GENDER OPEN REPOSITORIUM

Strukturelle Diskriminierung durch KI. Ein Blick auf EU-Governance

Domres, Sophie

2024-11-19

https://doi.org/10.25595/2424

Angenommene Version / accepted version
Working Paper

Empfohlene Zitierung / Suggested Citation:

Domres, Sophie: *Strukturelle Diskriminierung durch KI. Ein Blick auf EU-Governance*. Universität Duisburg-Essen, 2024-11-19. DOI: https://doi.org/10.25595/2424.

Erstmalig hier erschienen / Initial publication here: https://doi.org/10.17185/gender/20241119

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY 4.0 Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:

This document is made available under a CC BY 4.0 License (Attribution). For more information see:

https://creativecommons.org/licenses/by/4.0/deed.en

GenderOpen – Repositorium für die Geschlechterforschung: www.genderopen.de

FREIE UNIVERSITÄT BERLIN

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Strukturelle Diskriminierung durch KI. Ein Blick auf EU-Governance

19.11.2024 Sophie Domres

Die digitale Transformation schreitet in Europa rasch voran und mit ihr der Einsatz künstlicher Intelligenz (KI) in allen Bereichen des gesellschaftlichen Lebens – von der Gesundheitsversorgung über Bildung bis hin zu Justiz und Strafverfolgung. Die Europäische Union (EU) steht dabei vor der Herausforderung, ihre digitalen Technologien so zu gestalten, dass sie die Grundrechte und Werte der Union respektieren und schützen. KI-Systeme bergen jedoch Risiken, die bereits bestehende Diskriminierungen und soziale Ungleichheiten verschärfen können. Frauen und marginalisierte Gruppen sind von dieser Dynamik besonders betroffen.

KI und Diskriminierung: Was sagt die aktuelle Forschung?

Die Entwicklung und der Einsatz von KI-Systemen sind oft eng mit Voreingenommenheit und Diskriminierung verbunden. Studien zeigen, dass KI, wenn sie auf unzureichend diversifizierten Datensätzen trainiert wird, bestehende Vorurteile und Stereotype reproduzieren oder sogar verstärken kann (O'Neil 2016; Benjamin 2019). Ein Beispiel sind algorithmische Entscheidungssysteme, die in der Kreditvergabe, im Personalmanagement und in der Strafverfolgung eingesetzt werden. Hier wurden wiederholt Diskriminierungen von Frauen und Minderheiten dokumentiert, die auf historische Verzerrungen in den zugrundeliegenden Daten zurückzuführen sind (Buolamwini/Gebru 2018). Diese Verzerrungen spiegeln häufig strukturelle Ungleichheiten wider, die tief in unserer Gesellschaft verwurzelt sind.

Forscherinnen wie Safiya Noble und Ruha Benjamin argumentieren, dass diese technologischen Reproduktionen sozialer Ungleichheiten nicht zufällig entstehen, sondern auf systemische Machtstrukturen zurückzuführen sind, die den Prozess der Entwicklung und Nutzung von KI durchziehen (Noble 2018; Benjamin 2019). Wenn die EU das Ziel verfolgt, digitale Souveränität zu erlangen und ihre Autonomie im digitalen Bereich zu sichern (Floridi 2020), muss sie daher besonders darauf achten, dass diese Souveränität nicht auf Kosten der sozialen Gerechtigkeit und der Grundrechte geht.

Intersektionalität für eine gerechte KI-Entwicklung

Um die tiefgreifenden sozialen Implikationen von KI zu verstehen, bietet die Intersektionalitätstheorie von Kimberlé Crenshaw ein wertvolles Werkzeug. Intersektionalität beschreibt das Zusammenwirken verschiedener Diskriminierungsformen aufgrund von Geschlecht, Ethnizität, Klasse und anderen sozialen Kategorien, die individuelle Erfahrungen prägen (Crenshaw 1989). Aus einer intersektionalen Perspektive wird deutlich, dass die Auswirkungen von KI auf Menschen unterschiedlich sind und stark von ihrer sozialen Positionierung abhängen.

