



Algorithmische Diskriminierung und Künstliche Intelligenz

- Wie gefährlich ist ChatGPT?



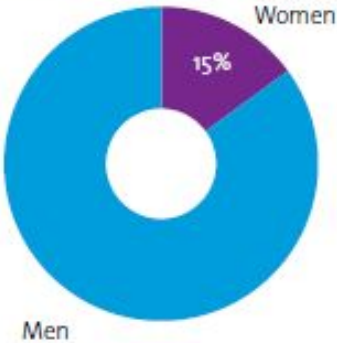
Wer schreibt den Code?

Technik und Gesellschaft stehen seit jeher im Wechselverhältnis: Technik kann bestehende Rollen festschreiben, sie kann ebenso Räume bieten aus diesen auszuberechnen.

Feministische Perspektiven auf KI

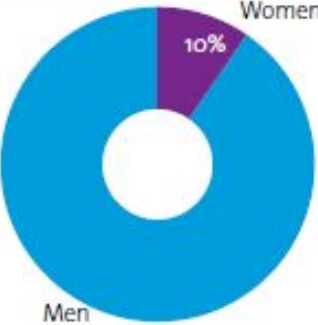


FIGURE 4.
Facebook's estimated AI workforce



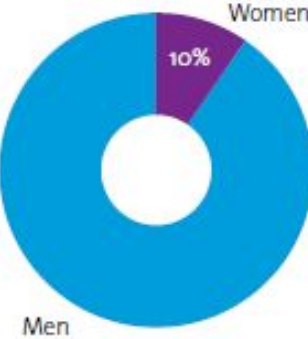
Source: Simonite 2018.

FIGURE 5.
Google's estimated AI workforce



Source: Simonite 2018.

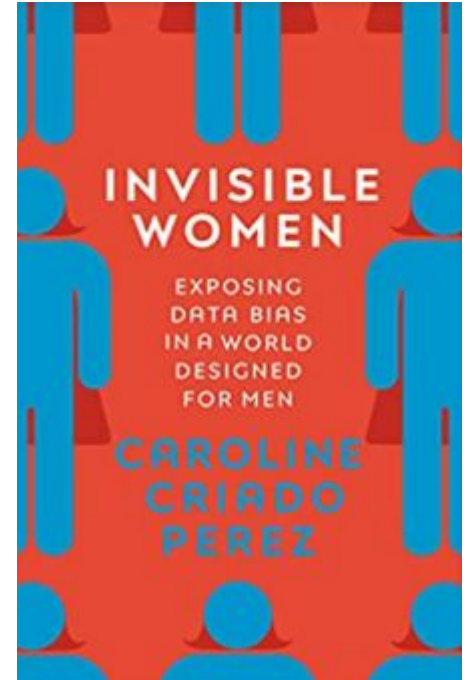
FIGURE 6.
Global AI workforce based on LinkedIn data



Source: World Economic Forum 2018

Künstliche Intelligenz, die auf Big Data basiert, manifestiert diskriminierende Stereotypen innerhalb des Codes

- Verzerrungen in den Trainingsdatensätzen
 - Biometrische Gesichtserkennung war lange Zeit nicht in der Lage, Schwarze Menschen und People of Colour zu erkennen, da sie in erster Linie auf Trainingsdatensätzen *weißer* Menschen basierte.
 - KI-Trainingsdatensätze für autonome Fahrzeuge berücksichtigen nicht die Trainingsdaten von nicht-normalisierten Körpern wie Rollstuhlfahrern



Gesichtserkennung

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
 Microsoft	94.0% 	79.2% 	100% 	98.3% 	20.8% 
 FACE++	99.3% 	65.5% 	99.2% 	94.0% 	33.8% 
 IBM	88.0% 	65.3% 	99.7% 	92.9% 	34.4% 



2018: Das Forschungsprojekt *Gender Shades* verglich Software von drei kommerziellen Anbietern: Microsoft, IBM und Face++ (US)

