



Faire KI?

Grundlagen Algorithmischer Fairness

- und was das für Geflüchtete
bedeutet

Clara Strasser Ceballos (LMU München)

Vortrag

17. Oktober 2025



Clara Strasser Ceballos (sie/ihr)

Ganz kurzer Werdegang:

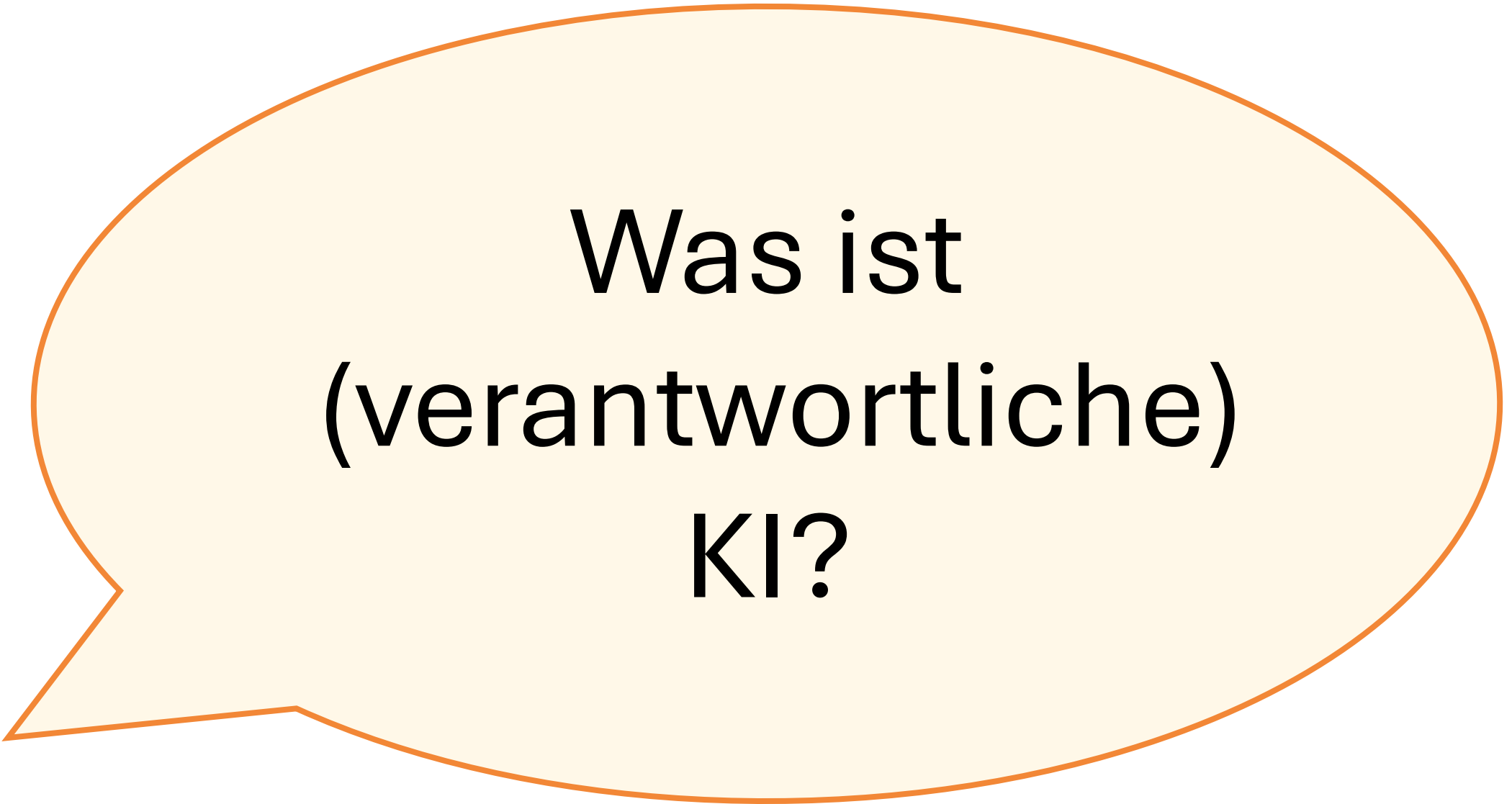
- Abitur an der Deutschen Schule Mailand, Italien
- Bachelor in **Volkswirtschaftslehre (VWL)** an der LMU
- Master in **Statistik** an der LMU
- Seitdem PhD über **Algorithmen in Migration & Asyl** an der LMU



Kontakt Daten:

E-Mail: Clara.StrasserCeballos@stat.uni-muenchen.de

!!! Ich freue mich immer sehr über Eure Fragen !!!



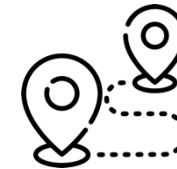
Was ist
(verantwortliche)
KI?

KI: Definition, Anwendungsbereiche, Chancen & Risiken



Künstliche Intelligenz (KI)

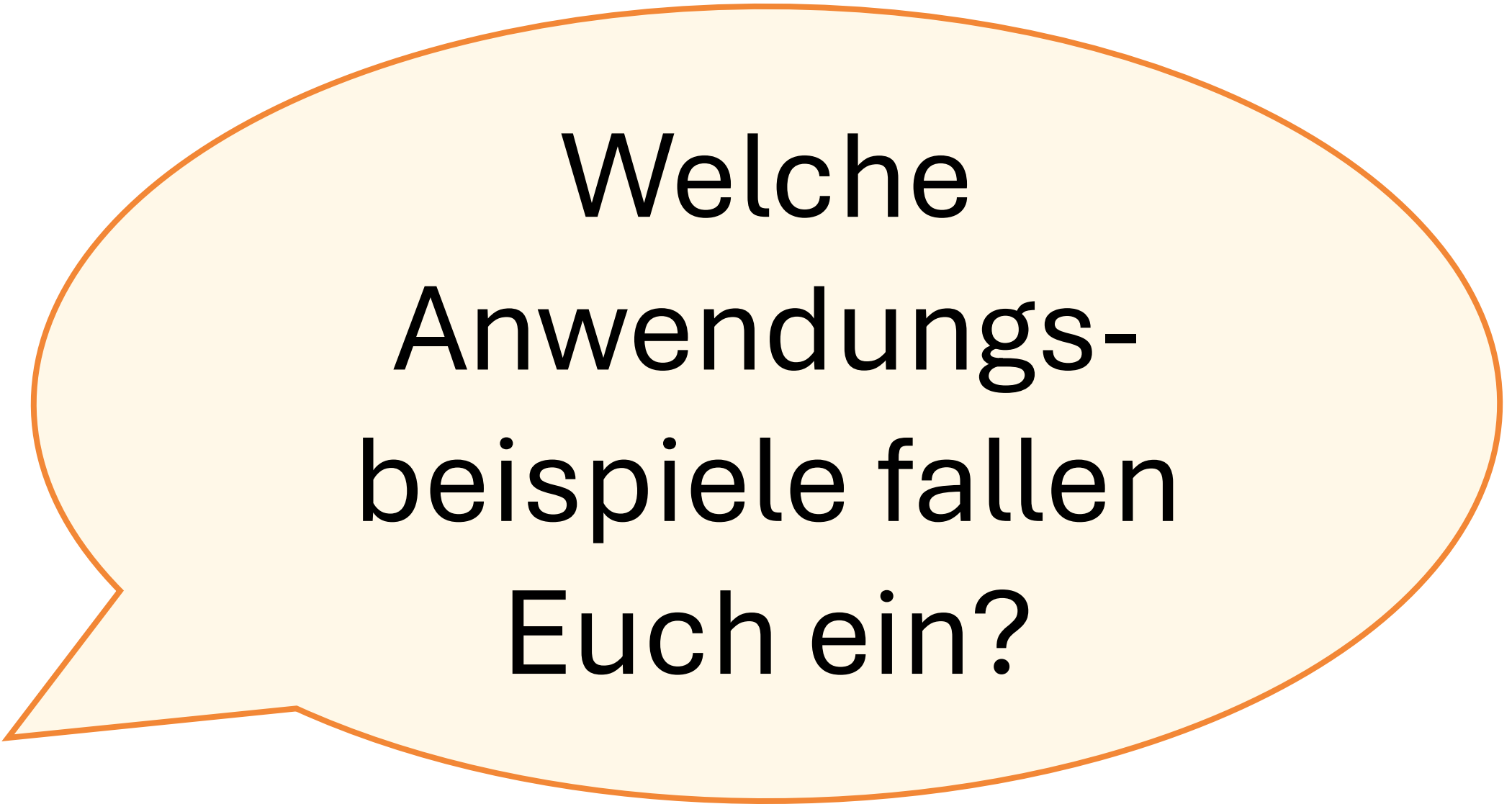
Systeme, die mithilfe von Algorithmen aus Daten lernen und Entscheidungen entweder unterstützen oder teilweise/vollständig automatisieren.



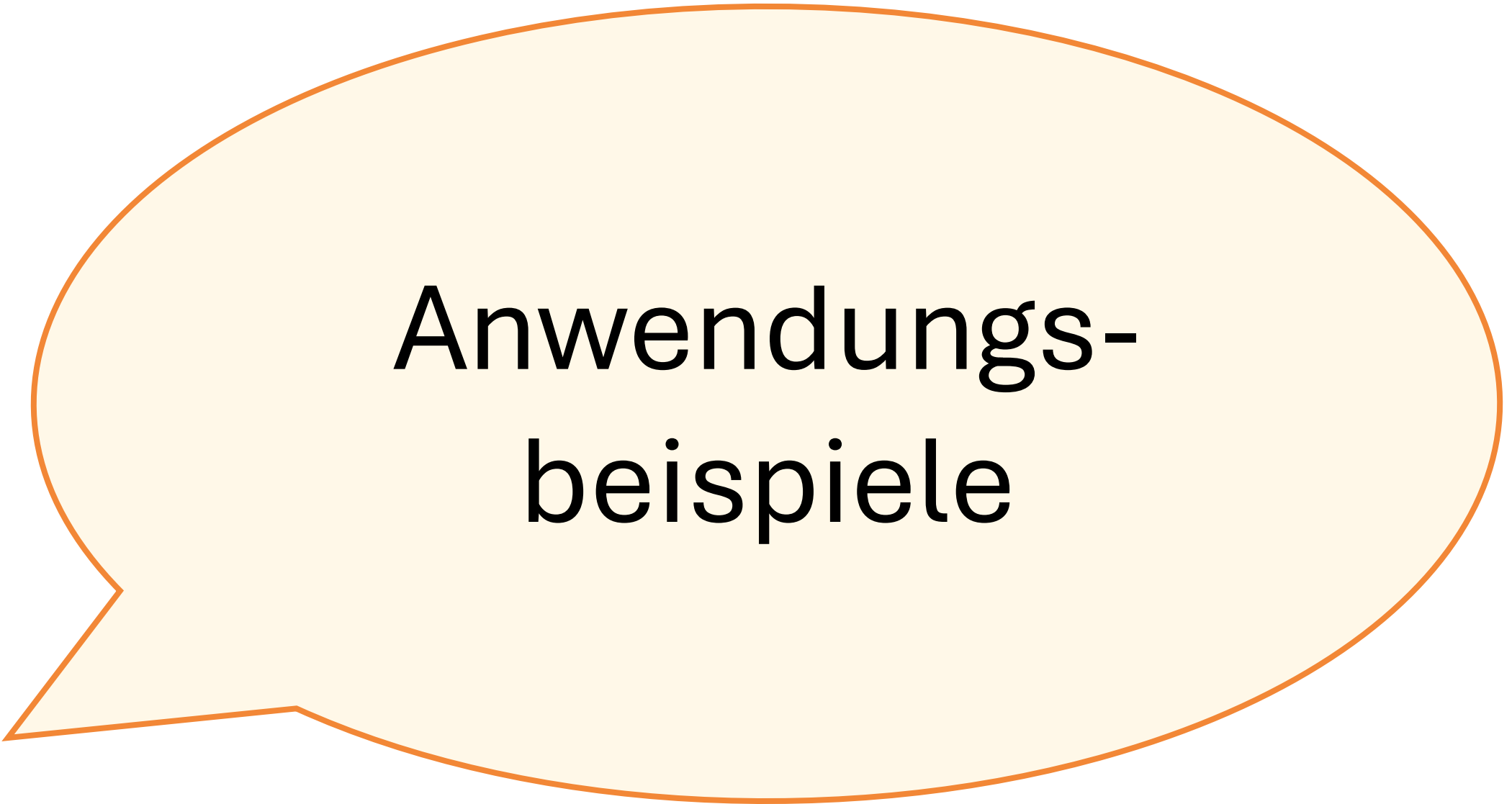
- **Schnell**
- **Effizient**
- **Konsistent**
- **Persönlich**
- ...



