GENDER BIAS IN DER KI – NUR EINE UTOPIE?

DR.-ING. FERESHTA YAZDANI

Technology Consultant / Data Scientist

Lufthansa Industry Solutions

Aktuell

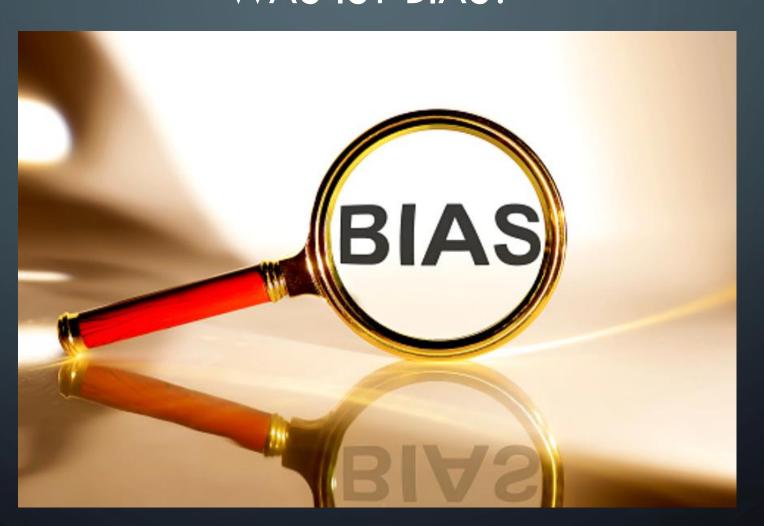
- Technology Consultant / Data Scientist bei der Lufthansa Industry Solutions
- Über 8 Jahre Erfahrung in der Kl und Robotik
- Mitglied und Bremer City Lead Women in Al
- Ausgezeichnet als GI Junior Fellows 2021

Background

- Diplom in Technischer Informatik @ Universität Bremen
- PhD in KI und Robotik@ Institute for Artificial Intelligence, Universität Bremen



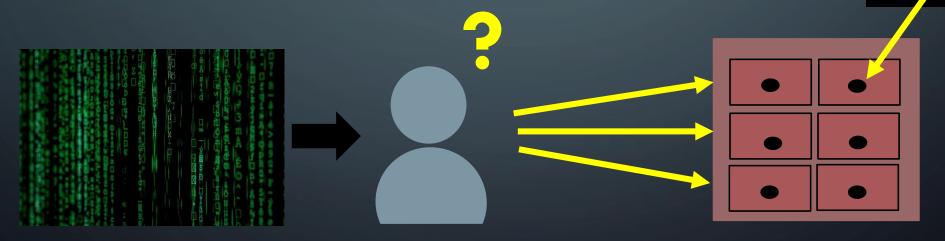
WAS IST BIAS?



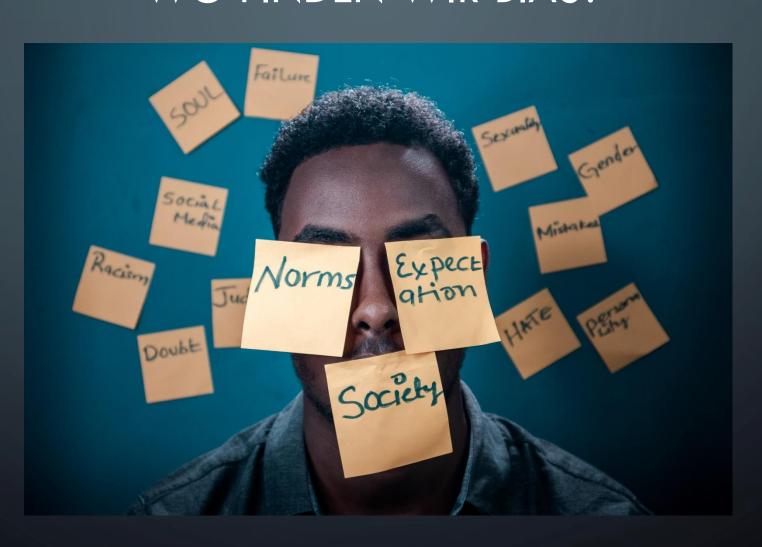
UNCONSCIOUS BIAS

B098

- Unbewusste kognitive Verzerrungen
 - Automatische Stereotypen
 - Unbewusste Denkmuster



WO FINDEN WIR BIAS?