- bis zu 34,4 % schlechtere Genauigkeit bei der Unterscheidung zwischen „hellerem Mann“ und „dunklerer Frau“
- fast ein Drittel aller Frauen* mit dunklerer Hautfarbe wurden regelmäßig überhaupt nicht erkannt oder fälschlicherweise als Männer identifiziert
- Schwarze Menschen werden im Allgemeinen häufiger falsch identifiziert

ChatGPT



ChatGPT basiert auf GPT-3 (mittlerweile ist zum Teil auch die Nachfolgerversion [GPT-4](#) verfügbar, die noch mehr Funktionen bietet). Das ist ein KI-Sprachmodell des amerikanischen Unternehmens "OpenAI". GPT-3 wurde mit Hunderten Milliarden Wörtern gefüttert, darunter Text aus Büchern, Zeitungen und dem Internet. Anschließend wurden gigantische Rechenleistungen aufgewendet, um diese Textmengen per [Machine Learning](#) auszuwerten und darin Muster zu erkennen.



Die KI hat keine Verbindung zur echten Welt, versteht also nicht, was Worte wie "Boris Becker" bedeuten. Aber sie hat Muster menschlicher Sprache gut genug analysiert, um zu verstehen: Beginnt ein Satz mit "Ein berühmter Tennisspieler ist ...", dann geht er im deutschen Sprachraum mit hoher Wahrscheinlichkeit weiter mit "... Boris Becker".



Man kann sich Text-KIs also vorstellen wie eine viel komplexere und schlauere Version der Auto-Vervollständigung auf dem Smartphone. Zwar hat die KI kein Bewusstsein und auch keinen direkten Zugriff auf die Realität außerhalb der digitalen Welt, aber sie ist so gut darin, Texte zu produzieren, dass diese von menschengemachten Texten oft kaum zu unterscheiden sind.

Ist ChatGPT rassistisch?

OpenAI und Sama stellten unterbezahlte Arbeiter in Kenia ein, um brutale Inhalte für ChatGPT zu filtern

Die Personen, die von Sama im Auftrag von OpenAI beschäftigt wurden, erhielten laut „Time“ je nach Dienstalter und Leistung einen Lohn zwischen 1,32 und 2 Dollar pro Stunde. Ein Sama-Mitarbeiter, der mit dem Lesen und Beschriften von Texten für OpenAI betraut war, erzählte „Time“, dass er unter wiederkehrenden Visionen litt, nachdem er eine grafische Beschreibung eines Mannes gelesen hatte, der in Gegenwart eines kleinen Kindes Sex mit einem Hund hatte. „Das war Folter“, sagte er. „Sie werden im Laufe der Woche eine Reihe solcher Aussagen lesen. Wenn es dann Freitag wird, ist man schon ganz verstört, wenn man dieses Bild durchdenkt.“ Der traumatische Charakter der Arbeit führte schließlich dazu, dass Sama im Februar 2022, acht Monate früher als geplant, seine gesamte Arbeit für OpenAI einstellte.

Ist ChatGPT sexistisch?



Dr. Hadas Kotek 
@HadasKotek



#ChatGPT doubles down on gender stereotypes even when they don't make sense in context

[Tweet übersetzen](#)

HA

The doctor yelled at the nurse because she was late. Who was late?



According to the sentence, the nurse was late.

HA

The nurse yelled at the doctor because she was late. Who was late?



In this sentence, it is still the nurse who was late.

HA

The doctor yelled at the nurse because he was late. Who was late?



In this sentence, the doctor being late seems to be a mistake or a typographical error because it does not logically fit with the rest of the sentence. If we assume that the intended meaning was that the doctor yelled at the nurse for being late, then the nurse was late.



ChatGPT und Datenschutz

Bei manchen Anfragen dürfte das kein Problem sein. Wenn es zum Beispiel um Eingaben wie "Erfinde eine Geschichte über eine grüne Banane" oder "Wie funktioniert eine Demokratie" geht, geben User*innen damit nichts über sich preis.

