

Digitale Diskriminierung?

Der AMS-Algorithmus in Österreich

Mag.^a Paola Lopez
Universität Wien / WZB Berlin
Wien, 24. September 2020

Überblick

- **Einleitung**
 - Diskriminierung, Daten und Algorithmen
- **Fallbeispiel: Der AMS-Algorithmus in Österreich**
 - Wie funktionieren solche Systeme?
 - Welche Probleme können auftreten?
 - Wie könnte man diese Probleme reparieren?
- **Fazit**

Diskriminierung und Objektivität

- **Unconscious Bias**
 - Strukturelle Ungleichheiten bündeln sich in individuellem Bias
 - Fortschreibung durch Menschen in Entscheidungspositionen
 - Diskriminierung
- **Lösungsvorschlag:** Objektive statt subjektive Kriterien
 - Daten
 - Zahlen
 - Algorithmen

Vermeidung von Bias durch objektive Daten?

- **Beispiel: Recruiting-Algorithmus von Amazon**
 - Automatisierter Gender Bias
 - Erwähnung von „*women*“ im Lebenslauf bewirkt Punkteabzug
 - Algorithmus abgeschafft
- **Beispiel: Automatisierter Lebenslauf-Scanner**
 - „*Jared*“ und „*high school lacrosse*“ größte Performance-Indikatoren
 - System abgeschafft
- **Im Folgenden: Der AMS-Algorithmus in Österreich**

Der AMS-Algorithmus in Österreich

- **Arbeitsmarktservice Österreich (AMS)**
- **Herbst 2018:** Algorithmisches System zur Einteilung von Erwerbsarbeitslosen in drei Gruppen vorgestellt
- **Effizienz und Effektivität**
 - Beratungsprozess
 - Vergabe von Förderressourcen
- **Assistenzsystem**
- Einteilung entlang algorithmisch prognostizierter „**Arbeitsmarktchancen**“

Drei Gruppen

- „Wie hoch sind die Chancen, in der nächsten Zeit einen Job zu finden?“
- **Gruppe A – hohe Chancen**
 - Weniger Zugang zu Förderinstrumenten
- **Gruppe B – mittlere Chancen**
 - Am förderungswürdigsten
 - Voller Zugang zu Ressourcen
- **Gruppe C – niedrige Chancen**
 - Betreuung in externen Formaten: Beratungs- und Betreuungseinrichtungen
Neu (BBEN)

„Arbeitsmarktchancen-Modell“

```
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
```

- **Statistische Analyse** (mittels logistischer Regression)
- Herzstück: **Koeffizienten**
 - Ca. 90 Modelle
- **Inkludierte Merkmale:**
 - Geschlecht
 - Alter
 - Staatsangehörigkeit
 - Ausbildung
 - Betreuungspflichten*
 - Wohnort
 - Beeinträchtigung

„Arbeitsmarktchancen-Modell“

$$\begin{aligned} \text{BE_INT} &= f(0,10 \\ &- 0,14 \times \text{GESCHLECHT_WEIBLICH} \\ &- 0,13 \times \text{ALTERSGRUPPE_30_49} \\ &- 0,70 \times \text{ALTERSGRUPPE_50_PLUS} \\ &+ 0,16 \times \text{STAATENGRUPPE_EU} \\ &- 0,05 \times \text{STAATENGRUPPE_DRITT} \\ &+ 0,28 \times \text{AUSBILDUNG_LEHRE} \\ &+ 0,01 \times \text{AUSBILDUNG_MATURA_PLUS} \\ &- 0,15 \times \text{BETREUUNGSPFLICHTIG} \\ &- 0,34 \times \text{RGS_TYP_2} \\ &- 0,18 \times \text{RGS_TYP_3} \\ &- 0,83 \times \text{RGS_TYP_4} \\ &- 0,82 \times \text{RGS_TYP_5} \\ &- 0,67 \times \text{BEEINTRÄCHTIGT} \end{aligned}$$