- **Bias und Diskriminierung**
- **Intransparenz**
- **Verantwortungsdiffusion**
- **Sicherheits-/Privacy-Risiken**
- ...



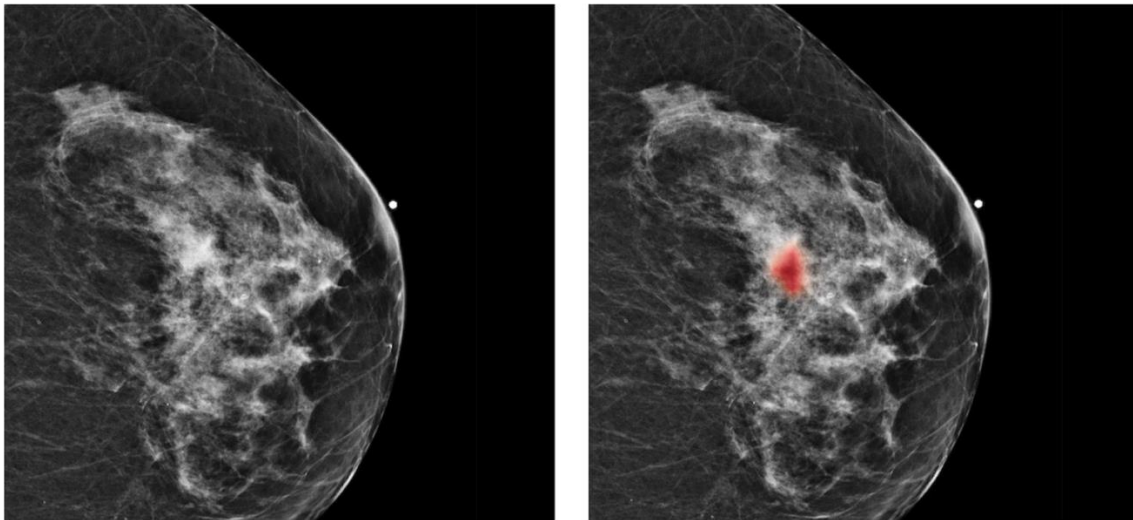
Welche
Anwendungs-
beispiele fallen
Euch ein?



Anwendungs- beispiele

KI wird in der **medizinischen Bildgebung** („Medical Imaging“) eingesetzt, beispielsweise bei Röntgen, CT, MRT, Ultraschall und in der Pathologie. Mithilfe von KI können Auffälligkeiten erkannt und segmentiert, Fälle priorisiert, Verläufe verglichen und Befundvorschläge erstellt werden.

Can Artificial Intelligence Perfect Mammography?



Radiologists missed a subtle mass on a mammography image (left). But it was identified by Perlmutter Cancer Center researchers using artificial intelligence (right, in red) as highly likely to be cancerous. A biopsy confirmed that the lesion was malignant. SCAN COURTESY ARTIE SHEN

<https://nyulangone.org/news/can-artificial-intelligence-perfect-mammography>

HELMHOLTZ MUNICH

EN | DE  

About Us Research Transfer Career Newsroom

Computational Health Center

Institute of Machine Learning in Biomedical Imaging

<https://compai-lab.github.io/>

Unsere Vision & Mission

Discovering Future Health.

Unsere Welt verändert sich ständig, mit großen Auswirkungen auf unsere Gesundheit. Volkskrankheiten wie Diabetes, Allergien und Lungenerkrankungen nehmen weltweit rasant zu. Gleichzeitig entstehen neue Krankheiten, beispielsweise aufgrund des Klimawandels.

Wir entwickeln Lösungen für eine gesündere Zukunft. Unsere Spitzenforschung bildet die Grundlage für medizinische Innovationen. Gemeinsam mit unseren Partnern beschleunigen wir den Transfer unserer Forschung, damit neue Ideen aus den Laboren noch schneller in der Gesellschaft ankommen.

Medicine & Healthcare



AI has the potential to fundamentally transform the future of medicine and healthcare by enabling earlier and more accurate diagnosis and better treatment, leading to improved outcomes for patients and increased efficiency in healthcare. The emergence of AI for medicine and healthcare also offers a number of transformative opportunities for economic growth. Examples cover prevention and early detection, e.g. AI for wearable devices as well as AI for screening (e.g. mammography). In relAI, we combine the expertise in AI healthcare and medicine to successfully tackle these challenges.

A key requirement for the successful deployment of AI in clinical environments is the development of safe, secure, and trustworthy ML techniques. In particular, advances are required in robust and data efficient learning, privacy preservation, and interpretable deep learning.

<https://zuseschoolrelai.de/>

Risiko: Ungleiche Genauigkeit der Vorhersagen für bestimmte Gruppen (z. B. Geschlecht, Alter, Haut-/Gewebebetöne).

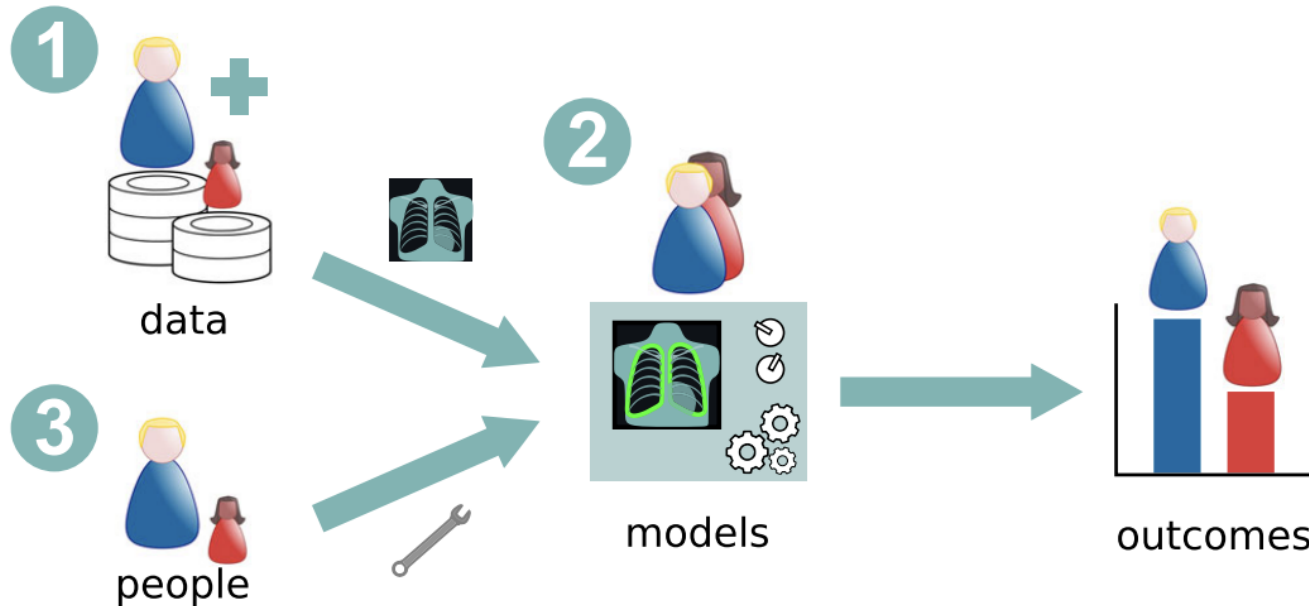


Fig. 2 | Main potential sources of bias in AI systems for MIC. The data being fed to the system during training (1), design choices for the model (2), and the people who develop those systems (3), may all contribute to biases in AI systems for MIC.

Gründe für Verzerrungen in (1):

- Ungleiche Repräsentation (z.B. Röntgenbilder stammen aus Militär-oder Arbeitsmedizin-Kontexten)
- Unterschiede in der Datenerhebung (Bildqualität variiert nach Krankenhaus)
- Unterschiede in der Gesundheitsversorgung
- ...

KI wird von Krankenkassen und Krankenhäuser verwendet um zu entscheiden, welche Patient*innen zusätzliche medizinische Betreuung oder Programme (z. B. für chronisch Kranke) bekommen sollen.

2019

Dissecting racial bias in an algorithm used to the health of populations

ZIAD OBERMEYER , BRIAN POWERS, CHRISTINE VOGELI, AND SENDHIL MULLAINATHAN  [Authors Info & Affiliations](#)

SCIENCE • 25 Oct 2019 • Vol 366, Issue 6464 • pp. 447-453 • DOI: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342)

↓ 166,259 💬 3,549



Racial bias in health algorithms

The U.S. health care system uses commercial algorithms to guide health decisions. Obermeyer *et al.* find evidence of racial bias in one widely used algorithm, such that Black patients assigned the same level of risk by the algorithm are sicker than

Healthcare algorithm used across America has dramatic racial biases

System sold by Optum estimates health needs based on medical costs, which are much less than for white patients, report finds



KI war rassistisch verzerrt!

Grund: KI wurde nicht auf tatsächliche Gesundheit, sondern auf **medizinische Kosten** trainiert.

→ Schwarze Patient*innen mussten deutlich kränker sein als weiße, um das gleiche Risikoniveau zugeteilt zu bekommen.

<https://www.science.org/doi/10.1126/science.aax2342>

KI wird im Arbeitsmarkt eingesetzt, um z.B. Arbeitsmarktanzeigen automatisch zuzuordnen, Lebensläufe (CVs) automatisch zu erkennen und zu bewerten, Bewerber*innen zu klassifizieren, Förderprogramme zuzuordnen,...

2018

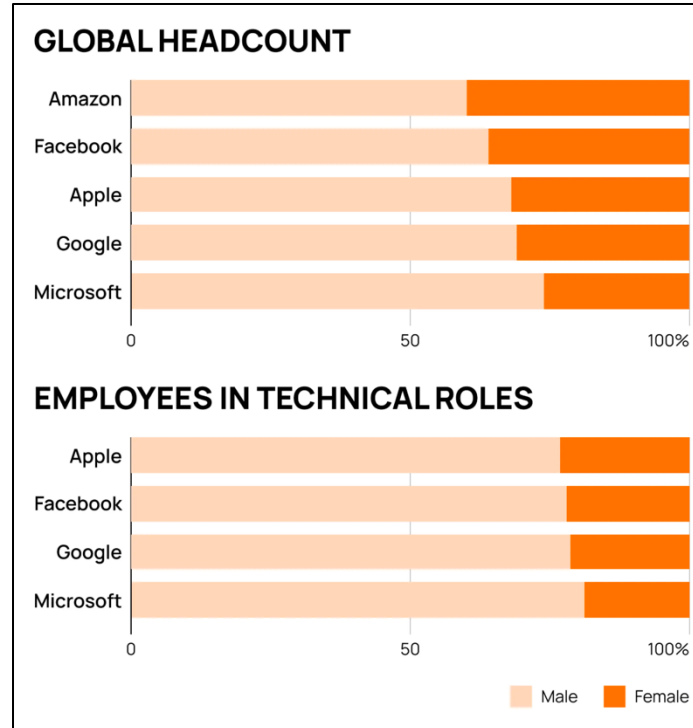
Künstliche Intelligenz diskriminiert (noch)

Der Bewerbungsroboter von Amazon hat Frauen diskriminiert. Wie konnte das passieren? Und wie können Algorithmen geeignete Kandidaten für einen Job erkennen?

