

## THEMA

REGARDS

FORSCHERIN ZU ALGORITHMISCHER DISKRIMINIERUNG

# „In Wirklichkeit entscheiden Menschen, wie Algorithmen funktionieren“

Interview: Joël Adami

**Die Forscherin Yasaman Yousefi beschäftigt sich mit der Frage, wie algorithmische Systeme und künstliche Intelligenz gerechter gestaltet werden können. Nächste Woche ist sie in Luxemburg zu Gast, die woxx hat sich mit ihr über den KI-Hype unterhalten.**

**woxx:** Was ist algorithmische Diskriminierung?

**Yasaman Yousefi:** Das ist ein Phänomen, das dann passiert, wenn automatisierte Systeme Menschen aufgrund bestimmter Eigenschaften unfair behandeln. Algorithmen nehmen Korrelationen aufgrund dieser Eigenschaften an – das könnte zum Beispiel Ethnie, Geschlecht, sozioökonomischer Status oder ganz neue Kategorien wie „traurige Teenager“ oder „Hundehalter“ sein. Algorithmen tun genau das: Sie finden Gemeinsamkeiten und gruppieren Dinge. Diese Differenzierung, die an sich harmlos ist, kann diskriminierend werden, weil das Training der Algorithmen auf vorurteilsbehafteten Daten oder Praxen basiert. Das kann zu einem ungleichen Zugang zu Chancen, Diensten oder gar Informationen führen. Ein Beispiel ist der Fall von Amazons Einstellungsalgorithmus. Der war zwar nicht dafür entwickelt, aber er hatte Vorurteile gegenüber Frauen. Warum? Weil in die Trainingsdaten die Einstellungspraxis der letzten 20 Jahre war, und darin wurden mehr Männer als Frauen eingestellt. Also entschied der Algorithmus, dass CVs, auf denen „Frau“ stand, nicht gut genug waren und ignoriert werden konnten. So passiert algorithmische Diskriminierung, auch wenn sie nicht unbedingt beabsichtigt ist.

**Der Einfluss, den Algorithmen von sozialen Netzwerken und Shopping-Websites auf unser Leben haben, ist etwas subtiler. Wie funktioniert das?**

Die Algorithmen dieser Websites und Apps sollen „Engagement“ und Profit

optimieren, besonders beim Shopping. In sozialen Netzwerken werden Inhalte priorisiert, die starke Reaktionen hervorrufen – normalerweise also polarisierende Ansichten oder auch Falschinformationen. Im Fall von Onlineshopping beeinflussen Algorithmen unsere Wahl, indem sie Empfehlungen personalisieren. Durch unsere Daten können sie uns gezielte Werbung zeigen. Ein Beispiel: Ich bin Katzenbesitzerin. Mein ganzes Instagram, wie Sie sich vorstellen können, ist voll mit Katzeninhalten. Genau so funktioniert das. Manchmal hat man das Gefühl, die Algorithmen lesen unsere Gedanken, weil wir gerade über ein Produkt nachgedacht oder mit Freunden darüber geredet haben – und dann sieht man schon eine Werbung dafür. Aber in Wirklichkeit korrelieren sie unser Onlineverhalten mit anderen Faktoren wie Altersgruppe, Einkommen, Interessen und Ähnlichem. So wird unser Verhalten gesteuert, ohne dass es uns bewusst ist.

**Wir interagieren also – wissend oder nicht – täglich mit vielen algorithmischen Systemen. Vielen ist auch nicht bewusst, dass diese das Potenzial haben, soziale Ungleichheiten zu vertiefen. Wie funktioniert das?**

Das Problem sind nicht Algorithmen per se, sondern die Daten, mit denen sie trainiert werden. Diese Daten enthalten bereits viele Ungleichheiten. In den USA gab es den Fall eines Algorithmus, der in Gerichtsentscheidungen benutzt wurde, um die Wahrscheinlichkeit einer Wiederholung einer Straftat zu berechnen. Der Algorithmus bevorzugte weiße Menschen gegenüber schwarzen, auch wenn das begangene Verbrechen komplett gleich war. So empfahl es den Richtern, schwarze Menschen länger ins Gefängnis zu stecken. Das sind keine neuen Vorurteile, oder etwas, was die Algorithmen selbst kreiert hätten – die Vorurteile in den Daten werden verstärkt. Das kann überall zu Problemen

führen: Marginalisierte Menschen, Frauen, oder rassifizierte Menschen könnten schlechtere Angebote für Kredite oder Hypotheken erhalten, weniger Jobangebote und müssten sich strengeren Kontrollen am Flughafen unterziehen.

