



Automatisierte Entscheidungssysteme und Diskriminierung?

Clara Helming | Tagung „Künstliche Intelligenz und rassistische Diskriminierung in der Arbeitswelt“

Gelbe Hand, 10. November 2023



/ Was ist KI?

- Künstliche Intelligenz (KI)
- Automatisierte Entscheidungssysteme (ADM-Systeme)
- Algorithmus



~~„Künstliche Intelligenz“~~



Automatisierung von Entscheidungen mit
Software

Konkrete Anwendung im Betrieb!



/ Auswirkungen auf Menschen

- Algorithmische Systeme, die eingesetzt werden (können),
 - für Entscheidungen über Menschen
 - um Inhalte zu generieren, die Entscheidungen von / über Menschen beeinflussen.
- vs Fließband, Wetterprognose, Rechtschreibprogramm, ...





/ Wie entsteht rassistische KI?

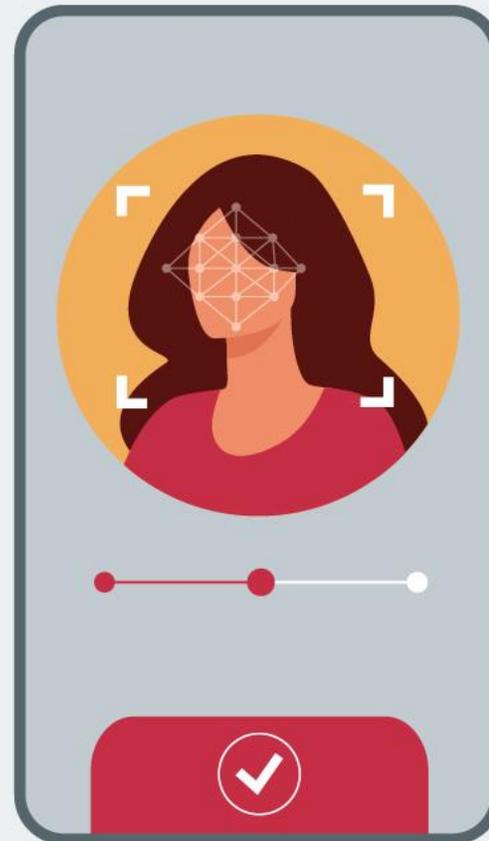
... durch Menschen

→ Technik ist nie neutral

DATEN

Sind die Daten repräsentativ für die Gruppen, auf die das ADM-System einen Effekt haben wird?

Wer wählt die Daten aus und warum wählt die Person diese Daten (z. B. Qualität der Daten, Verfügbarkeit oder Kosten)?



Sind die Daten aktuell und ausreichend? Wer prüft das?

Welche Fragen können mit den Daten beantwortet werden?

Wer hat die Daten annotiert und nach welchen Kriterien?



ZIEL

Welcher Zweck soll durch den Einsatz eines ADM-Systems erfüllt werden?

Welche (kritischen) Fragen werden gestellt?
Welche Alternativen in Betracht gezogen?

Welche Annahmen über ADM-Systeme stecken hinter der Einschätzung, dass dies die beste Lösung für das identifizierte Problem ist?

Wer definiert das Ziel des Einsatzes, das Problem und die Lösung? Welche Annahmen hat diese Person?

Welche Perspektiven auf die Gesellschaft hat die Entscheidungsperson?



MODELL



Auf welchen Annahmen über das Problem, die Realität oder das Ziel basiert das Modell?

Wer übersetzt das Problem in ein Modell und wie wird es übersetzt?

Welchen Ausschnitt der Realität bildet das Modell ab? Ist die Vereinfachung dem Problem angemessen?

Ist das Modell geprüft worden? Nach welchen Qualitätsstandards? Unternehmensintern oder von unabhängiger Stelle?

ANWENDUNG

Sind Anwendende geschult, das System anzuwenden? Kennen sie die Möglichkeiten und Grenzen?