Von [Felicitas Wilke](#)

18. Oktober 2018, 20:05 Uhr

<https://www.zeit.de/arbeit/2018-10/bewerbungsroboter-kuenstliche-intelligenz-amazon-frauen-diskriminierung>



***** Hinweis: Amazon gibt keine Angaben zur Geschlechterverteilung seiner technischen Arbeitskräfte bekannt.**

<https://www.cut-the-saas.com/ai/case-study-how-amazons-ai-recruiting-tool-learns-gender-bias>

KI bewertete Lebensläufe von Frauen systematisch schlechter! Bestrafte z.B. Begriffe wie “Women’s Chess Club Captain” oder “Women’s College”

→ KI wurde überwiegend auf Lebensläufe von Männern trainiert

2020



Newsletters Contact Press Deutsch

DONATE

BECOME A SUPPORTER

ABOUT / PROJECTS / PUBLICATIONS / STORIES / POSITIONS / BLOG



/ Austria's employment agency rolls out discriminatory algorithm, sees no problem

by *Nicolas Kayser-Bril*

AMS, Austria's employment agency, is about to roll out a sorting algorithm that gives lower scores to women and to the disabled. It is very likely illegal under current anti-discrimination law.

<https://algorithmwatch.org/en/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/>

Der AMS-Algorithmus war strukturell diskriminierend!

→ Frauen, ältere Menschen und Personen ohne österreichische Staatsbürgerschaft erhielten deutlich niedrigere Wahrscheinlichkeiten, innerhalb von sechs Monaten eine Arbeit zu finden.

BE_INT

= f (0,10

– 0,14 x GESCHLECHT_WEIBLICH
– 0,13 x ALTERSGRUPPE_30_49
– 0,70 x ALTERSGRUPPE_50_PLUS
+ 0,16 x STAATENGRUPPE_EU
– 0,05 x STAATENGRUPPE_DRITT
+ 0,28 x AUSBILDUNG_LEHRE
+ 0,01 x AUSBILDUNG_MATURA_PLUS
– 0,15 x BETREUUNGSPFLICHTIG
– 0,34 x RGS_TYP_2
– 0,18 x RGS_TYP_3
– 0,83 x RGS_TYP_4
– 0,82 x RGS_TYP_5
– 0,67 x BEEINTRÄCHTIGT
+ 0,17 x BERUFSGRUPPE_PRODUKTION
– 0,74 x BESCHÄFTIGUNGSTAGE_WENIG
+ 0,65 x FREQUENZ_GESCHÄFTSFALL_1
+ 1,19 x FREQUENZ_GESCHÄFTSFALL_2
+ 1,98 x FREQUENZ_GESCHÄFTSFALL_3_PLUS
– 0,80 x GESCHÄFTSFALL_LANG
– 0,57 x MN_TEILNAHME_1
– 0,21 x MN_TEILNAHME_2

An excerpt from the AMS algorithm's documentation.

2021

MIT
Technology
Review

Featured Topics Newsletters Events Audio

SIGN IN

SUBSCRIBE

ARTIFICIAL INTELLIGENCE

LinkedIn's job-matching AI was biased. The company's solution? More AI.

ZipRecruiter, CareerBuilder, LinkedIn—most of the world's biggest job search sites use AI to match people with job openings. But the algorithms don't always play fair.

By Sheridan Wall & Hilke Schellmann

June 23, 2021

<https://www.technologyreview.com/2021/06/23/1026825/linkedin-ai-bias-ziprecruiter-monster-artificial-intelligence/>

Männer und Frauen bekamen unterschiedliche Jobvorschläge

- Männer klickten und bewarben sich häufiger auf Führungsjobs und der Algorithmus lernte, diese eher Männern zu zeigen.
- LinkedIn optimierte für Klickwahrscheinlichkeit - nicht für gleiche Chancen.

Years ago, LinkedIn discovered that the recommendation algorithms it uses to match job candidates with opportunities were producing biased results. The algorithms were ranking candidates partly on the basis of how likely they were to apply for a position or respond to a recruiter. The system wound up referring more men than women for open roles simply because men are often more aggressive at seeking out new opportunities.



2016



Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions):

- Entwickelt in den USA
- Risikobewertungssoftware zur Einstufung des Risikos eines Angeklagten, erneut straffällig zu werden.
- Verwendet von Richter*innen für Entscheidungen über Kaution, Bewährung und Strafmaß.

→ Test auf Genauigkeit und Vorurteile gegenüber bestimmten ethnischen Gruppen durch ProPublica

Weitere Beispiele...

For example, translating Turkish, which is a gender-neutral language, can prove difficult.

Détecter la langue **Turc** Français Anglais ▼

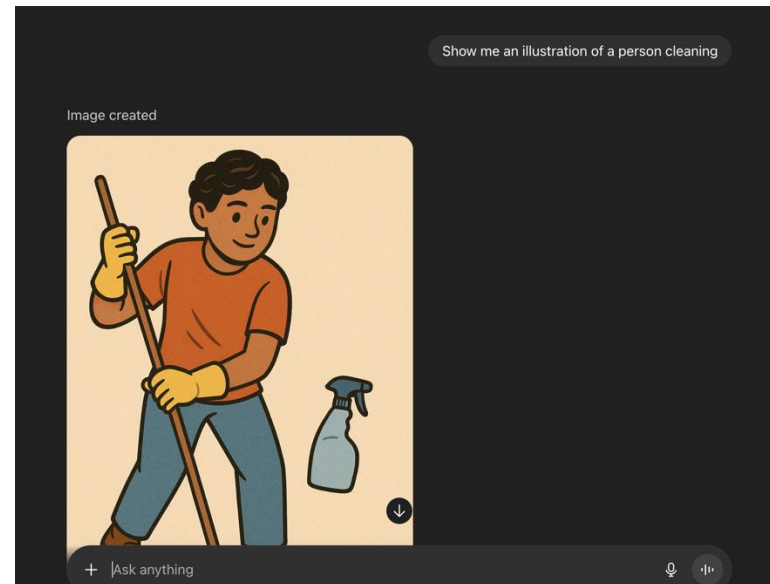
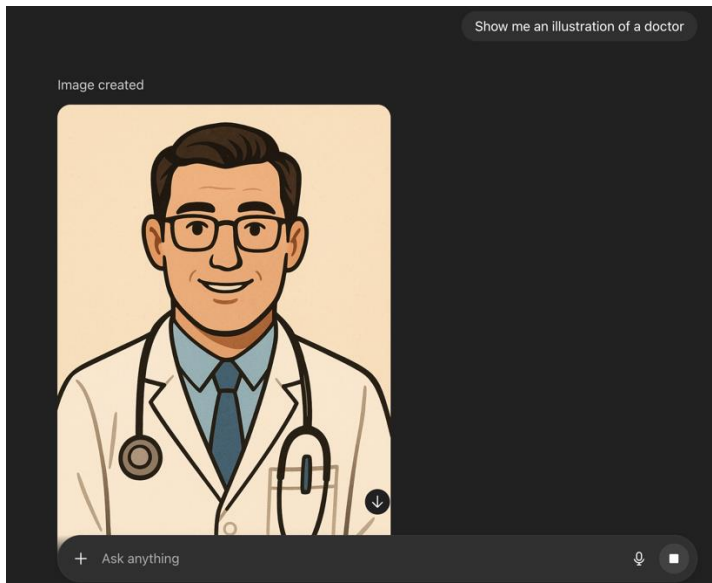
↔ Allemand Gaélique (Écosse) **Anglais** ▼

O bir doktor ve o bir hemşire.

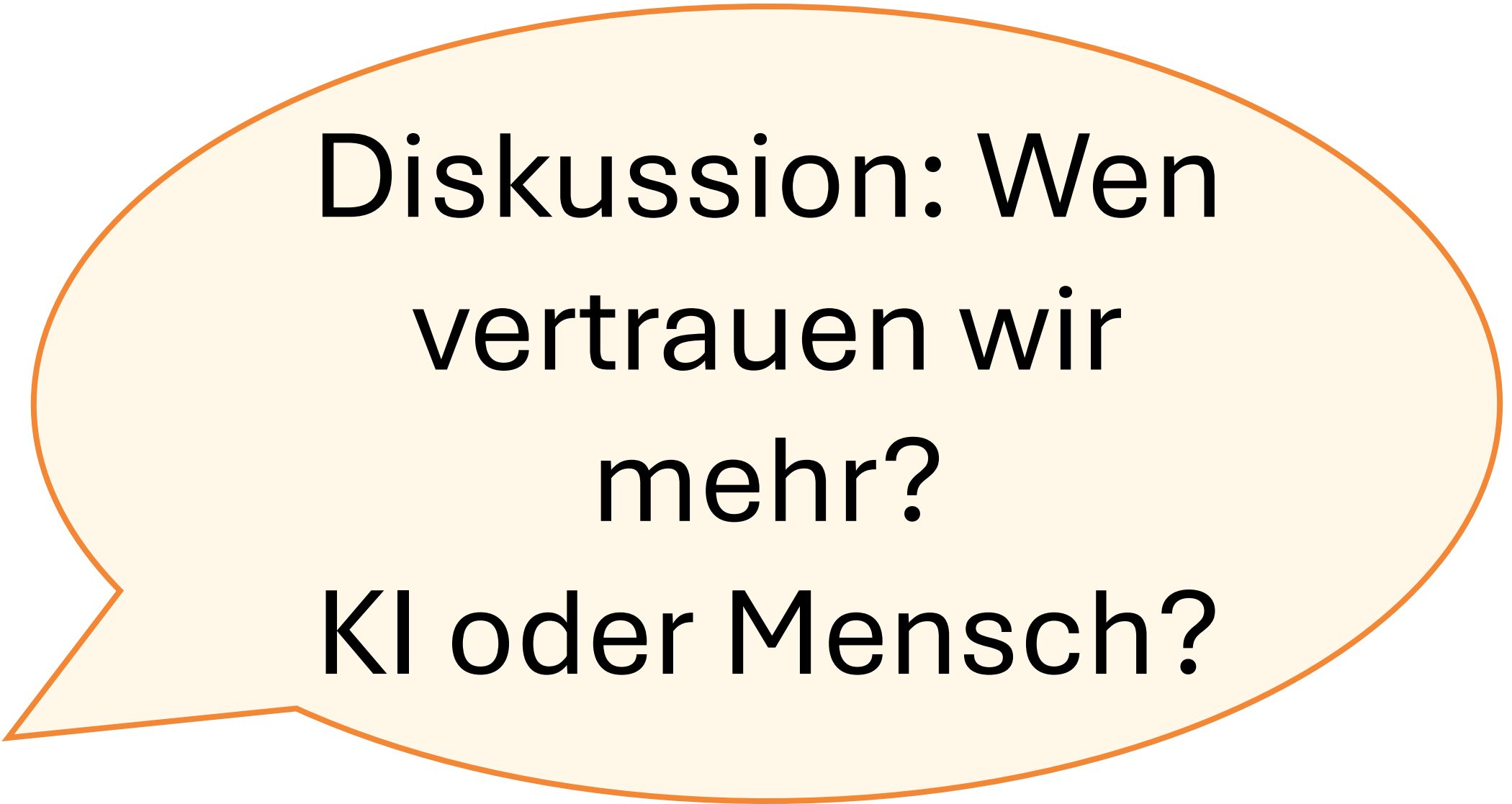
×

He is a doctor and she is a nurse.