**Das Problem sind nicht Algorithmen per se, sondern die Daten, mit denen sie trainiert werden.**

**Es gibt den Trend, algorithmische Systeme zu vermenschlichen, und das nicht nur mit KI, sondern auch bei sozialen Netzwerken. Zum Beispiel wird gesagt „Instagram mag diese Art von Fotos, Instagram mag diese Inhalte nicht“. Verschleiert das, wer wirklich für die Aktionen dieser Algorithmen verantwortlich ist?**

Das ist ein guter Punkt. Wenn wir Algorithmen vermenschlichen, geben wir uns der Illusion hin, dass diese Plattformen einen eigenen Willen haben. In Wirklichkeit entscheiden Menschen, wie Algorithmen funktionieren – was wiederum die Ungleichheiten in unserer Gesellschaft spiegelt. Der Eindruck, dass es ein Algorithmus oder eine App ist, die etwas mag oder nicht mag, ist manipulativ. Es nimmt alle Menschen, die hinter der Technologie stehen, aus der Verantwortung: Entwickler, Ingenieure, Designer. Dieses Narrativ von „Siri sagt jenes“ oder „ChatGPT denkt dies“ ist etwas, wovor wir wirklich Angst haben sollten.

**Mit sogenannter KI ist diese Tendenz noch größer. Ein Chatbot wirkt, als habe er Intelligenz oder sogar Bewusstsein. Ist das ein cleverer Trick, um der Verantwortung zu entgehen?**

Absolut. Als ich das erste Mal über mein Forschungsthema nachgedacht

## Veranstaltungshinweis:

### Philo-Workshop mit Yasaman Yousefi und Nora Schleich

Am 29. April diskutieren Yasaman Yousefi und Nora Schleich ab 18:30 Uhr im Rahmen eines „Philo-Workshops“ in der Erwuessebildung (5, avenue Marie-Thérèse, Luxemburg-Stadt) über die Hintergründe und Konsequenzen algorithmischer Diskriminierungen. Es ist auch möglich, die Veranstaltung online zu verfolgen. Um Anmeldung via E-Mail an [info@ewb.lu](mailto:info@ewb.lu) wird gebeten.



FOTO: PRIVAT

Yasaman Yousefi forscht interdisziplinär zur Fairness in algorithmischer Entscheidungsfindung. Am 29. April ist sie in Luxemburg zu Gast.

habe, war ich noch Masterstudentin. Ich las einen Artikel darüber, dass Siris Stimme weiblich ist. Das hat bei mir die Frage ausgelöst, warum diese Stimme weiblich ist, und wer das entschieden hat. Siri soll deine Assistentin sein, und du bist der Chef. Also wurde Siri mit weiblicher Stimme entwickelt, um untergeben zu wirken. Das baut auf dem Stereotyp auf, dass Frauen für Assistenzrollen geeignet sind – und verstärkt diesen zugleich. Wenn also eine KI eine „Persönlichkeit“ bekommt, ist das ein Weg, um zu vermeiden, dass über die Designentscheidungen, die Datenquellen und die Machtstrukturen, die all diese Systeme formen, nachgedacht wird. Im Grunde handelt es sich um ein wunderschönes Hintertürchen, damit wir die Technologie und nicht die Unternehmenspolitik, die hinter ihrer Entwicklung steht, verantwortlich machen.

#### **Welche Risiken könnte das verbergen?**

Ich möchte betonen, dass ich nicht denke, dass KI nur Risiken mit sich bringt. Mit guter Bildung, ethischem Design und solider Gesetzesgrundlage könnte KI sehr nützlich für die Gesellschaft sein. Neben dem Festschreiben von Vorurteilen und Diskriminierung sehe ich die größten Risiken in der mangelnden Transparenz und dem Problem der „Black Boxes“. Ein Problem wäre auch, wenn es weniger menschliche Kontrolle gäbe. Wir könnten zu sehr auf Automatisierung und KI-Empfehlungen vertrauen, was Machtdynamiken unsichtbar machen könnte. Es könnte durch KI auch neue Formen der Diskriminierung geben, auf die wir nicht vorbereitet sind. Privatsphäre und Datenschutz sind weiterhin Felder, die trotz DSGVO und AI Act (beides EU Verordnungen, Anm.d.Red.) nicht gänzlich gelöst sind. Letzteres Gesetz ist zwar ein guter Anfang, aber die Legislation entwickelt sich immer langsamer als die Technologie.