- **Statistische Analyse** (mittels logistischer Regression)
- Herzstück: **Koeffizienten**
 - Ca. 90 Modelle
- **Inkludierte Merkmale:**
 - **Geschlecht**
 - **Alter**
 - **Staatsangehörigkeit**
 - **Ausbildung**
 - **Betreuungspflichten***
 - **Wohnort**
 - **Beeinträchtigung**

Datenbasierte Systeme

- **Daten** sind das Fundament datenbasierter Systeme
- **Mathematische „Substanz“**
- **Problemfelder**
 - Was wird in Daten gemessen? Was bleibt unsichtbar?
 - Welche Kategorien gibt es?
 - Was ist überhaupt messbar? („*Soft skills*“ versus „*hard facts*“)
 - Heterogene Sachlagen wirken homogen
 - Ex-post Perspektive: Daten kennen keine Zukunft
 - „Prognose“ ist nur eine Bezeichnung

AMS-Algorithmus: Was bedeuten die Minuszeichen?

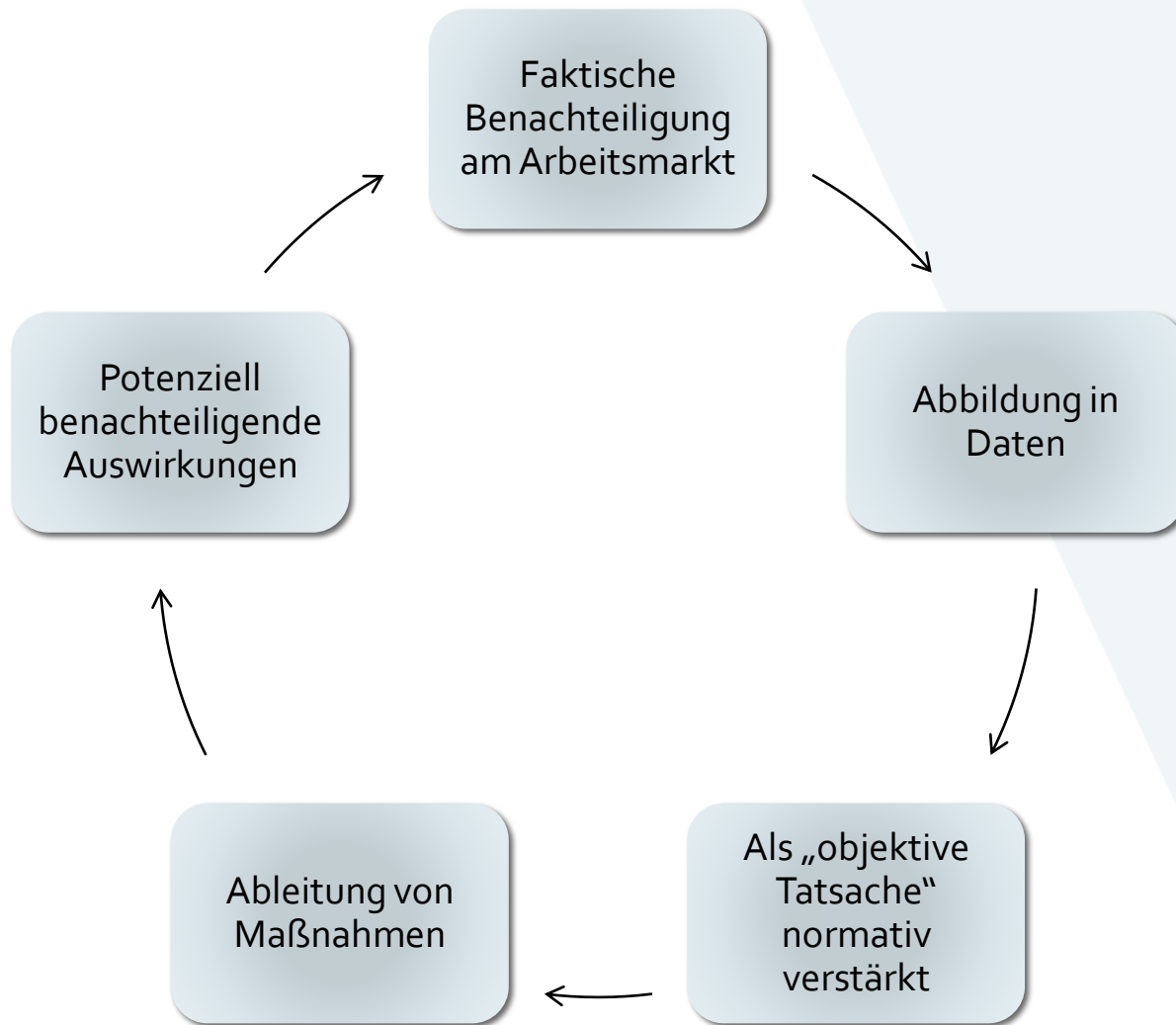
- **Analyse von Daten aus der Vergangenheit**
 - AMS erhebt schon immer Daten
- *„Welche Menschen mit welchen Dateneinträgen wurden wie schnell und wie nachhaltig in den Arbeitsmarkt integriert?“*
- **Minuszeichen** bei bestimmten Merkmalen
 - Gebündelter Bias des österreichischen Arbeitsmarkts
- **Mathematisch:**
 - Statistischer „Beweis“ für die faktische Benachteiligung von Frauen