<https://knowledge-centre-translation-interpretation.ec.europa.eu/en/content/data-discrimination-when-words-carry-prejudice>



Eigene Graphiken



**Diskussion: Wen
vertrauen wir
mehr?
KI oder Mensch?**

Verzerrungen in Algorithmen beheben?

EMERGING TECHNOLOGIES

Why AI bias may be easier to fix than humanity's

Jun 30, 2023



<https://www.weforum.org/stories/2023/06/why-ai-bias-may-be-easier-to-fix-than-humanity-s/>

“Die Erkenntnis, dass KI dazu neigt, Ungleichheiten zu perpetuieren, kann uns im Kampf für Fairness einen Vorteil verschaffen. Letztendlich wäre es zweifellos einfacher, die Vorurteile der KI zu mildern, als die von Menschen perpetuierten Vorurteile zu beseitigen.

Das liegt daran, dass **mangelnde Fairness in der KI systematisiert und quantifiziert werden kann, wodurch sie transparenter ist als menschliche Entscheidungen**, die oft von unbewussten Vorurteilen und Mythen geprägt sind. KI schafft keine Vorurteile. Vielmehr dient sie als Spiegel, um Beispiele dafür aufzudecken – und **es ist einfacher, etwas zu stoppen, das man sehen und messen kann.**”

Regierungen und Unternehmen müssen die Fairness der KI zu einer Priorität machen! Dafür braucht es strenge Regulierungen!

EU AI Act

 EU Artificial Intelligence Act

[Das Gesetz](#) ▾ [Umsetzung](#) ▾ [Einhaltung der Vorschriften](#) ▾ [Über uns](#) [Newsletter](#) ↗  DE ▾

Zusammenfassung des AI-Gesetzes auf hoher Ebene

27 Feb, 2024

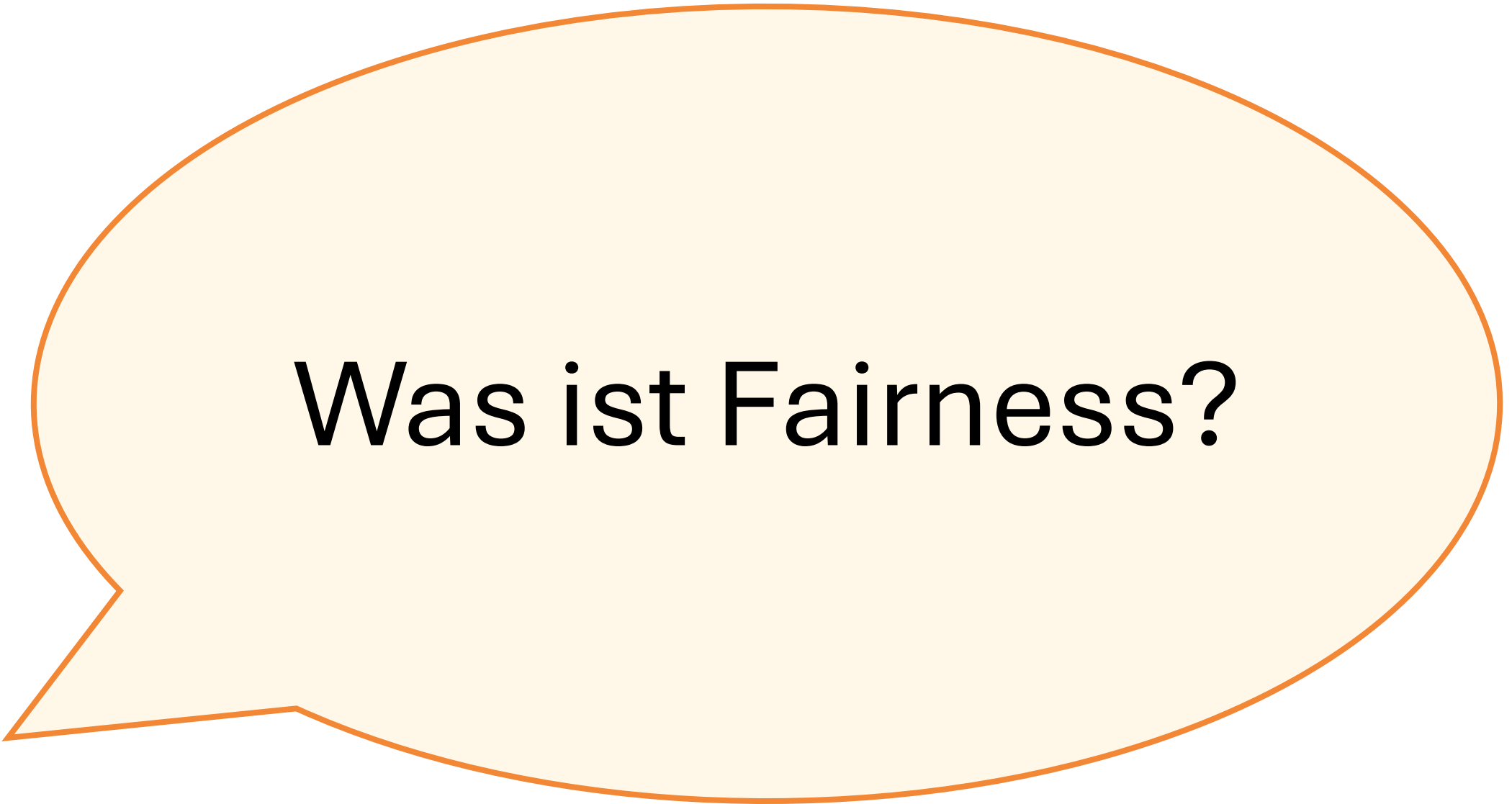
Aktualisiert am 30. Mai in Übereinstimmung mit der berichtigten Fassung des AI-Gesetzes.

In diesem Artikel geben wir Ihnen eine Zusammenfassung des Gesetzes über künstliche Intelligenz, wobei wir die Teile ausgewählt haben, die für Sie am ehesten von Bedeutung sind, unabhängig davon, wer Sie sind. Dort, wo es relevant ist, geben wir Links zum Originaldokument an, damit Sie jederzeit auf den Gesetzestext zugreifen können.

<https://artificialintelligenceact.eu/de/high-level-summary/>

“Der AI Act gehört zum Bereich der „Produktregulierung“. Er **regelt, welche KI-Systeme in der Europäischen Union (EU) erlaubt und welche verboten ist.** Die Verordnung stellt einheitliche Regeln für alle 27 Mitgliedstaaten auf – für den Verkauf und die Verwendung von KI-Systemen in der EU.

<https://www.mpg.de/24096506/faq-was-regelt-der-ai-act>



Was ist Fairness?

„Im Zusammenhang mit der Entscheidungsfindung bedeutet Fairness das Fehlen jeglicher Vorurteile oder Bevorzugungen gegenüber einer Person oder einer Gruppe aufgrund ihrer angeborenen oder erworbenen Eigenschaften.“ (Mehrabi et al. 2019)

Fairness = Abwesenheit von Verzerrung („Bias“)

Wo können Verzerrungen entstehen?

1. Gesellschaftlich vorhanden

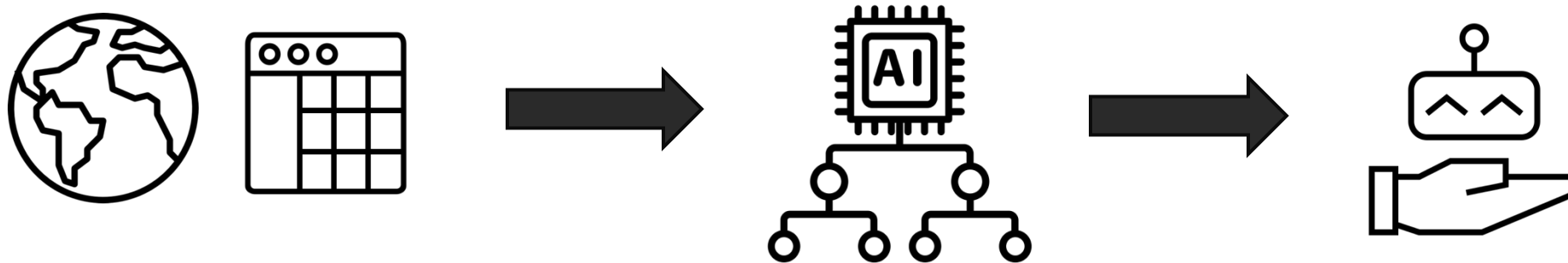
- existieren unabhängig vom Algorithmus
- haben ihren Ursprung in gesellschaftlichen Strukturen (z. B. Diskriminierung, ungleiche Chancen)

3. Kontextabhängig neu erzeugt

- entstehen durch den Nutzungskontext oder die Anwendung der KI (z. B. falsche Interpretation, unfaire Entscheidungspraxis)

2. Technisch bedingt

- entstehen oder werden verstärkt durch die technischen Eigenschaften des Algorithmus oder der Datenverarbeitung (z. B. unbalancierte Trainingsdaten, Modellarchitektur)



