

»Der weltweite Anteil von Frauen an
im KI-Bereich Tätigen beträgt 22%.

Ein positiver Zukunftstrend ist
nicht zu erkennen.«

World Economic Forum, Global Gender Gap Report 2018¹



Mehr zu **Geschlecht** und **Künstlicher Intelligenz**



I Geschlecht

Geschlechtsbezogene Verzerrung (gender bias)

Die meisten KI-Verfahren beruhen auf Trainingsdaten, die eine bestimmte Perspektive auf die Realität aufweisen:

- Trainingsdaten können eine historisch inhaltlich korrekte Ungleichverteilung beispielsweise von Macht, Geld, Rechten oder politischer und gesellschaftlicher Teilhabe beinhalten.
- Trainingsdaten können einen verengten Ausschnitt der Wirklichkeit reproduzieren, da sie nur auf bestimmte Diskurse fokussieren.
- Trainingsdaten können durch unzureichende Auswahl der Quellen auch per se ungleich gewichtet sein und damit verzerrend wirken.

I Geschlecht

Geschlechtsbezogene Verzerrung (gender bias)

- Verzerrungen in den Trainingsdaten wirken sich auf die Ergebnisse aus, seien es Texte, Bilder, oder Entscheidungen.
- Beispiele sind etwas die Fortschreibung geschlechtsbezogener Stereotype sein, aber auch die konkrete Benachteiligung beispielsweise von Frauen.
- Prinzipiell sind Verzerrungen in Trainingsdaten identifizierbar, vermeidbar und korrigierbar, aber der Aufwand ist hoch.
- Nicht einfach zu beantworten ist aber die Frage, in welchem Umfang und mit welcher Zielsetzung Verzerrungen korrigiert werden sollen – hierzu gibt es verschiedene Ansichten.

I Geschlecht

Entscheidungssysteme

Trainingsdaten-Bias oder algorithmischer Bias kann zu Benachteiligungen führen, die nicht wünschenswert oder sogar rechtlich unzulässig sind, zum Beispiel:

- Nicht-männliche Bewerber werden seltener eingestellt, weil auch historisch mehr Männer rekrutiert wurden.
- Medizinische Diagnosesysteme funktionieren bei Männern besser, weil mehr Daten vorliegen und die Modelle besser auf männliche Patienten optimiert sind.
- Kredite: Nahezu alle Entscheidungen, die auf Daten beruhen, die historische Ungleichheiten beinhalten, benachteiligen Frauen.

I Geschlecht

Männer, Frauen - und sonst?

Hinsichtlich non-binärer Geschlechtsidentitäten lässt sich vereinfacht sagen:

- Tendenziell werden (auf Basis von Bias) Männer (cis*) bevorzugt, und alle anderen benachteiligt.
- Nonbinäre Geschlechter spielen in Trainingsdaten nur eine verschwindend geringe Rolle, weil entsprechende Merkmale nur in sehr geringem Umfang überhaupt erfasst wurden.

Es besteht möglicherweise die Gefahr, dass non-binäre Personen aufgrund politischer Ideologie durch gezielte Algorithmen benachteiligt werden.

I Geschlecht

Arbeitsmarkt

Laut Bundesagentur für Arbeit (2024)² sind in Deutschland nur 18% der IT-Fachkräfte weiblich.

- Diese Unterrepräsentierung beeinflusst, welche Forschungsfragen priorisiert werden und welche ethischen Standards gesetzt werden.
- Sie wirkt sich ebenso auf Auswahl der Daten und Details der Algorithmen aus.
- Ob Frauen stärker von einem möglichen Arbeitsplatzverlust durch KI bedroht sind, scheint gegenwärtig noch unklar.

I Geschlecht

Maßnahmen & Perspektiven

Folgende Maßnahmen könnten sich positiv für mehr Geschlechtergerechtigkeit auswirken:

- Größere Diversität in Entwicklerteams
- Bessere / geschlechtergerechtere Datensätze
- Fairness-Algorithmen, die Bias in Daten aktiv ausgleichen können.
- Transparente KI-Modelle, deren Entscheidungen nachvollzogen werden können
- Regulierungen & Richtlinien gegen Diskriminierung
- Bewusstsein & Aufklärung: Mehr Forschung, Konferenzen und ethische Guidelines zur Förderung gendergerechter KI.

I Geschlecht

Referenzen

1

World Economic Forum, Global Gender Gap Report 2018

<https://www.weforum.org/publications/the-global-gender-gap-report-2018/>

2

BITKOM: IT- und Digitalberufe: Frauen weiter unterrepräsentiert

https://www.bitkom.org/Presse/Presseinformation/IT-Digitalberufe-Frauen-weiter-unterrepraesentiert#_

Image Prompts

Dall-E Prompts, ohne Tuning, jeweils erster Versuch, zu:

„female ai, no cyborgs“

non-binary tech CEOs

female pope