...und die Recruiting-Algorithmen?

- Amazon-Algorithmus: Mit **Mitarbeiterdaten** aus den letzten 10 Jahren „trainiert“
- Welche Mitarbeiter*innen waren im Unternehmen erfolgreich?
- Bilden den **vorhandenen Bias** ab
- **Nicht:** „Wer ist die fähigste Person?“
- **Eher:** „Welche Art von Person war in der Vergangenheit im Unternehmen erfolgreich?“
 - Jared, der in der High School Lacrosse gespielt hat (laut Lebenslaufscanner)
- „Der Computer macht, **was man ihm sagt** – aber es ist vielleicht nicht das, **was man von ihm wollte.**“ (Joseph Weizenbaum)

AMS-Algorithmus: Diskriminierung durch Minuszeichen?

- Die Minuszeichen sind **an sich nicht diskriminierend**
- Erst die **Koppelung an Förderressourcen** problematisch
- **Gruppe C:** Keine AMS-internen Förderinstrumente wegen „zu vieler Minuszeichen“?
- **Rechtlicher Diskriminierungsschutz:** Gleichbehandlungsgesetz untersagt Benachteiligung aufgrund bestimmter Merkmale
 - Geschlecht
 - Elternschaft
 - Alter
 - Beeinträchtigung



Kann man das „reparieren“?

- **Diskriminierungsrechtlich relevante Merkmale entfernen?**
 - Prognose nicht mehr akkurat
 - Proxy-Effekt
- **Andere Arten von Daten sammeln?**
 - Dauert Jahre
- **Assistenzsystem**
 - Routinemäßige Übernahme der Ergebnisse
 - „Automation Bias“
- **Fokus auf Gruppe C?**

Fazit: Anwendungszusammenhang

- **Strukturelle Analyse** des Arbeitsmarkts
- **„Ungleichheiten-Modell“** statt **„Arbeitsmarktchancen-Modell“**
- Sinnvolles **Diagnosewerkzeug**
- **Vergleich von Gruppen** über die Jahre
 - Datenbasierte Algorithmen „sehen“ das Individuum nicht
 - Keine Entscheidung über (vulnerable) Individuen
- **„Technologie ist weder gut noch böse – noch neutral.“** (Melvin Kranzberg)

Herzlichen Dank für Ihre Aufmerksamkeit!

Mag.a Paola Lopez
Universität Wien / WZB Berlin
Paola.Lopez@univie.ac.at