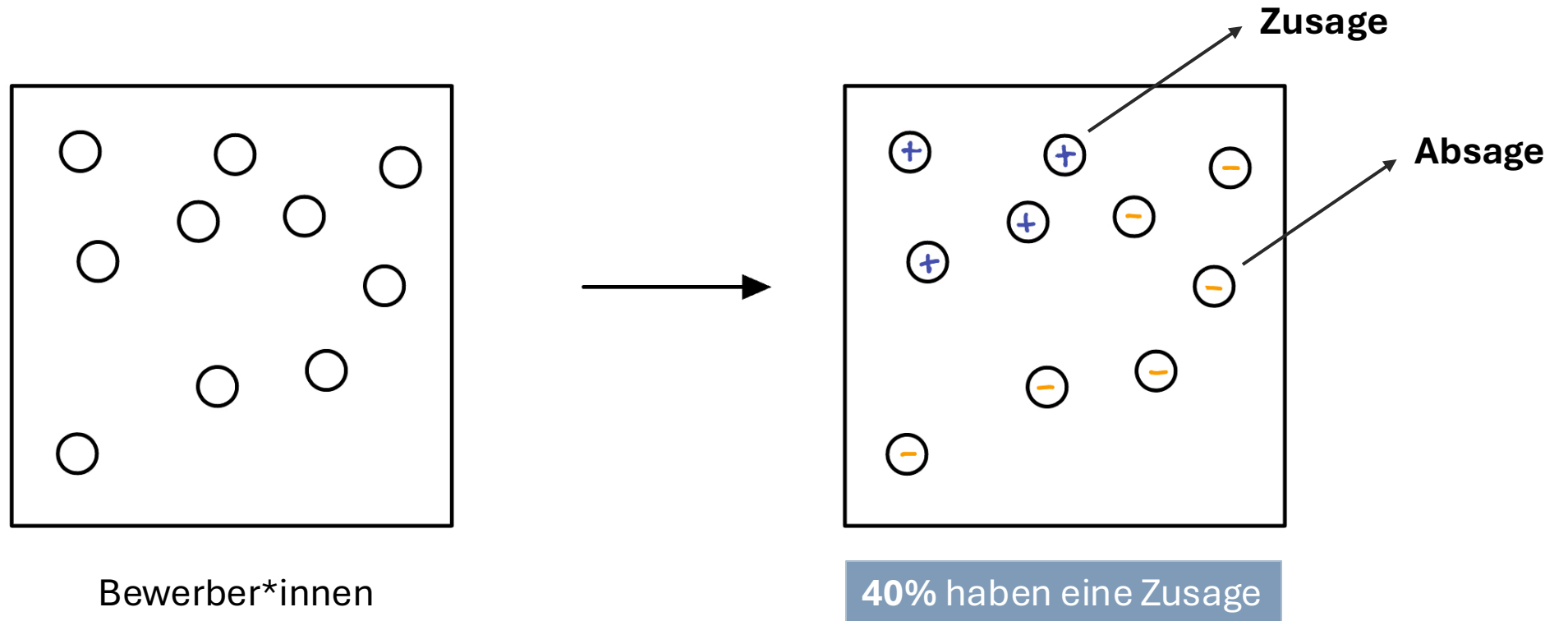
Friedman & Nissenbaum 1996



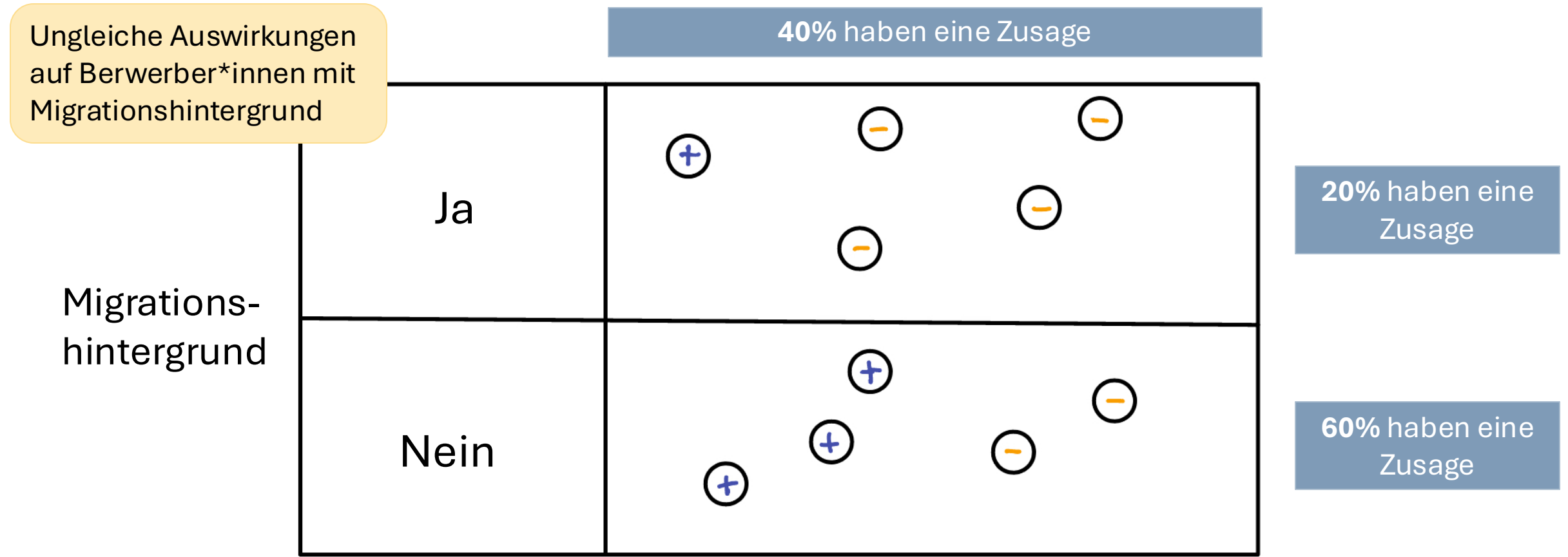
Wie misst man
Fairness?

Fairness-Messung: Beispiel

Beispiel: Die Qualifikationschancen von Bewerber*innen werden von einer KI vorhergesagt. Wenn die Chancen **hoch** sind, erhalten sie eine **Zusage**, andernfalls eine **Absage**.

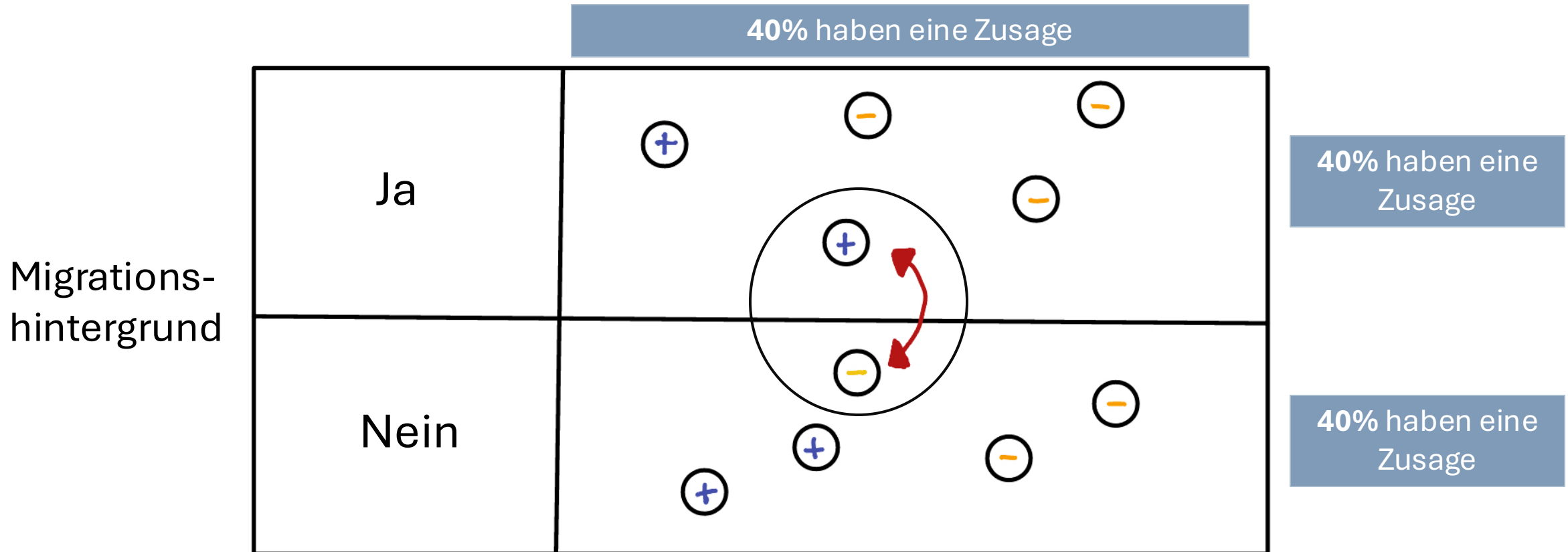


Fairness-Messung: Beispiel



Fairness-Messung: Beispiel

Statistische Parität (ein beliebtes Maß für Gruppenfairness): Alle Gruppen haben die gleiche Wahrscheinlichkeit auf eine positive Entscheidung (z. B. Zusage, Kreditbewilligung etc.).



Fairness-Messung: Beispiel

Das Entfernen des Merkmals „Migrationshintergrund“ aus dem Algorithmus verhindert **keine** Diskriminierung!

Postleizahl

		Postleizahl	
		0123	4567
Migrations- hintergrund	Ja		   
	Nein	  	 

20% haben eine Zusage

60% haben eine Zusage

Fairness-Messung: Beispiel

Statistische Parität:

- 1. Kann den Zielen der Entscheidungsträger*innen widersprechen.
- 2. Verletzt **individuelle Fairness**: Zwei Personen, die sich ähneln, sollten ähnliche Ergebnisse erhalten.

Qualifiziert

Migrations- hintergrund	Ja	Nein
	Ja	Nein
Ja		   
Nein	  	 

40% haben eine
Zusage

40% haben eine
Zusage

Zwei Vorstellungen von Fairness

Individuelle Fairness



Gleichberechtigung

Gruppenfairness



Gleichstellung

Weitere Fairness-Messungen (für Gruppenfairness)

Migrations-
hintergrund

		Qualifiziert	
		Ja	Nein
Ja	Ja	+ - - +	-
	Nein	+ +	- -

G1

Qualifiziert	Zusage Ja	Zusage Nein
Ja	2	2
Nein	0	1

G2

Qualifiziert	Zusage Ja	Zusage Nein
Ja	2	0
Nein	0	3

Weitere Fairness-Messungen (für Gruppenfairness)

G1

Qualifiziert	Zusage Ja	Zusage Nein
Ja	2	2
Nein	0	1

Ein Individuum möchte, dass alle mit gleicher Qualifikation die gleiche Chance auf eine Zusage haben – unabhängig von der Gruppenzugehörigkeit (Equal Opportunity).

Gruppe 1: Trefferquote = 2 von 4 also 50%

Gruppe 2: Trefferquote = 2 von 2 also 100%

→ Equal Opportunity = 50%

G2

Qualifiziert	Zusage Ja	Zusage Nein
Ja	2	0
Nein	0	3

Ein Entscheider möchte, dass alle Gruppen gleich behandelt werden – also gleiche Fehler- und Trefferquoten haben (Equalized Odds).

Gruppe 1:

Trefferquote = 2 von 4 also 50%

Fehlerrate = 0 von 1 also 0 %

Gruppe 2:

Trefferquote = 2 von 2 also 100%

Fehlerrate = 0 von 3 also 0%

→ Equalized Odds verletzt bei Trefferquote

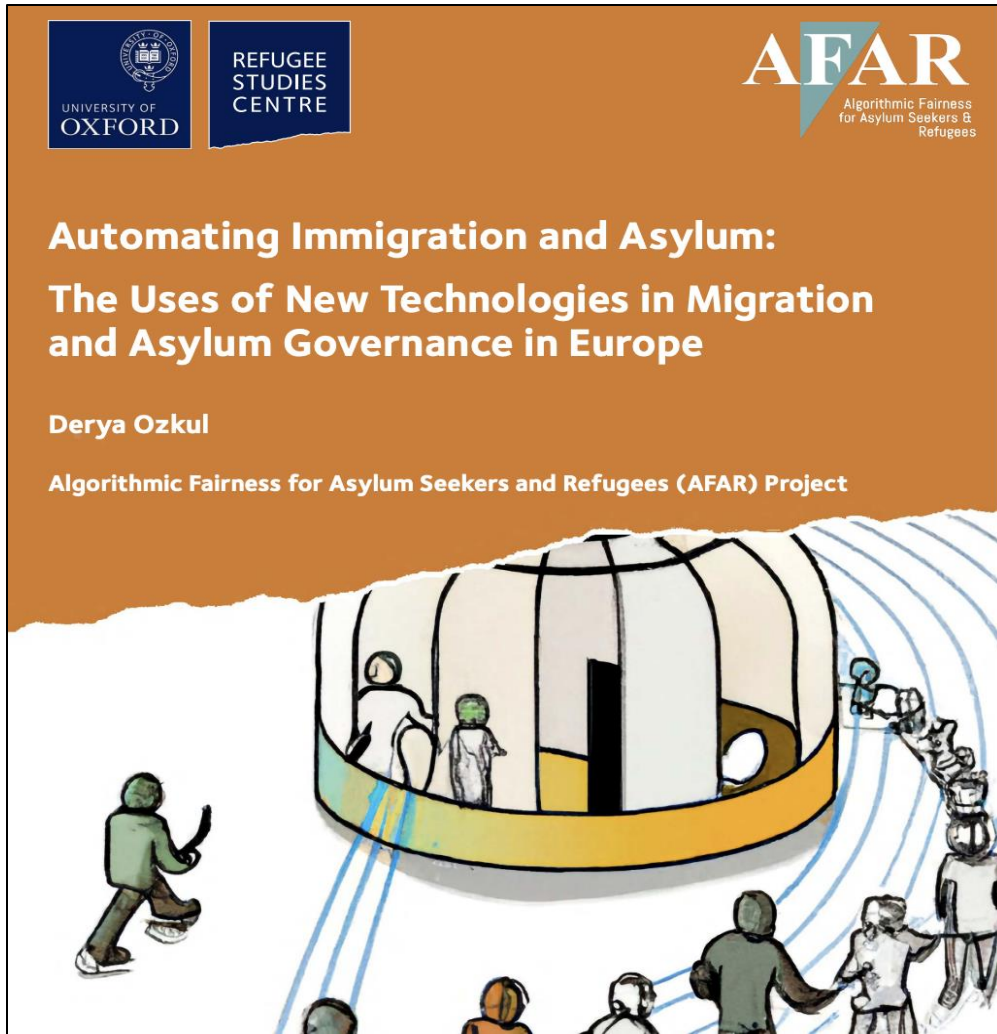
Fazit

- “Fairness” ist nicht eindeutig definiert
→ Unterschiedliche Verständnisse je nach Kontext
- Es gibt verschiedene Fairness-Metriken
→ Diese Metriken können sich gegenseitig widersprechen
- Fairness braucht aktive Gestaltung
→ Verzerrungen (vor allem in den Daten) erkennen, messen und gezielt korrigieren!