BEISPIELE FÜR BIAS

- Erfahrungen, kulturelle Einflüssen und Darstellungen in den Medien
- Stereotypen: typische Frauen- und Männerberufe bei Eignungsbeurteilung
- Arbeitsmodell: Teil-/Vollzeit bei Leistungsbeurteilung



KI-ANWENDUNGEN

- KI künftig in allen Branchen integriert
 - Gesundheitswesen Data Science
 - Marketing Kundensegmentierung
 - Industrie 4.0 Predictive Maintenance
 - IoT Smart Buildings













Die Welt besitzt Daten

+

Die Welt hat einen BIAS

Daten haben einen BIAS





BEISPIELE FÜR BIAS IN TECHNOLOGIEN

BEISPIEL FÜR BIAS IN AUTOMOTIVE

"Eine Durchschnitts-Dummy-Frau existiert nicht."

Frauen und Beifahrerinnen sterben etwa

17 Prozent häufiger

bei Autounfällen.

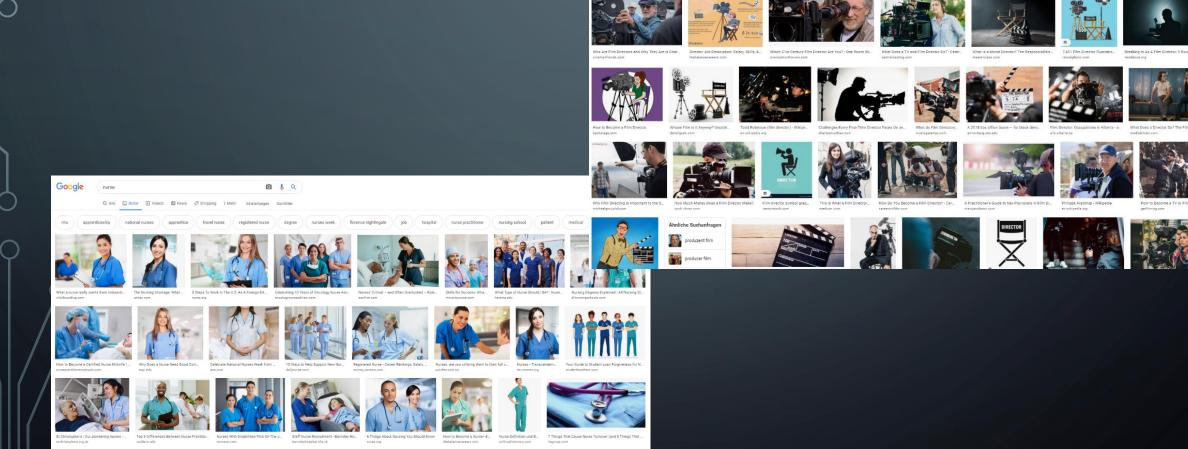


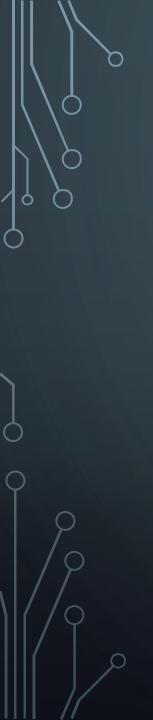
Angeschnallte Frauen haben eine höhere Wahrscheinlichkeit von

73 Prozent

ernsthaft verletzt zu werden.

