

Diversity Monitoring – aber wie?

Zwischen Diskriminierungs- und Datenschutz

Helena Mihaljević / 14. September 2022



**Hochschule für Technik
und Wirtschaft Berlin**

University of Applied Sciences

Digitale Personalauswahl

IT Unternehmen möchte Stellenanzeigen inklusiver gestalten und mehr Frauen zu einer Bewerbung bewegen. Hierfür sollen Technologien zur automatischen Textverbesserung eingesetzt werden.

Organisation X will mehr Bewerbungen auf Ausbildungsstellen bekommen und setzt auf eine Gamification-basierte Technologie, die sogar verspricht, fair zu sein.

Unternehmen Y möchte eine auf Persönlichkeitsdiagnostik basierende Software einsetzen, um den Bewerbungsprozess effizienter, objektiver und wissenschaftlicher zu gestalten.

Wer überprüft Technologien?

Unternehmen verlassen sich auf Technologie-Anbietende; selbst können sie keine Evaluationen der jeweiligen Produkte vornehmen:

- Sie erheben nicht systematisch und strukturiert Diversity-Angaben
- Für Monitoring des Auswahlprozesses und Technologie-Einsatzes ist techn.-math. Know-how notwendig (z.B. A/B Tests, vorher/nachher Vergleiche)

Es wird angenommen, dass Software-Anbietende ihre Produkte hinsichtlich (Effizienz sowieso) Fairness und Diskriminierungspotenzialen überprüfen.

Was können Software-Anbietende leisten?

Umfassende und zuverlässige Überprüfung der Technologien hinsichtlich Fairness* nicht möglich:

- I.d.R. sind nur **Laborstudien** möglich (allein aus ethischen Gründen)
- **Kontext** der Nutzung relevant: Technologie kann anders eingesetzt werden als beabsichtigt oder vorhergesehen
- Wenn mit einzelnen Kunden-Unternehmen kooperiert wird:
 - anfallende Daten i.d.R. zu **wenige**, um statistisch zuverlässige Schlüsse zu ziehen
 - **Verzerrung** bzgl. Branche, Region o.ä. zu erwarten → erschwert Rückschlüsse auf andere Situationen

Zugang für Dritte

Rechtlich: KI-Systeme (weit gefasster Begriff) sollen durch unabhängige Dritte evaluiert werden können

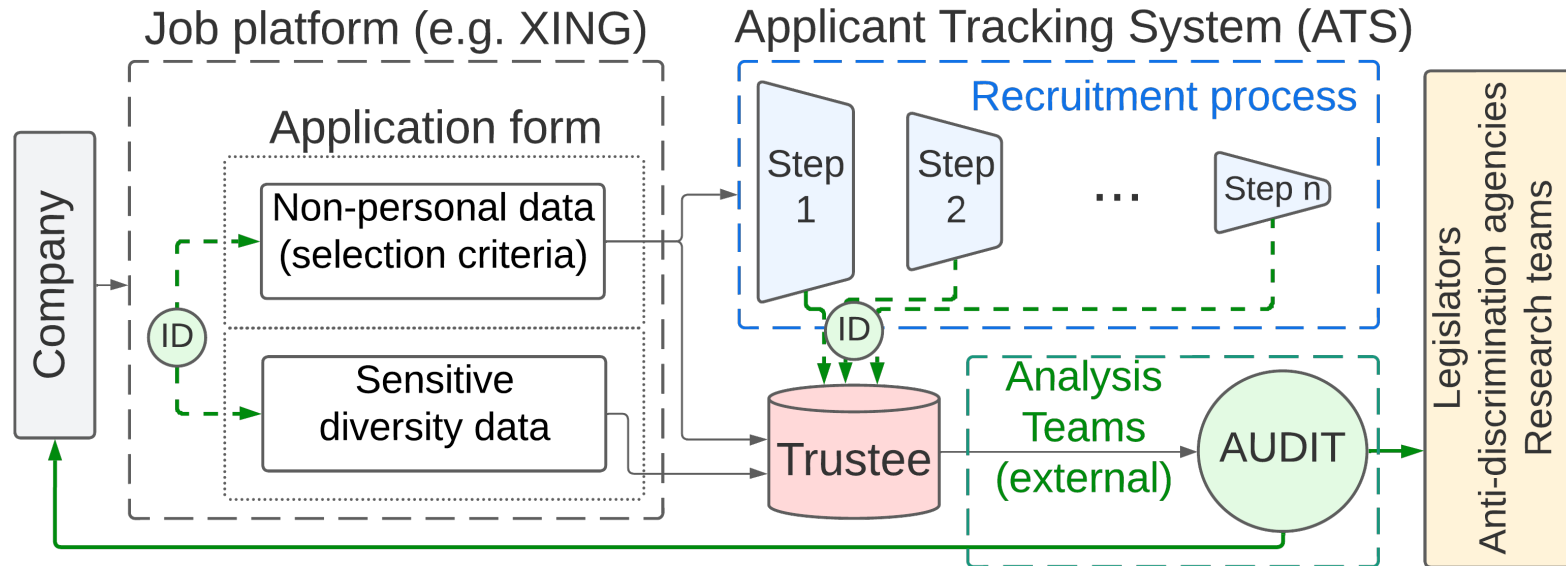
(Zivil-)gesellschaftlich: Es gibt einen gesellschaftlichen Bedarf, bei bedenklichen technologischen Entwicklungen intervenieren zu können

Algorithmus teilt Jobsuchende in drei Gruppen

Der AMS-Algorithmus teilt Jobsuchende in drei Kategorien ein. Um ihn zu entwickeln, [fütterte eine Firma den Algorithmus mit Daten aus der Vergangenheit](#), darunter Alter, Geschlecht, Wohnort, der bisherigen Laufbahn, Betreuungsverpflichtungen oder Staatsangehörigkeit. Auf Basis dieser Daten sagt das System voraus, ob die Chancen einer Person für den Wiedereinstieg hoch, mittel oder niedrig sind. Aus einem Menschen macht es so eine statistische Wahrscheinlichkeit.

Netzpolitik.org, 28.1.2021

Kontextualisierte, partizipative Auditierung der Fairness beim Einsatz von Personalauswahl-Technologien



Offene Fragen

- (Statistische) Fairness-Maße: was ist das und wie messen?
- Welche Diversity-Merkmale erheben und wie?
- Welche Rechtsgrundlage? Falls Einverständnis: wie mit Sampling Bias umgehen?
- Welche Stakeholder sollen Zugang zu welchem Wissen / Analysen bekommen?
- Wer soll den Data Trust betreiben?
- Welche Daten-Privacy und Daten-Sicherheits Mechanismen sollten implementiert werden?
- ...

Beispiel: statistische Fairness-Maße

- Es gibt verschiedene Zugänge, u.a. Gruppen-Fairness (aber auch Individual-Fairness, u.m.)
- Demographische Parität vs. bedingte demographische Parität
- Zshg. Mit Diversity-Zielen?
- Für bedingte Parität: lassen sich objektive sachliche Kriterien immer festlegen?
- Sind solche Kriterien (aus gesellschaftlicher Perspektive) an sich unfair?

Vielen Dank.



**Hochschule für Technik
und Wirtschaft Berlin**

University of Applied Sciences