KI in Migration & Asyl?

KI in Migration & Asyl



https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf

Vor Ankunft:

- Frühwarn-und Prognosemodellen
- Automatisierte Bearbeitung von Visas
- ...

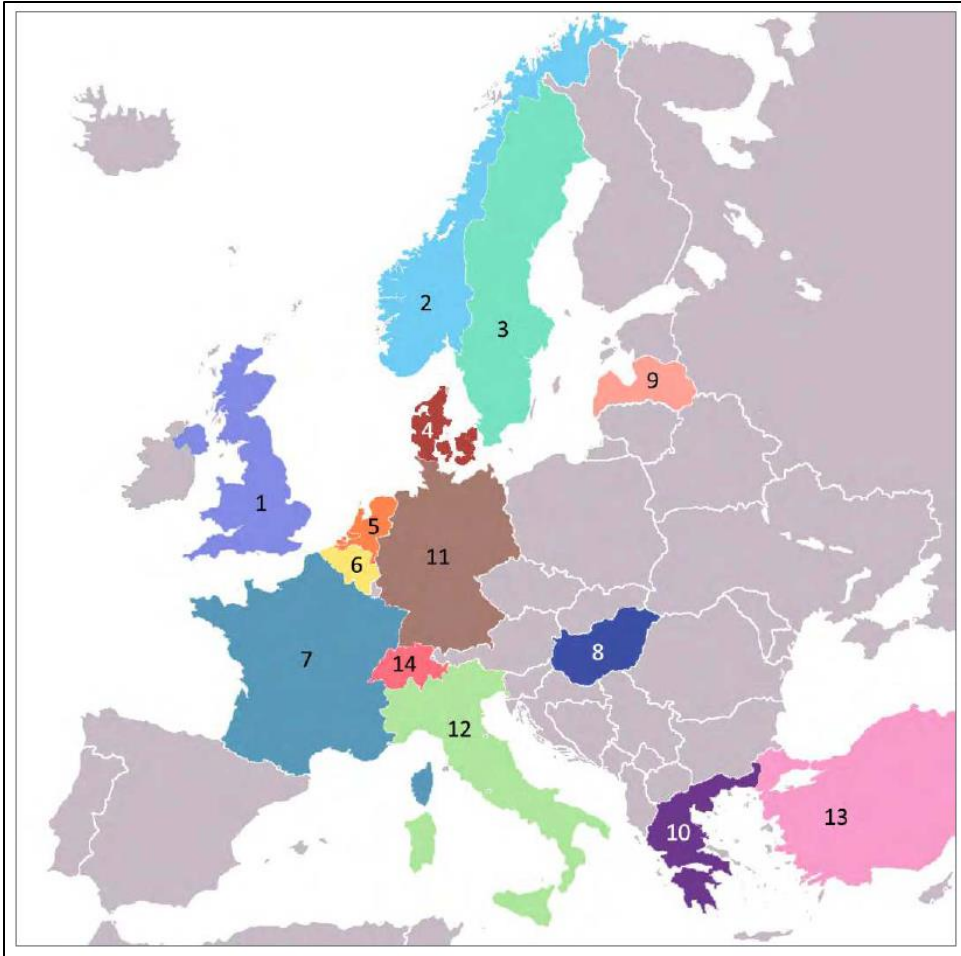
Bei Ankunft:

- Europäisches Grenzüberwachungssystem
- Dokumentenprüfungs-Technologien
- ...

Nach Ankunft:

- Sprach- und Dialekterkennung im Asylverfahren
- **Algorithmische Verteilungssysteme**
- ...

KI in Migration & Asyl



- 1 UK:** risk assessment for the processing of visitor visa applications (halted); risk assessment of applications for marriages; categorisation of applications for the EU Settlement Scheme; identification and prioritisation of irregular migrants; electronic monitoring; mobile phone data extraction
- 2 Sweden:** partly automated processing of residency applications; processing of citizenship applications
- 3 Norway:** processing of residency applications for family migration of skilled/posted workers; processing of citizenship applications; mobile phone data extraction; distribution of welfare benefits to asylum seekers; matching tool for allocation of reception centres
- 4 Denmark:** mobile phone data extraction
- 5 Netherlands:** screening of employment sponsorship (currently under revision); document verification; assessment of appeal cases' type and complexity (under development); mobile phone data extraction; matching tool for screening similar asylum applications; matching tool for settlement (under development for testing)
- 6 Belgium:** document verification (under development)
- 7 France:** document verification (under development)
- 8 Hungary:** lie detection (tested)
- 9 Latvia:** lie detection (tested); speech recognition to help applicants with citizenship applications
- 10 Greece:** lie detection (tested)
- 11 Germany:** name transliteration; dialect recognition; mobile phone data extraction; matching tool for settlement (under development)
- 12 Italy:** speech-to-text technology for the transcription of interviews with asylum seekers
- 13 Turkey:** speech and dialect recognition (tested)
- 14 Switzerland:** matching tool for settlement (tested)

https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf

(KI-gestützte) Verteilung von Schutzsuchenden

mit



Location matching on shaky grounds: Re-evaluating algorithms for refugee allocation

Clara Strasser Ceballos
LMU
Munich, Germany
Clara.StrasserCeballos@stat.uni-muenchen.de

Christoph Kern
LMU
Munich, Germany
Munich Center for Machine Learning (MCML)
Munich, Germany
christoph.kern@stat.uni-muenchen.de

Abstract

The initial location to which refugees are assigned upon arrival in a host country plays a key role in their integration. Several research groups have developed tools to optimize refugee-location matching, with the overall aim of improving refugees' integration outcomes. Four primary tools are already being piloted across various countries: GeoMatch, Annie™ Moore, Match'In, and Re:Match. The first two tools combine supervised machine learning with optimal matching techniques, while the latter two rely on heuristic methods to match refugee preferences with suitable locations. These tools are used in a highly sensitive context and directly impact human lives. It is, therefore, not only desirable but critical to (re-)evaluate them through the lens of algorithmic fairness. We contribute in three key aspects: First, we provide a comprehensive overview and systematization of the tools aimed at the algorithmic fairness community. Second, we identify sources of biases along the tool design stages that can contribute to disparate impacts downstream. Finally, we simulate the application of the GeoMatch tool using German survey data to empirically illustrate the impact of target variable choice on matching outcomes. While GeoMatch optimizes economic integration, we demonstrate that the integration gains differ substantially when social integration is prioritized instead. With our use case, we highlight the susceptibility of algorithmic matching tools to design decisions such as the operationalization of the integration outcome and emphasize the need for more holistic evaluations of their social impacts.

CCS Concepts

• Computing methodologies → Machine learning; • Applied computing → Law, social and behavioral sciences.

1 Introduction

Wars, violence, political persecution, famine, and other destructive events force people to flee their homes in search of refuge in new host societies. According to the United Nations High Commissioner for Refugees (UNHCR), around 2 million refugees needed resettlement in 2023 [119]. However, less than 5% of those in need successfully resettle to third countries [46, 119]. Upon arrival in the host countries, the locations to which refugees are eventually assigned depend on the legal and administrative framework in place, with decisions being largely influenced by location-specific constraints [99]. In some countries like Switzerland and Germany, location allocations are even made (quasi-)randomly [12, 40]. As a result, most allocation processes fail to incorporate refugees' characteristics and preferences [5]. Yet, research has shown that initial placement plays a critical role in the integration outcomes for refugees [5, 8]. To address this gap, several research groups have developed tools that take refugees' characteristics and preferences into account when allocating refugees, with the aim of improving their integration outcomes. These tools are: GeoMatch, Annie™ Moore, Match'In and Re:Match¹ [5, 13, 103, 112].

The first two tools, GeoMatch and Annie™ Moore, assign refugees to locations based on predictions of integration outcomes [5, 13]. In short, the tools train machine learning models on historical data containing information on refugee characteristics, assigned locations, and specific measures of integration (e.g., employment status). Integration predictions are generated for newly arriving refugees. These refugees are assigned to locations that maximize an optimality criterion (e.g., global average employment) subject to constraints (e.g., location capacity). These tools are currently being piloted by two U.S. resettlement agencies, Global Refugee, and

1

Wie funktionieren KI-gestützte Verteilungssysteme?

2

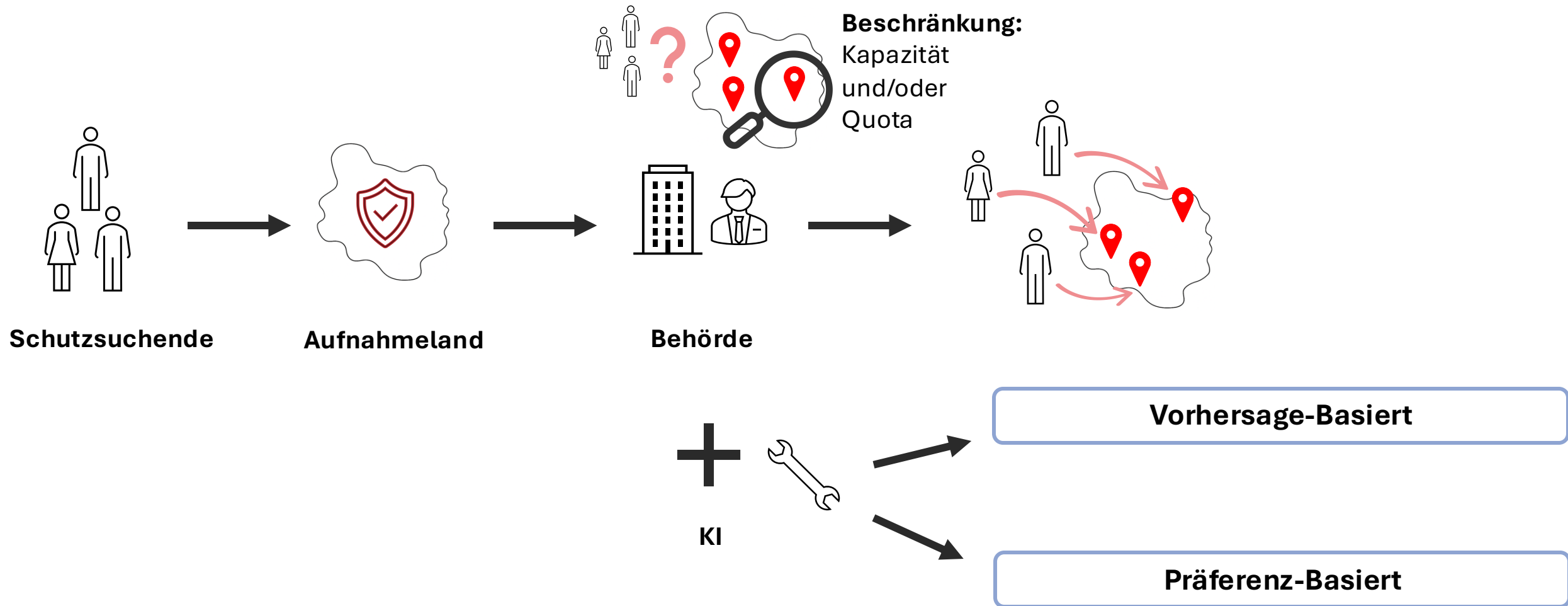
Welche KI-gestützte Verteilungssysteme existieren?

