufbereitet. Dazu kommen massive Umweltauswirkungen und eine extreme Machtkonzentration von grossen Technologiefirmen, die Märkte mit ihren Quasi-Monopolen steuern.

Es scheint noch viel zu tun zu geben. Was kann AlgorithmWatch CH beitragen?
Wir wollen Wissen bereitstellen und Licht ins Dunkel bringen, einerseits durch Recherchen und Untersuchungen zu den Auswirkungen von algorithmischen Systemen, andererseits durch die Analyse der gesetzlichen Rahmenbedingungen. Auf Basis der Ergebnisse machen wir Advocacy-Arbeit, d.h. wir entwickeln Empfehlungen und Vorschläge für die Politik. So haben wir etwa ein Positionspapier verfasst, das konkrete Empfehlungen zum Schutz

«Gerechtigkeit muss als explizites Ziel definiert werden.»

«Es braucht eine breite gesellschaftliche Debatte, an der alle teilnehmen können.»

vor diskriminierenden Algorithmen formuliert.[6] Gleichzeitig machen wir auch Öffentlichkeitsarbeit, weil wir überzeugt sind, dass es eine breite gesellschaftliche Debatte braucht, an der alle teilnehmen können. Dieses Jahr haben wir eine grosse Kampagne zu Diskriminierung durch Algorithmen durchgeführt und im Sommer einen Appell an den Bundesrat lanciert, den Schutz vor Diskriminierung durch Algorithmen und KI zu einer Priorität der kommenden Regulierungen rund um KI zu machen.[7] Mit diesem breit abgestützten Signal wollten wir zeigen, dass es um Fragen der Gerechtigkeit geht, für die wir als Gesellschaft Verantwortung übernehmen können und müssen.

Isabel Knobel ist Politologin und sucht als Projektleiterin beim Migros-Pionierfonds mutige Ideen für eine zukunftsfähige Gesellschaft. Sie schreibt regelmässig für die Fachzeitschrift «Frauenfragen».

Anmerkungen

1 AlgorithmWatch CH: Wie ChatGPT, Bing Chat und Co. die öffentliche Meinungsbildung gefährden könnten, 04.10.2023. https://algorithmwatch.ch/de/bing-chat-wahlen-2023 (abgerufen am: 04.07.2024).

2 AlgorithmWatch CH: Algorithmenbasierte Diskriminierung in Job-Auswahlverfahren vermeiden, 16.11.2024. https://algorithmwatch.ch/de/findhr (abgerufen am: 04.07.2024).

3 European Parliament, E-000780/2023: Rotterdam 'fraud prediction' algorithms automating injustice: Dutch Government violating fundamental rights and the rule of law, [Parlamentarische Frage] 07.03.2023. https://www.europarl.europa.eu/doceo/document/E-9-2023-000780_EN.html#def1 (abgerufen am: 04.07.2024).

4 Kayser-Bril, Nicolas: Automatisierte Diskriminierung: Facebook verwendet grobe Stereotypen, um die Anzeigenschaltung zu optimieren. AlgorithmWatchCH, 18.09.2024. https://algorithmwatch.org/de/automatisierte-diskriminierung-facebook-verwendet-grobe-stereotypen-um-die-anzeigenschaltung-zu-optimieren (abgerufen am: 04.07.2024).

5 Hufschmid, Samuel: Für die soziale Durchmischung teilt in Uster ein Computer Klassen ein. AlgorithmWatchCH, 03.07.2023. https://algorithmwatch.ch/de/uster-computer-teilt-klassen-ein (abgerufen am: 04.07.2024).

6 AlgorithmWatch CH: Diskriminierende Algorithmen: So gelingt der Schutz, 19.09.2023. https://algorithmwatch.ch/de/diskriminierende-algorithmen (angerufen am: 04.07.2024).

7 AlgorithmWatch CH: Künstliche Intelligenz mit Verantwortung … ohne Diskriminierung, 06.2024. https://algorithmwatch.ch/de/ki-ohne-diskriminierung (abgerufen am: 04.07.2024).

Abstracts

«Il ne s'agit pas de diaboliser la technologie»

Angela Müller dirige l'organisation de la société civile AlgorithmWatch CH qui s'occupe de l'impact des algorithmes sur les droits humains et sur la société. Dans son interview, elle présente des exemples montrant que la discrimination des femmes dans les décisions algorithmiques est imputable pour une part à la technique, mais aussi systématiquement à la manière dont les systèmes ont été développés et sont utilisés. Elle préconise qu'au lieu de lutter contre les symptômes nous prenions l'initiative, en tant que société, d'imposer des conditions au développement technologique afin qu'il bénéficie à l'ensemble de la population.

«Non si tratta di demonizzare la tecnologia»

Angela Müller è la direttrice esecutiva dell'organizzazione della società civile AlgorithmWatch CH che si occupa dell'impatto degli algoritmi sui diritti umani e sulla società. Nell'intervista mostra con esempi come la discriminazione contro le donne attraverso decisioni algoritmiche non sia mai riconducibile solo alla tecnologia in quanto tale, ma anche al modo in cui viene sviluppata e utilizzata. Anziché combattere i sintomi, ci esorta come società a formulare proattivamente le condizioni dello sviluppo tecnologico affinché vada a beneficio di tutti.