Bei Anfragen wie "Korrigiere mir folgenden Text und verbessere die Grammatik" oder "Stelle mir für eine Präsentation Folien aus folgendem Text zusammen", also solchen, die personenbezogene Daten oder **vertrauliche Informationen** enthalten, allerdings schon.

Unternehmen wie Samsung sensibilisieren ihre Mitarbeitenden daher aktuell auch auf datenschutzrelevante Probleme wie diese. Andere Firmen haben den Zugriff auf Chat GPT in ihren Netzwerken komplett gesperrt oder verboten. "Wenn da Infos über mich oder interne Firmenmemos drinstehen, kann im ungünstigsten Fall passieren, dass der nächste Chat-GPT-User diese Infos bei einer thematisch ähnlichen Anfrage als Ergebnis geliefert bekommt", erklärt Michael Gessat.

ChatGPT und Fakten



Teresa Kubacka @paniterka_ch · 5. Dez. 2022



I asked about more publications about multiferroics in general. I got a second shortlist – and it was all fake, again. Many names agreed again (Spaldin, Khomskii), but none of the citations was correct. Some had wrong DOIs, others wrong authors etc.



pretend you are a scientist in condensed matter physics. You are writing a dissertation about multiferroics. What are the publications you would cite in your dissertation? include full bibliographic data with DOI



Here are some publications that I would consider citing in a dissertation about multiferroics, along with their full bibliographic data and DOIs:



- Dagotto, E. and Hotta, T. "Multiferroic materials: New frontiers in physics." *Physics Reports*, vol. 442, no. 5-6, 2007, pp. 1-153, doi:10.1016/j.physrep.2007.05.002.
- Yanase, A. and Nagaosa, N. "Theory of multiferroics: Recent progress and future perspectives." *Journal of the Physical Society of Japan*, vol. 74, no. 10, 2005, pp. 2734-2740, doi:10.1143/JPSJ.74.2734.
- Khomskii, D.I. "Multiferroics: Different ways to combine magnetism and ferroelectricity." *Physics*, vol. 2, no. 4, 2009, pp. 20-36, doi:10.1103/Physics.2.20.
- Ramesh, R. and Spaldin, N.A. "Multiferroics: Progress and prospects in thin

»Dringender Handlungsbedarf«

Journalistenverband verlangt klare Regeln für Künstliche Intelligenz

»Kollege KI« darf die Redaktion nicht ersetzen: Der Journalistenverband formuliert ein Positionspapier, wie mit Inhalten umzugehen sei, die von Künstlicher Intelligenz generiert werden.



Wir brauchen feministische KI (Künstliche Intelligenz), weil KI-Systeme unsere Entscheidungen und unser Verhalten beeinflussen und somit Auswirkungen auf die Geschlechtergerechtigkeit haben können. KI-Systeme werden von Menschen entwickelt und trainiert und spiegeln somit die Vorurteile, Stereotypen und Ungleichheiten wider, die in unserer Gesellschaft existieren.



Feministische KI zielt darauf ab, diese Ungleichheiten und Vorurteile zu beseitigen, indem sie eine inklusive und diverse KI-Entwicklung und -Anwendung fördert. Feministische KI-Systeme berücksichtigen die geschlechtsspezifischen Auswirkungen von KI auf Frauen und andere unterrepräsentierte Geschlechter und bemühen sich, sie zu minimieren oder zu beseitigen.

Ein Beispiel dafür ist die Verwendung von Daten, die geschlechtsspezifische Unterschiede berücksichtigen, um sicherzustellen, dass KI-Systeme nicht unbeabsichtigt diskriminierend sind. Ein weiteres Beispiel ist die Förderung der Teilhabe von Frauen und anderen unterrepräsentierten Gruppen in der KI-Entwicklung, um sicherzustellen, dass ihre Bedürfnisse und Perspektiven berücksichtigt werden.

Insgesamt trägt feministische KI dazu bei, dass KI-Systeme nicht nur technisch effektiv sind, sondern auch ethisch und sozial verantwortlich, um sicherzustellen, dass sie der Förderung der Geschlechtergerechtigkeit dienen.