Wer wendet das ADM-System an und wofür?

Welche Kompetenzen und welches Wissen haben die Anwendenden?

Haben Anwendende die Möglichkeit oder Freiheit, sich anders zu entscheiden als vom System empfohlen?

Falls formal möglich: Ist dies im Arbeitsalltag realistisch, z. B. angesichts knapper Zeit- oder Personalressourcen? Welche Konsequenzen ergeben sich bei einem Widerspruch für Anwendende?





/ Ursachen für diskriminierende Ergebnisse

ADM-Systeme sind nicht neutral oder objektiv:

- Daten
- Ziele
- (mathematisches) Modell
- Anwendung

→ Für soziale Probleme gibt es keine technischen Lösungen



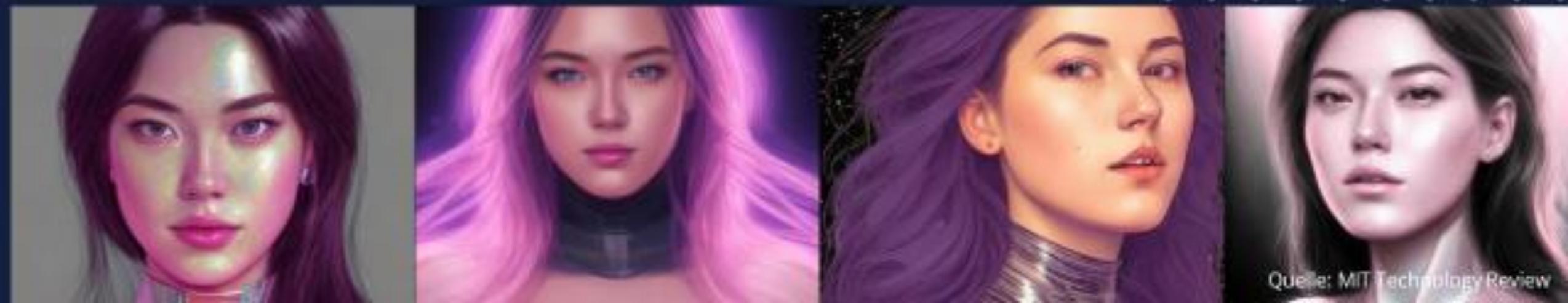
/ Beispiele

The viral AI avatar app Lensa undressed me—without my consent

My avatars were cartoonishly pornified, while my male colleagues got to be astronauts, explorers, and inventors.

By **Melissa Heikkilä**

December 12, 2022



Wrongfully Accused by an Algorithm

In what may be the first known case of its kind, a faulty facial recognition match led to a Michigan man's arrest for a crime he did not commit.





/ „Kindergeldskandal“ Niederlande

- Rassistische Diskriminierung bei der Überprüfung von Kindergeldansprüchen
- Fälschlicherweise: Forderung von Rückzahlungen des Kindergeldes