Ein Beispiel ist die Geschlechterdiskriminierung in der Arbeitswelt, die durch algorithmische Entscheidungen verstärkt wird. So haben Studien gezeigt, dass Frauen seltener für technische Berufe ausgewählt werden, wenn KI-basierte Systeme bei der Personalauswahl eingesetzt werden (Bogen/Rieke 2018). Diese Systeme tendieren aufgrund der Überrepräsentation männlicher Datensätze dazu, Bewerber zu bevorzugen, die dem stereotypen Bild eines „technischen Experten“ entsprechen. Diese Verzerrungen wirken sich negativ auf die Karrierechancen von Frauen aus und verfestigen traditionelle Geschlechterrollen. Die Theorie der Intersektionalität hilft, solche Wechselwirkungen zwischen Geschlecht, Ethnizität und anderen Faktoren zu identifizieren und zu analysieren. Ein besonders eklatantes Beispiel ist die Ungleichbehandlung bei Gesichtserkennungstechnologien. Studien zeigen hier, dass Frauen und insbesondere People of Color deutlich höhere Fehlerraten erleben als weiße Männer (Buolamwini/Gebru 2018). Diese technologischen Verzerrungen können in sensiblen Bereichen wie der Strafverfolgung schwerwiegende Folgen haben, wenn Personen falsch identifiziert oder unverhältnismäßig überwacht werden.

Ein kritischer Blick auf die KI-Entwicklung

Die Feminist Technology Studies (FTS) bieten einen kritischen Rahmen, um die Prozesse der KI-Entwicklung und -Implementierung zu hinterfragen und geschlechtergerechte Lösungen zu finden. Diese Studien gehen davon aus, dass Technologien nicht neutral sind, sondern von den sozialen und politischen Kontexten geprägt werden, in denen sie entstehen (Wajcman 2010). KI wird von Menschen entwickelt, die ihre eigenen Vorurteile und Perspektiven einbringen, was zur Reproduktion patriarchaler und rassistischer Strukturen führen kann.

Ein wichtiges Anliegen der FTS ist es, die Machtstrukturen zu analysieren, die den Einsatz und die Entwicklung von KI-Systemen prägen. Die ungleiche Verteilung von Frauen und Männern in technischen Berufen und Entscheidungspositionen führt häufig dazu, dass die Perspektiven und Bedürfnisse von Frauen und Minderheiten in der Entwicklung von KI unterrepräsentiert sind. Der Mangel an Diversität im Technologiebereich wirkt sich negativ auf die Gerechtigkeit und Inklusivität von KI-Systemen aus und unterstreicht die Notwendigkeit, unterschiedliche Perspektiven in die Technikgestaltung einzubeziehen (D’Ignazio/Klein 2020).

Der rechtliche Rahmen der EU

Die EU ist bestrebt, sich als globale Akteurin im Bereich der digitalen Souveränität zu positionieren. Mit Initiativen wie dem AI Act und der Datenschutz-Grundverordnung (GDPR) setzt die EU Standards, um die Kontrolle über ihre digitalen Infrastrukturen und Daten zu behalten und gleichzeitig die Rechte ihrer Bürger*innen zu schützen (European Commission 2020). Die digitale Souveränität der EU soll sicherstellen, dass technologische Innovationen im Einklang mit den europäischen Werten Demokratie, Menschenrechte und soziale Gerechtigkeit stehen.

Doch inwieweit berücksichtigt diese Governance-Strategie tatsächlich die Rechte und Bedürfnisse marginalisierter Gruppen? Die Analyse bestehender und geplanter Regulierungsansätze zeigt Lücken, insbesondere hinsichtlich der Berücksichtigung intersektionaler Diskriminierungen und der spezifischen Auswirkungen auf Frauen und Minderheiten (Anwar/Henshaw 2023; Mügge 2024). Ein rein marktregulierender Ansatz reicht nicht aus, um die gesellschaftlichen Implikationen von KI umfassend zu adressieren. Vielmehr muss die EU-Politik sicherstellen, dass die Grundrechte aller Bürger geschützt und technologiebedingte Diskriminierungen aktiv bekämpft werden.

