V29 Mehr KI, weniger Diskriminierung! Diskriminierungsfreiheit als Bestandteil der Berliner Digital-Strategie

Antragsteller*in:	Laura Sophie Dornheim (KV Berlin-Kreisfrei)
Tagesordnungspunkt:	TOP 8 Verschiedenes
Status:	Zurückgezogen

Schon heute entscheiden Algorithmen
An unzähligen Stellen treffen wir als Bürger*innen und Menschen auf bestimmte und vordefinierte Entscheidungsprozesse. Bei der Beantragung staatlicher Leistungen (z.B. Kitagutschein, ALG aber auch Asylgesuch) genau wie in der Privatwirtschaft (z.B. Jobbewerbung, Kreditvergabe, Handyvertrag). Diese Entscheidungen werden anhand spezifischer Verfahrensanweisungen getroffen, die derzeit meist noch von Menschen ausgeführt werden.

Diese Verfahrensanweisungen, oder auch Algorithmen, an einen Computer zu übertragen, könnte nicht nur die Bearbeitungszeit stark beschleunigen, sondern potentiell auch Diskriminierung und subjektive Entscheidungen eliminieren.

Selbstverständlich ist es dazu essentiell, den Algorithmus diskriminierungsfrei und so objektiv wie möglich zu programmieren.

Zukunftstechnologien diskriminierungsfrei gestalten

Wir Bündnis-Grüne sind stark im Einsatz gegen jede Form der Diskriminierung und Benachteiligung marginalisierter Menschen. Daher wollen wir auch Zukunftstechnologien wie algorithmen-basiertes Entscheiden (Algorithmic Decision Making, ADM) so gestalten, dass sie unsere Welt gerechter und inklusiver macht.

Diskriminierungsfreiheit als Bestandteil der Berliner Digital-Strategie

Wir fordern daher, dass im Rahmen der Berliner Digitalstrategie die Förderung diskriminierungsfreier ADM- und KI-Technologien explizit berücksichtigt wird.

Forschungsvorhaben und Start-Ups, die entsprechende Schwerpunkte haben, sollen besonders gefördert werden können.

Dafür soll ein Gremium aus Expert*innen Qualitätsstandards definieren, die ADM und KI auf ihre Diskriminierungsfreiheit hin bewertbar machen. Dabei soll ein Fokus auf die anzustrebenden Standards beim Einsatz von ADM und KI in der Verwaltung und anderen öffentlichen Stellen liegen.

Vorbild ist hier die Stadt New York, die mit der New York City Automated Decision Systems Task Force (ADS Task Force) per Gesetz eine eigene Institution geschaffen hat, die als eine Art “Algorithmen-TÜV” agiert. Die ADS Task Force überprüft alle in der Verwaltung eingesetzten Tools auf ihre Diskriminierungsfreiheit.

Solche Qualitätsstandards könnten sich beispielsweise auf die Entwicklung der Algorithmen, die Zusammensetzung des Entwicklungsteams, die verwendeten Trainingsdaten (insbesondere bei KI), die Überprüfbarkeit und Transparenz (bspw. Open Source) beziehen.

Die Sprecher*innen der LAG Digitales und Netzpolitik bieten gerne an, ihre Kompetenzen in diesen Prozess einzubringen.

Begründung

Begründung:

Jeder Algorithmus ist nur so gut, wie der Mensch, die ihn programmiert hat. “Der Fehler sitzt meistens vor dem Bildschirm” lernen Informatikstudierende auch heute noch.

Um dafür zu sorgen, dass Algorithmen wirklich objektiv entscheiden, müssen Entwickler*innen nicht nur gut programmieren können, sondern auch für existierende Vorurteile und Stereotypen sensibilisiert sein. Dies gilt für “einfache” Algorithmen und hat auf die Entwicklung von KI einen noch deutlich größeren Einfluss. Wenn die Datenmengen, mit denen eine KI trainiert wird, nicht-bereinigtes Abbild einer diskriminierenden Realität sind, wird mit sehr großer Wahrscheinlichkeit diese KI am Ende ebenfalls diskriminieren, möglicherweise sogar noch schlimmer als in den zugrundeliegenden Daten.

Bestes Beispiel dafür ist die gescheite Recruiting-KI von Amazon, die eingestellt wurde, nachdem sich herausstellte, dass sie systematisch Frauen benachteiligte. [1] Verantwortlich dafür ist aber nicht die KI, denn eine Software hat keinerlei eigene Intentionen. Vielmehr ist der Grund die seit Jahren bestehende diskriminierende Einstellungs- und Beförderungspolitik [2], deren Ergebnisse die Datengrundlage für die Recruiting-KI waren. Wenn 78% der Führungspositionen von Männern besetzt sind, dann schliesst eine KI daraus, dass Männer bessere Führungskräfte sind. Jedenfalls wenn niemand darauf achtet, dem Algorithmus beizubringen, dass erwiesenermaßen Frauen mindestens ebenso gut solche Positionen ausfüllen können.

Diskriminierungsfreie Technik ist leider noch nicht Standard, aber die Thematik bekommt immer größeren Stellenwert, etwas verzögert zum gesellschaftlichen Diskurs über Diskriminierung. So haben zwei japanischer Forscher*innen kürzlich einen Algorithmus vorgestellt, der es ermöglicht, die Diskriminierung einer KI zu limitieren. [3]

Sowohl einfache als auch ML-basierte Algorithmen können Fehlentscheidungen treffen oder diskriminierend entscheiden. In beiden Fällen liegt die Ursache bei den Menschen, die sie entwickelt haben.