/ In der Arbeitswelt



X-GMBH

NEUIGKEITEN

SERVICE

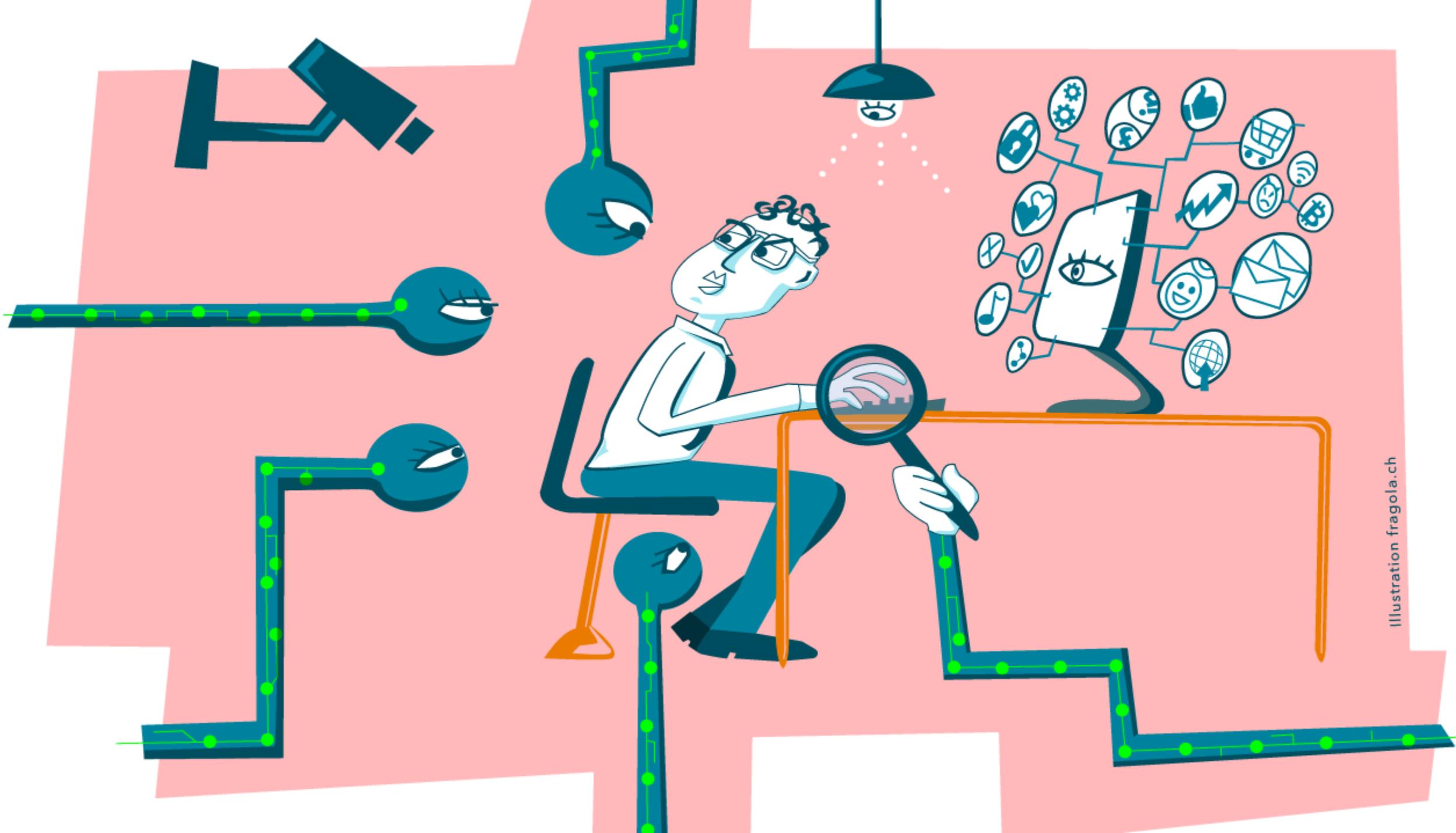
ÜBER UNS

KONTAKT

**WIR
SUCHEN ...**

BEWERBEN









/ Sollten wir Algorithmen und KI überhaupt nutzen?

- Ja. So dass sie uns (allen) nützen.
- Aber: Algorithmen und KI ...
 - ... sind ein Werkzeug, das von Menschen entwickelt und eingesetzt wird.
 - ... sind kein Zweck an sich.
 - ... sind keine Naturgewalt.



/ Lösungsansätze

- Transparenz
- Nachvollziehbarkeit
- Verantwortlichkeit



/ Mitbestimmungsrecht!

- Welche Entscheidungen sollen automatisiert werden und warum?
- Wer profitiert davon und für wen hat es Nachteile?
- Wer trägt die Verantwortung?
- Gibt es eine Möglichkeit zu widersprechen oder Fragen zu stellen?



VIELEN DANK!

Kontakt: <https://algorithmwatch.org/de>

helming@algorithmwatch.org