Forderungen an eine gerechte und inklusive KI-Politik

Um KI-Systeme gerechter zu gestalten, müssen die zugrunde liegenden Datensätze diversifiziert werden. Die Daten sollten eine breite gesellschaftliche Repräsentation in Bezug auf Geschlecht, ethnische Zugehörigkeit und andere soziale Merkmale aufweisen. Nur durch die Berücksichtigung vielfältiger Perspektiven und Lebensrealitäten können diskriminierende algorithmische Entscheidungen reduziert werden.
Auch die Zusammensetzung der Entwicklungsteams spielt eine entscheidende Rolle. KI-Technologien müssen von Teams entwickelt werden, die unterschiedliche soziale Hintergründe und Perspektiven mitbringen. Der Zugang von Frauen und Minderheiten zu technischen Berufen muss aktiv gefördert und Barrieren müssen abgebaut werden. Eine intersektionale Perspektive sollte in die Aus- und Weiterbildung von KI-Entwicklern integriert werden.
Die EU sollte außerdem verbindliche Standards zur Kontrolle von Biases in KI-Systemen einführen und deren Umsetzung regelmäßig überprüfen. Der AI Act muss um einen solchen Rahmen ergänzt werden, um algorithmische Diskriminierung systematisch zu verhindern und die Transparenz von Entscheidungsprozessen zu fördern.
Eine inklusive KI-Politik erfordert auch die Einbindung der Zivilgesellschaft in Entscheidungsprozesse. Gesellschaftliche Organisationen, insbesondere solche, die sich für Frauen- und Minderheitenrechte einsetzen, sollten aktiv in die Formulierung und Umsetzung von KI-Gesetzen und -Richtlinien einbezogen werden. Dies könnte durch Konsultationsmechanismen und öffentliche Anhörungen geschehen.

Fazit

Der Schutz der Grundrechte in der digitalen Zukunft Europas erfordert einen integrativen und bereichsübergreifenden Ansatz bei der Entwicklung und Regulierung von KI-Technologien. Die EU hat bereits wichtige Schritte zur Stärkung ihrer digitalen Souveränität unternommen, muss aber sicherstellen, dass dieser Anspruch nicht nur wirtschaftliche, sondern auch soziale und ethische Dimensionen umfasst. KI-Systeme, die die vielfältigen Perspektiven und Lebensrealitäten der europäischen Gesellschaft widerspiegeln, sind notwendig, um Diskriminierung zu bekämpfen und die Werte der EU zu wahren. Die Integration feministischer und intersektioneller Perspektiven in die Technikgestaltung ist ein wesentlicher Schritt, um den Grundrechtsschutz im Zeitalter der KI zu gewährleisten.

Literatur

Anwar, Mhajne & Henshaw, Alexis L. (Eds.) (2023). Critical Perspectives on Cybersecurity: Feminist and Postcolonial Interventions. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780197695883.001.0001

Benjamin, Ruha (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge, UK: Polity.

Bogen, Miranda & Rieke, Aaron (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. Washington, DC: Upturn. Retrieved from https://www.upturn.org/work/help-wanted/ (05.11.24).

Buolamwini, Joy & Gebru, Timnit (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research, 81. Retrieved from https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf (05.11.24).

Crenshaw, Kimberlé (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics. University of Chicago Legal Forum, 1989(1), 139–167. Retrieved from http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8 (05.11.24).

D'Ignazio, Catherine & Klein, Lauren F. (2020). Data Feminism. Cambridge, MA: MIT Press.

European Commission (2020). White Paper on Artificial Intelligence - A European Approach to Excellence and Trust. Retrieved from https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-

artificial-intelligence-feb2020_en.pdf (05.11.24).

Floridi, Luciano (2021). The European Legislation on AI: A Brief Analysis of its Philosophical Approach. Philosophy & Technology, 34, 215–222. https://doi.org/10.1007/s13347-021-00460-9

Mügge, Daniel (2024). EU AI Sovereignty: For Whom, To What End, and To Whose Benefit? Journal of European Public Policy, 31(8), 2200–2225. https://doi.org/10.1080/13501763.2024.2318475

Noble, Safiya U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press.

O'Neil, Cathy (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishing Group.

Veale, Michael & Zuiderveen Borgesius, Frederik (2021). Demystifying the Draft EU Artificial Intelligence Act: Analyzing the Good, the Bad, and the Unclear Elements of the Proposed Approach. Computer Law Review International, 22(4), 97–112. https://doi.org/10.9785/cri-2021-220402

Wajcman, Judy (2010). Feminist theories of technology. Cambridge Journal of Economics, 34(1), 143–152, https://doi.org/10.1093/cje/ben057

Zitation

Sophie Domres: Strukturelle Diskriminierung durch KI. Ein Blick auf EU-Governance, in: blog interdisziplinäre geschlechterforschung, 19.11.2024, www.gender-blog.de/beitrag/strukturelle-diskriminierung-durch-ki/, DOI: https://doi.org/10.17185/gender/20241119

Beitrag lizensiert unter einer Creative Commons Namensnennung 4.0 International Lizenz

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

ub | universitäts bibliothek

Dieser Text wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt. Die hier veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

DOI: 10.17185/gender/20241119
URN: urn:nbn:de:hbz:465-20241119-085029-0

Dieses Werk kann unter einer Creative Commons Namensnennung 4.0 Lizenz (CC BY 4.0) genutzt werden.