Während bei einfachen Algorithmen noch leichter nachvollziehbar ist, welche Schritte ausgeführt werden, um zu einer Entscheidung zu kommen, ist dies bei ML-basierten KI-Algorithmen kaum möglich, solche Algorithmen sind hochkomplex und ohne tiefes technisches und mathematisches Verständnis kaum nachvollziehbar.

Doch für beide Arten von Algorithmen liessen sich, gerade für den Einsatz bei öffentlichen Stellen, Rahmenbedingungen definieren, die Diskriminierung und Fehlentscheidung weitgehend eliminieren. So könnten qualitätssichernde Prozesse und Tests vorgegeben werden, und Dokumentationspflichten, die neben der Funktion der Software auch die Bemühungen, einen möglichst objektiven, gesetzeskonformen Algorithmus zu entwickeln dokumentieren.

Natürlich soll nicht nur die Entwicklung sondern auch die Anwendung von Algorithmen, die (Grund-)rechtliche Entscheidungen empfehlen oder treffen, transparent sein. Für die Nutzer*innen und insbesondere für die Menschen, die von den Entscheidungen betroffen sind.

Menschenlesbare Erläuterung der Algorithmen würden das Vertrauen sowohl in den Rechtsstaat als auch in die eingesetzten Algorithmen bestärken.

Wenn solche Algorithmen zudem per se Open Source wären, hätten Entwickler*innen jederzeit die Möglichkeit, zu überprüfen, ob die angegebene Intention wirklich auch entsprechend in der Software umgesetzt ist.

Für ML-basierte, intelligente, Algorithmen könnte darüber hinaus eine Veröffentlichungspflicht für sämtliche Trainingsdaten gefordert werden.

Auch bereits vor der eigentlichen Entwicklung der Technologien müssen Kriterien der Diversität im Recruiting stärker in den Fokus gelangen. “Shit in, Shit out“ sagen die Programmierer*innen und Datenanalyst*innen zu schlechten Entscheidungssystemen, die aufgrund von unreinen oder vorurteilshaften Daten ungenaue Ergebnisse liefern, denn Maschinen lernen von Menschen und übernehmen deren Denk- und Handlungsweisen. Deshalb ist es essentiell, dass eine KI von einer diversen Gruppe entwickelt wird, sodass sie auch die tatsächliche Lebensrealität der Anwender widerspiegelt. Amazon ist bei weitem nicht allein mit der Entwicklung von sexistisch und rassistisch diskriminierenden KI-Systemen. In mehreren Fällen hat sich bereits gezeigt, dass mangelnde Diversität in den Entwicklungsteams zu Diskriminierung, zum Beispiel bei rassistischer Gesichtserkennungssoftware geführt hat. [4]

Auch da sich derzeit nur 12% der Datenwissenschaftler*innen und Programmierer*innen als weiblich identifizieren [5], ist diese Entwicklung leider nicht verwunderlich. Eine neutrale KI kann aber nur entstehen, wenn auch ihre Entwickler*innen aus einer diversen Gruppe stammen.

Die Gestaltung digitaler Schlüsseltechnologien wie KI und ADM ist eine zentrale Machtfrage des 21. Jahrhunderts. Der digitale Arbeitsmarkt, auf dem ein angeblicher Fachkräftemangel herrscht, bleibt jedoch vielen Frauen* und ethnischen Minderheiten verschlossen. Das trägt auch aufgrund der hohen Entlohnung der Berufe im Bereich der Software- und Technologieentwicklung zu finanzieller Ungleichheit von Frauen und marginalisierten Menschen bei. Die Gründe für diese Ungleichheit sind strukturell und kulturell gewachsen und können nicht von den individuellen Frauen allein beseitigt werden. Es müssen daher neben vereinzelten Veranstaltungen wie Girls-Day weitere Anreize wie eine Quote für die Zusammenstellung diverser Entwicklungsteams und eine Zertifizierung für eine inklusive KI geschaffen werden.

---

[1] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[2] https://www.realclearpolicy.com/articles/2018/10/09/report_amazon_falls_short_on_gender_equity_110846.html

[3] https://phys.org/news/2018-07-algorithm-limits-bias-machine.html

[4] https://towardsdatascience.com/https-medium-com-mauriziosantamicone-is-artificial-intelligence-racist-66ea8f67c7de

[5] weforum.org/agenda/2019/06/this-is-why-ai-has-a-gender-problem/

Unterstützer*innen

Vasili Franco (KV Berlin-Friedrichshain/Kreuzberg); Willi Junga (KV Berlin-Treptow/Köpenick); Jan-Klaus Oczenasek (KV Berlin-Kreisfrei); Michael Servatius (KV Berlin-Kreisfrei); Adrian Roeske (KV Berlin-Kreisfrei); Joana Zühlke (KV Berlin-Friedrichshain/Kreuzberg); Enad Altaweel (KV Berlin-Friedrichshain/Kreuzberg); Johannes Mihram (KV Berlin-Mitte); Silke Gebel (KV Berlin-Mitte); Aida Baghernejad (KV Berlin-Friedrichshain/Kreuzberg); Oliver Edgar Münchhoff (KV Berlin-Treptow/Köpenick); Ingrid Bertermann (KV Berlin-Mitte)