3

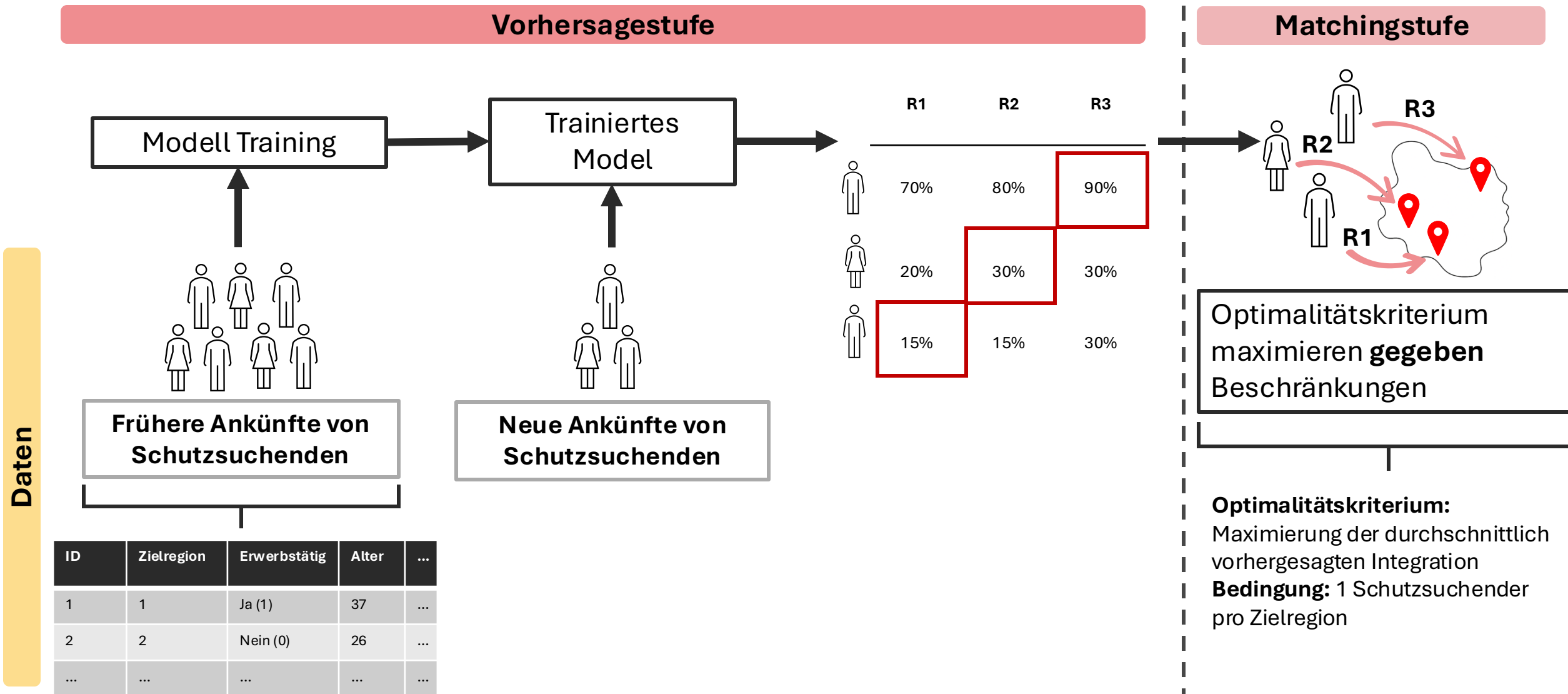
Was sind mögliche Quellen für Fehler und Verzerrungen?

<https://dl.acm.org/doi/full/10.1145/3715275.3732149>

(KI-gestützte) Verteilung von Schutzsuchenden



Vorhersage-Basierte KI Systeme



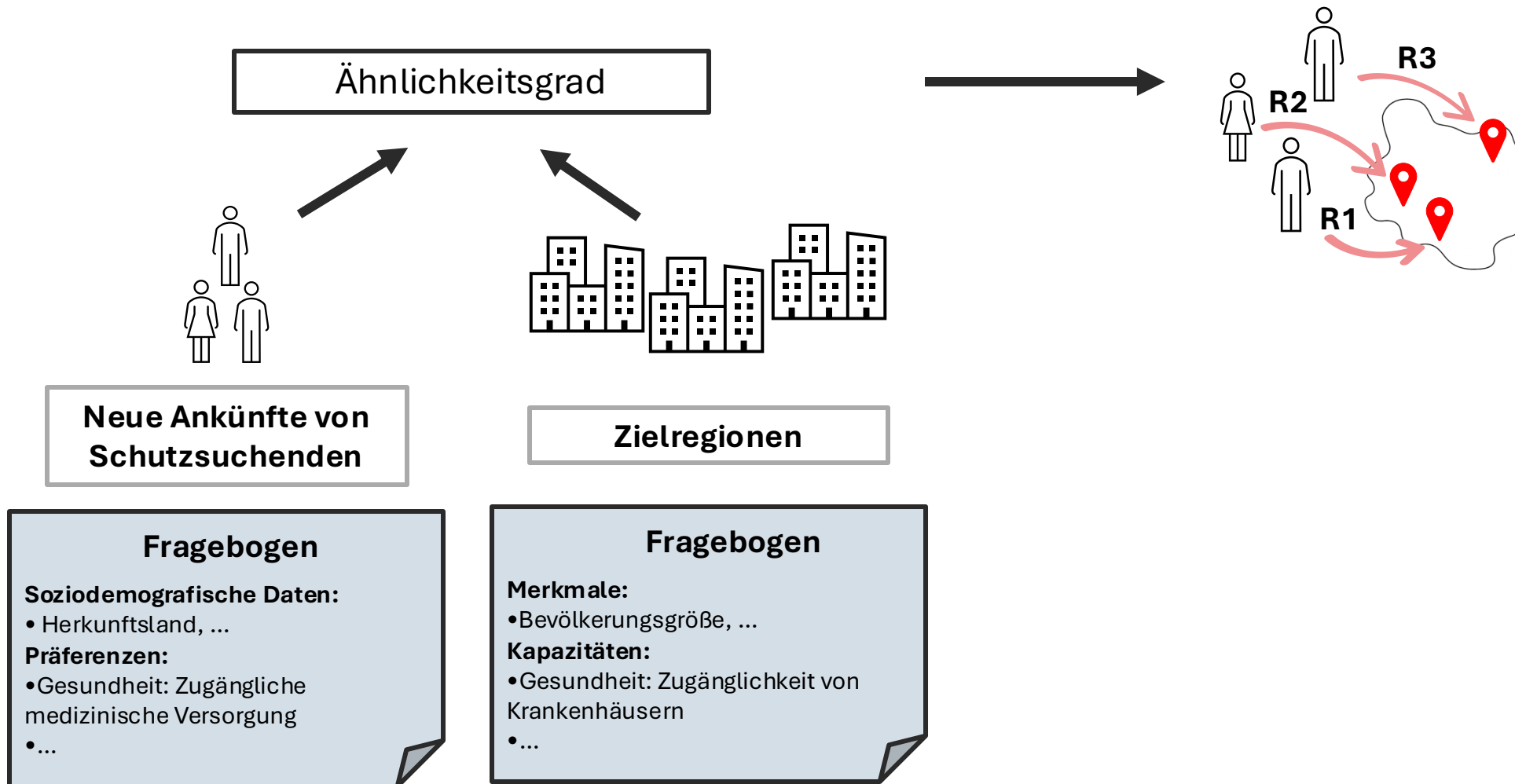
Vorhersage-Basierte KI Systeme

	Vorhersage-Basiert	
Systeme	GeoMatch	Annie™ MOORE
Studie:	Bansak et al. (2018)	Ahani et al. (2021)
Entwickelt von:	Universität Stanford & ETH Zürich	Universität Oxford, Lund & Worcester
Pilotprojekte:	2020: Schweiz Seit 2023: USA (Global Refuge) Seit 2024: Niederlanden Ab 2025: Kanada	Seit 2018: USA (HIAS)