BEISPIEL FÜR BIAS: STEREOTYPEN





Daten haben einen BIAS

+

KI arbeitet mit Daten

=

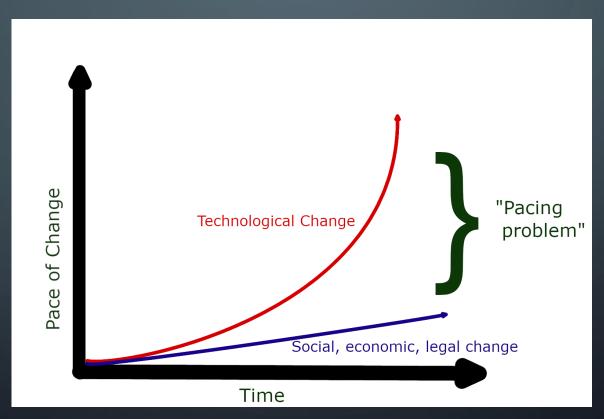
Die KI hat einen BIAS



WARUM REFLEKTIERT DIE KI UNSERE GESELLSCHAFTLICHEN PROBLEME WIEDER?



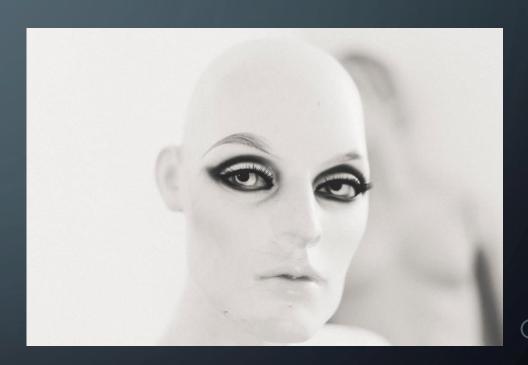
PACING PROBLEM



"Gesetze und Verordnungen können mit Technologien nicht mithalten!"

BEISPIELE GENDER BIAS IN DER KI

- Personal assistants (e.g.Alexa, Siri etc.) vs.
 IBM Watson, Salesforce Einstein
- Amazon Al recruiting process systematische
 Diskriminierung von Frauen
- Gesichtserkennungstechnologien von wichtige Technologie-Giganten zeigen Bias in Hautfarbe und im Geschlecht





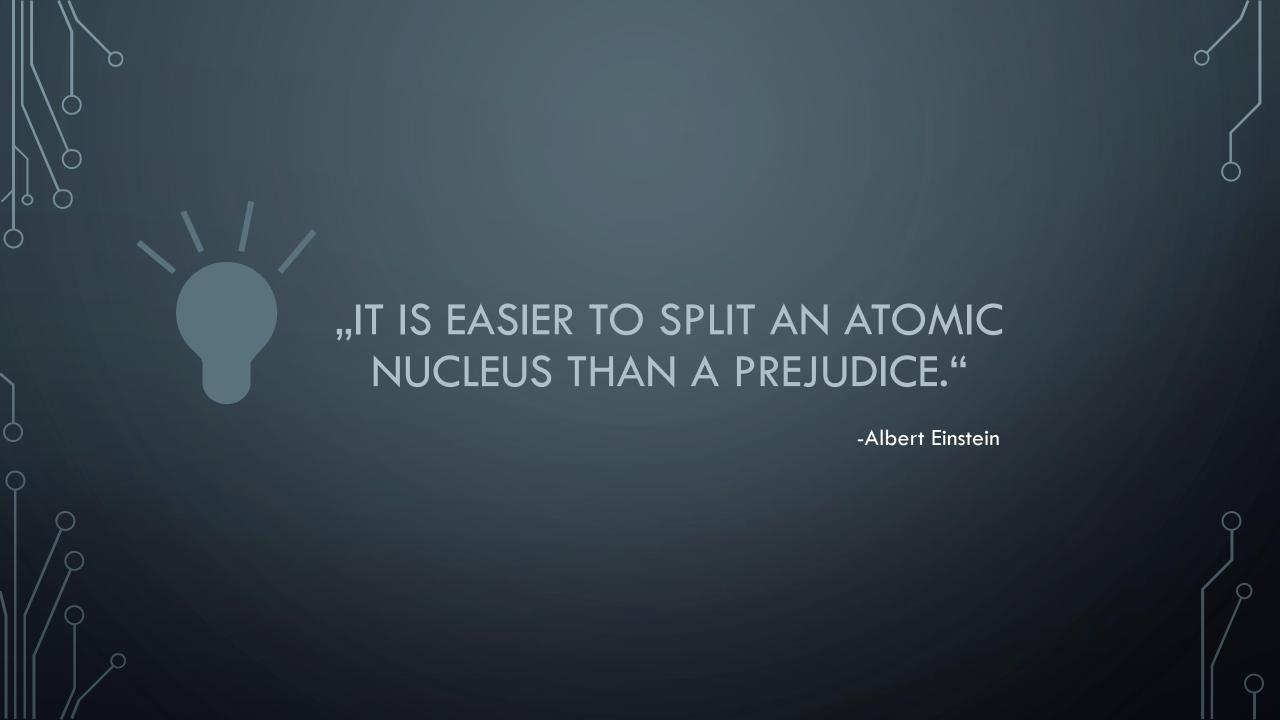
"BIAS in der KI resultiert vom Verhalten, der Einstellung und Entscheidung von Menschen, welche große Auswirkungen auf andere Menschenleben haben kann."