#### **Sie haben diese „Black Boxes“ erwähnt. Denken Sie, Firmen könnten dazu gezwungen werden, diese zu öffnen und zu zeigen, wie genau ihre Algorithmen funktionieren?**

Ich denke, die „Black Boxes“ ganz zu öffnen, wäre zu viel verlangt, weil manchmal wissen die Ingenieure selbst nicht, wie ein System eine Entscheidung getroffen hat. Gesetze könnten für mehr Transparenz und Rechenschaftspflichten sorgen, und das tun sie auch bereits. Zum Beispiel sind im AI Act Kriterien für Erklärbarkeit, Transparenz und menschliche Kontrolle angegeben, die die „Black Box“ nicht ganz öffnen, aber die Gedankengänge dahinter aufzeigen. Wir sollten aber bedenken, dass ein System, je komplexer es wird, umso schwieriger zu öffnen sein wird. Zudem auch der menschliche Geist eine Art „Black Box“ ist: Wir können nicht in ein Gehirn hineinschauen und sehen, wie es genau funktioniert. Menschliche Entscheidungen sind deshalb auch nicht immer völlig erklärbar. Ich denke, sowohl mit Menschen und mit KI ist es wichtig, weiterhin Fragen zu stellen

und die Argumentation hinter einer Entscheidung zu hinterfragen.

#### **Im Grunde handelt es sich um ein wunderschönes Hintertürchen, damit wir die Technologie und nicht die Unternehmenspolitik dahinter verantwortlich machen.**

#### **Im Moment gibt es einen riesigen Hype um alles, was mit KI zu tun hat. Industrie und Regierungen geben Millionen aus, um Rechenzentren zu bauen, die sie „KI-Fabriken“ nennen. Denken Sie, es ist möglich, diese Technologien in einer ethischen Art und Weise einzusetzen?**

Mit ethischem Design, inklusiver Entwicklung, der Einbindung der Zivilgesellschaft und starker Regulierung könnten wir ethisch entwickelte Technologien haben. Viele Organisationen

versuchen dies zu erreichen. Ich arbeite zum Beispiel mit einem multidisziplinären Team bei einer Firma namens Dexai, um ethische KI-Systeme zu entwickeln. Ich denke, neben dem Fokus auf ethischer Entwicklung sollten wir uns alle eine wichtige Frage stellen: Ist KI in diesem Einsatzgebiet wirklich nötig? In manchen Fällen geht es wirklich ohne. Im Buch „From Pessimism to Promise“ (von Payal Arora, Anm. d. Red.) wird ein interessantes Beispiel beschrieben. In mehreren afrikanischen Ländern gibt es große Investitionen in KI-Systeme, um Wilderei zu bekämpfen. Diese Systeme alarmieren die Ranger, sodass diese eingreifen können. Aber manche Ranger sagen, dass das Geld besser in neues Equipment investiert wäre, dann könnten sie auch ohne KI ihre Arbeit tun. Wir sollten KI nur dann nutzen, solange sie nützlich ist, und nicht nur immer mehr davon fordern, weil damit Geld zu verdienen ist. Doch das ist leider das, was gerade passiert.

#### **Gibt es auch Lichtblicke?**

Ich denke, ein positiver Aspekt ist, wenn es einen gleichberechtigten Zugang zu ihr gibt, Technologie im Generellen wirklich die Rettung für manche Gemeinschaften sein könnte. Meine persönliche Perspektive ist: Ich bin wegen Technologie da, wo ich bin. Die Sprachen, die ich gelernt habe, habe ich durch meinen Zugang zu Technologie gelernt. Marginalisierte Gemeinschaften sollten einen gleichberechtigten Zugang bekommen, um den Nutzen von Technologien zu maximieren. Ich sehe einen Hoffnungsschimmer für Gemeinschaften, in denen Ungleichheiten viel höher sind als in Europa – sie können dem Rest der Welt ihre Sichtweisen mitteilen. Ich denke, wir sollten uns mehr auf die positiven als auf die negativen Seiten fokussieren. Gleichberechtigter Zugang zu und inklusives Design von Technologien könnten sogar einige menschliche Fehler ausgleichen, was zu einer besseren Zukunft führen könnte.

#### **Zur Person**

Dr. Yasaman Yousefi hat gemeinsam an den Universitäten Bologna und Luxemburg in Recht, Wissenschaft und Technologie promoviert. Sie hat kürzlich ihre Dissertation mit dem Titel „Quest for AI Fairness: Ethical, Legal, and Technical Solutions“ verteidigt. Ihre Forschung verfolgt einen interdisziplinären Ansatz zur Fairness in algorithmischer Entscheidungsfindung und verbindet ethische, rechtliche und technische Perspektiven. Yasaman arbeitet jetzt als Postdoktorandin an der Universität Bologna und beschäftigt sich mit den Risiken von KI-Systemen.