Präferenz-Basierte KI Systeme

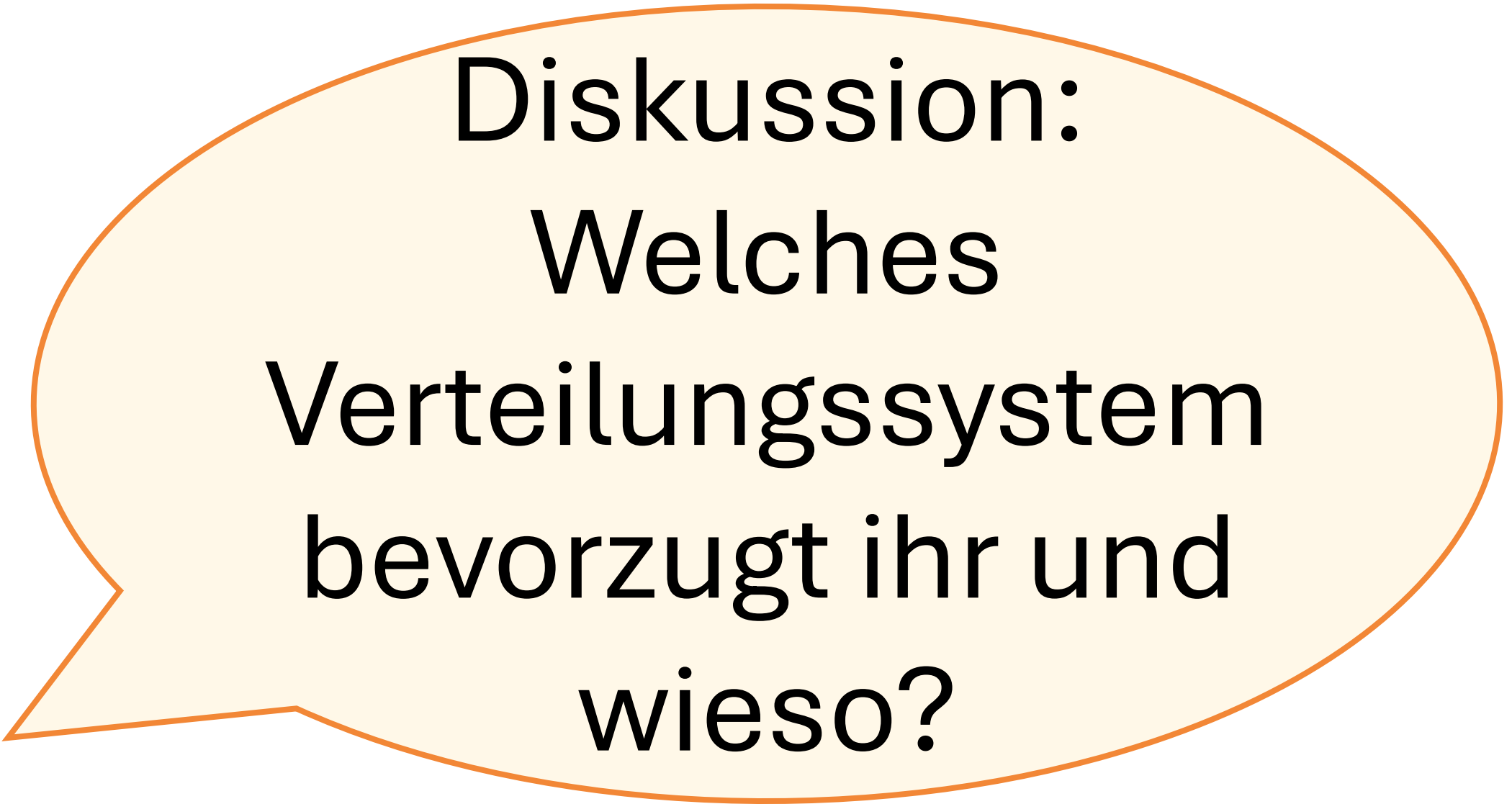
Matchingstufe

Daten



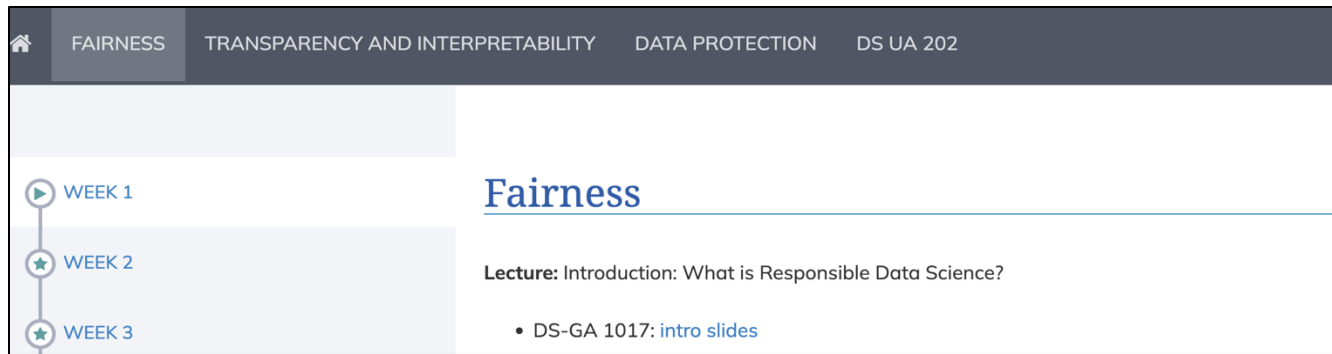
Präferenz-Basierte KI Systeme

	Präferenz-Basiert		
Systeme	Match'In	Re:Match	RUTH
Studie:	Sauer et al. (2024)	Smith et al. (2024)	Farajzadeh et al. (2023)
Entwickelt von:	Universität Hildesheim & FAU	Berlin Governance Platform & Pairity	Universität Oxford & Worchester
Pilotprojekte:	2023-2024: Deutschland	2022-2024: Deutschland	Seit 2022: USA (HIAS)

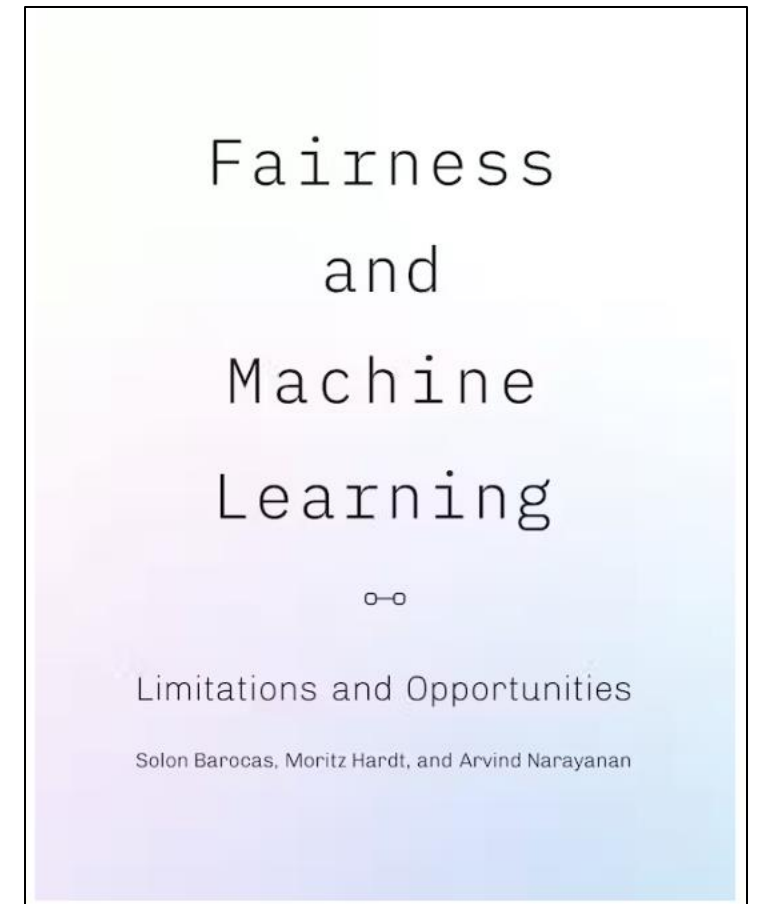
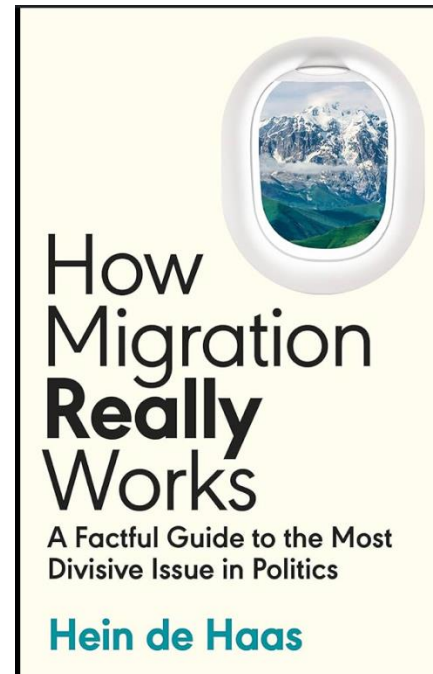
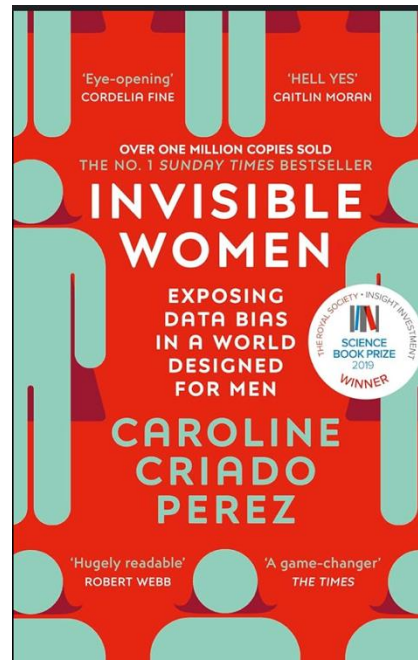
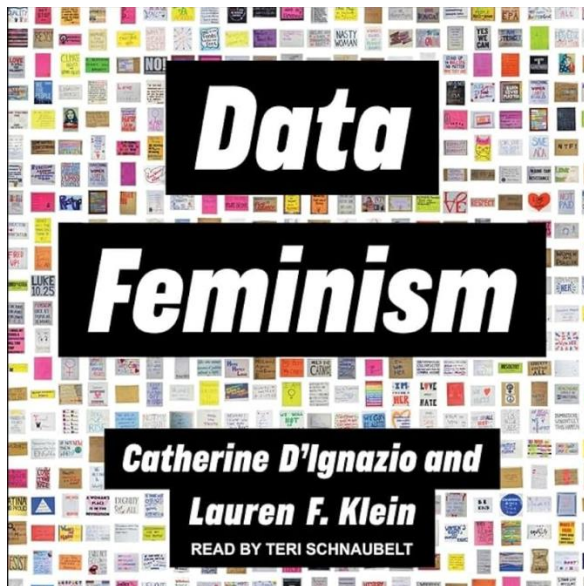


**Diskussion:
Welches
Verteilungssystem
bevorzugt ihr und
wieso?**

Literatur-/Kursempfehlung

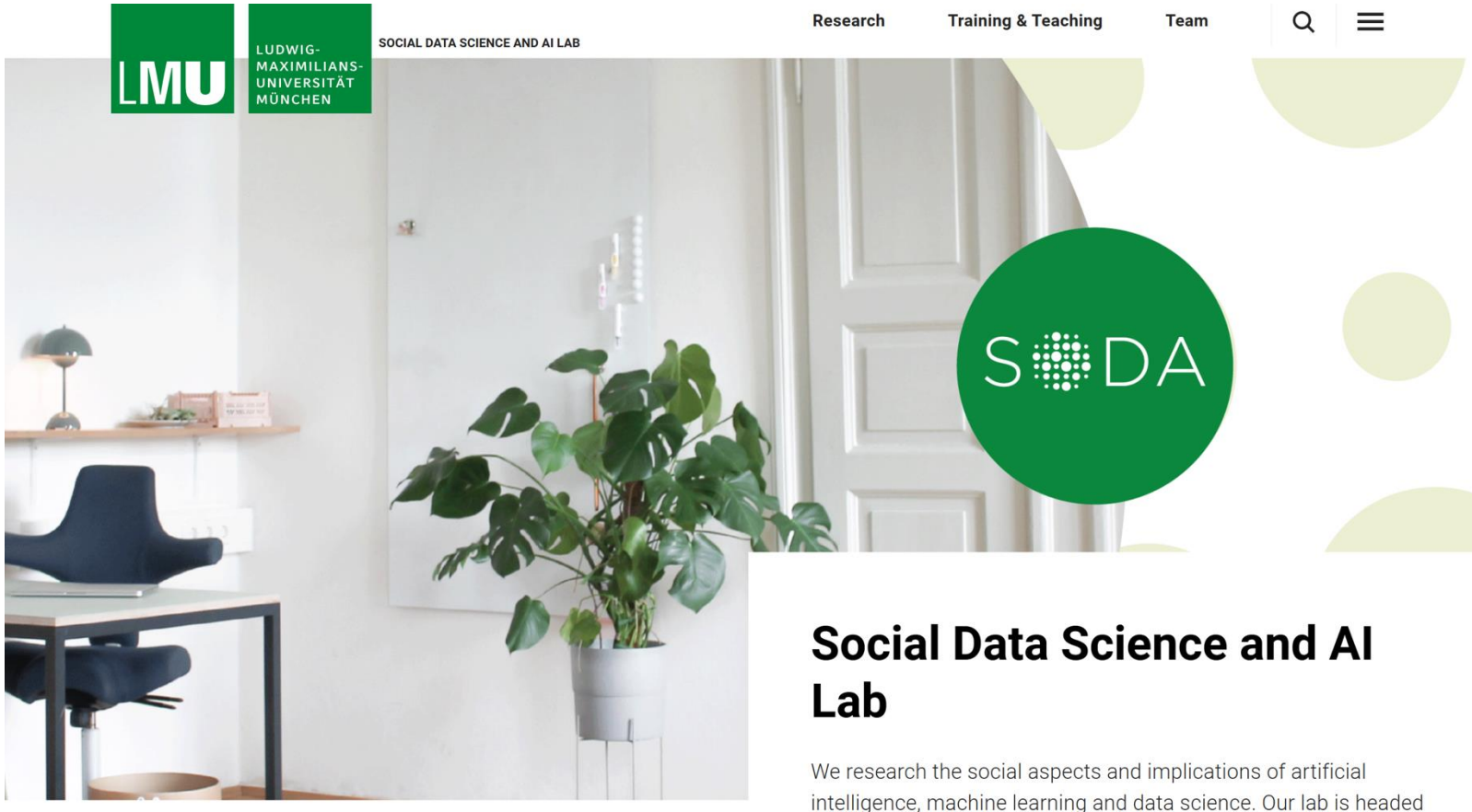


<https://dataresponsibly.github.io/rds25/modules/fairness/week1/>



<https://fairmlbook.org/pdf/fairmlbook.pdf>

Danke!



**Möchten Sie mehr
über unsere
Forschung erfahren?
Besuchen Sie uns
oder unsere
Website!
<https://www.stat.lmu.de/soda/en/>**