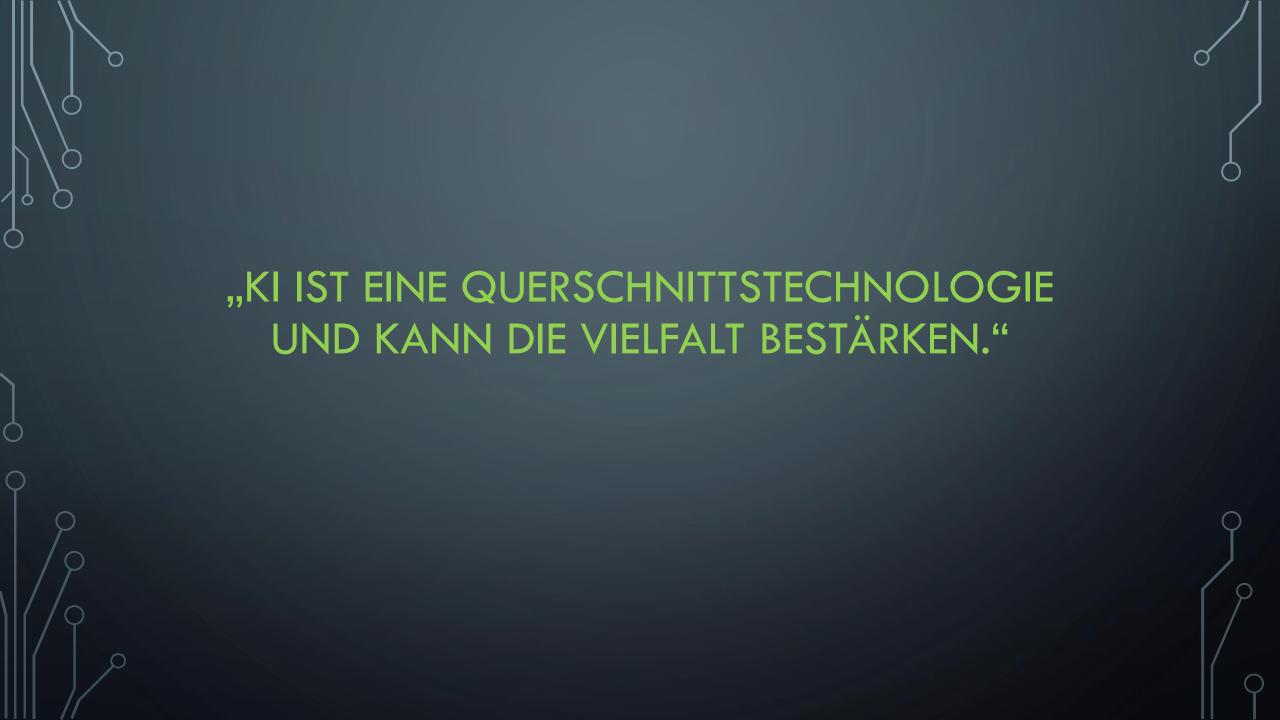
RISIKEN VON BIAS IN DER KI

- Projezieren von vergangenen und präsenten Vorurteilen in die Zukunft
- Gesundheitliche Folgen in der Zukunft (e.g. dummy crash tests)
- Historische und soziale Ungleichheiten (e.g. Kreditwürdigkeit)

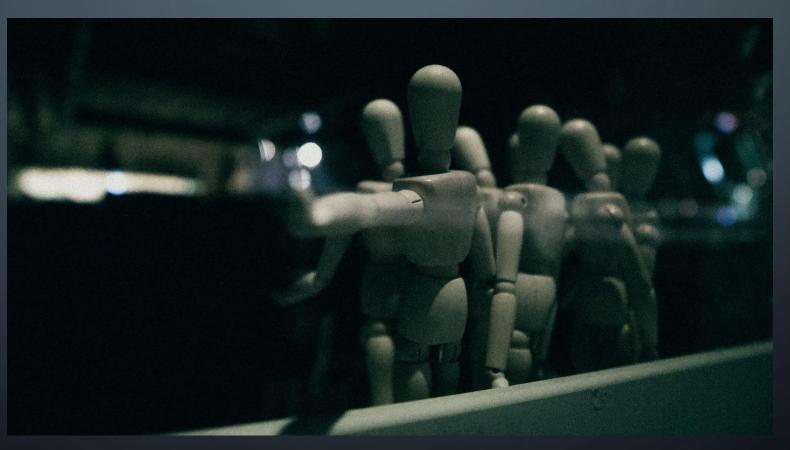








LÖSUNGEN FÜR DIE REDUZIERUNG DES BIAS IN DER KI





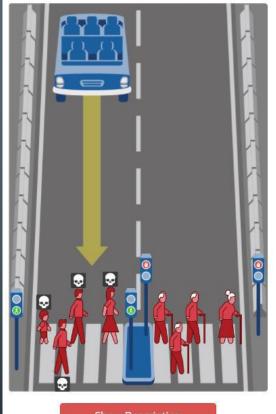
- Sicherstellung von gut erforschten und entwickelten Machine Learning technologien
- Ziel: Menschheit helfen, die Einführung von KI Systemen fair zu steuern

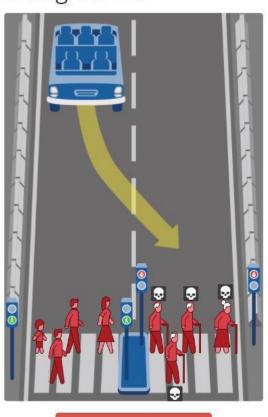
"Ethiks sind in jeder Kultur verschieden"



Kann ein selbst-fahrendes Auto
ethisch handeln, wenn
Menschen Ethik für sich
selber definieren?

What should the self-driving car do?





Show Description

Show Description

Ethik, Fairness und BIAS
 müssen am Start jedes
 Projektes erfolgen
 Kein Nachzügler!





 Diverse Teams und Perspektiven bei der Entwicklung von KI-Anwendungen

"Diversity prediction theorem"



• Train the kids!

- Problem geht auf Kindheit zurück: Spielzeuge, Farben
- Erziehung in Schulen und Universitäten: Vorbilder und Mentoren schaffen
- Unterrichten und zur Verfügung stellen von Ethik Trainings für Studenten



"If we want to have unbiased technologies, we need to work on our own mindsets and attitudes."