🔄 Regenerate response



beta

Why is the sky blue?



Synthesize

All years

Share This Page

Analogy now steps in and suggests that the colour of the sky is due to the reflection of sunlight from transparent bodies of such a size that the reflected light is the blue of the first order.

The Colour and polarisation of Blue Sky Light

Nature 1901

Cite Share

The theoretical and numerical results show that the scattering of the sunlight by gas molecules is the main reason for the blue sky, and the color of cloud depends on the thickness of the cloud, the size of the gas particles and the gas components.

Theoretical Analysis of Blue Sky and White Cloud

Goureneal of Jiyuan Vocational and Technical College | Zhu Jing-cheng et al. 2005

Cite Share

Christian

New Feature: Consensus Meter



Fragen?

WENN KI, DANN FEMINISTISCH IMPULSE AUS WISSENSCHAFT UND AKTIVISMUS

Hrsg. von netzforma* e.V.
Berlin 2021

EINLEITUNG

Wenn KI, dann hermitisch
"netzforma" – Wenn für hermitische Netzpolitik 7

ÜBERWACHUNG

1. — Überwachung und Künstliche Intelligenz
Wie überwacht man eigentlich wen?
Francesca Bonnet & Johanna Luise Mollathin 15

DIGITALE GEWALT

2. — Kontrollverlust und digitalisierte Ermordung
– Das Gewaltpotential Künstlicher Intelligenz
Lorenz Simon 31

ALGORITHMISCHE ENTSCHEIDUNGS- SYSTEME

3. — Effiziente Ungerechtigkeit
Adriano Gajardo Santos 47

4. — Automatisierte Ungerechtigkeit
Wie vergrößert sich die Einkommens- und
gesellschaftliche Machtungleichheit in Produktion
? Heiko Ludwig 65

5. — Künstliche Beschäftigung: Das Ende menschlicher Vermögensmehrheit
oder die Anfänge von Makroökonomie 2.0?
Dennis Linden 77

ROBOTIK

6. — Feminismus und autonome Waffensysteme
Dora Achawan 91

7. — Harmonizing Robots: The Future with Harmony
Christiane Fritsch 107



netzforma*e.V.
Verein für feministische Netzpolitik

Mosene/ Dinar / Koster / Schmidt, Zoff um Wiki, In: MISSY MAGAZIN #03/20 Unser Netz! 14 Seiten über das feministische Internet (Juni/Juli2020) <https://missy-magazine.de/blog/2020/05/11/zoff-um-wiki/>

AI Lab, HIIG (2020): ZUKUNFTSFAKTOR DIVERSITÄT - Positionspapier zum Roundtable „KI und Frauen**“ https://www.hiig.de/wp-content/uploads/2020/12/Positionspapier-KI-und-Frauen-WEB_V2.pdf

Mosene, Katharina (2021): Gutachten zur Recommendation on the Ethics of Artificial Intelligence, SHS/IGM-AIETHICS/2021/JUN/3 Rev.2., Politikbereich Gender In: Kettemann, Matthias C. (2022): UNESCO-Empfehlung zur Ethik Künstlicher Intelligenz. Bedingungen zur Implementierung in Deutschland <https://www.unesco.de/wissen/wissenschaft/ethik-und-philosophie/studie-umsetzung-ki-ethik-empfehlung>

Mohamed, Shakir; Png, Marie-Therese und Isaac, William (2020): Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. In: Philosophy & Technology, 33(4), S. 659–684. <https://doi.org/10.1007/s13347-020-00405-8>

Netzforma* e.V. (2020): Publikation: Wenn KI, dann feministisch. Impulse aus Wissenschaft und Aktivismus https://netzforma.org/wp-content/uploads/2021/01/2020_wenn-ki-dann-feministisch_netzforma.pdf

Katharina Mosene
Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI)

k.mosene@leibniz-hbi.de

Twitter [@mosenii](#)

CC BY-NC 4.0