

Big Data und Diskriminierung in der Arbeitswelt

Prof. Dr. Ingrid Schneider
Universität Hamburg, FB Informatik
Ethik in der Informationstechnologie (EIT)
Ingrid.Schneider@uni-hamburg.de

* * *

Neue Arbeitswelt und Digitalisierung. Welche Folgen haben neue Organisationsformen und Technologien? TA17, Wien, 19. Juni 2017

Gliederung

- Einführung und Literaturgrundlage
- Definition von Diskriminierung
- Profilbildung und Diskriminierungspotentiale
- Einfallstore für *bias* bei Big Data Anwendungen
- Normative Fragen und Forschungsfragen

Einführung

- Digitalisierung und Big Data
 - Inhalte und Organisationsformen von Arbeit
 - Bewertung der Arbeit
- **Art der Rekrutierung von Arbeitskräften**
- Job-Assessment-Center/ Auswahl von StellenbewerberInnen:
- algorithmische Methoden der Qualifizierungseinschätzung und prädiktive Analytik („Big Data“, Scores, Ranking)
- **→ Frage: Big Data: Beseitigen algorithmenbasierte Entscheidungen Diskriminierung oder wird sie fortgesetzt oder gar verstärkt?**

Google Bild-Suche „CEO“ Seite 1



CEO



Anmelden

Alle **Bilder** News Videos Bücher Mehr Einstellungen Tools

SafeSearch ▾



Ceo Clipart



CEO Portrait



Ceo Position



Ceo Auto



CEO Christia



<https://www.google.de/imgres?imgurl=http://www.lorenzanoandyanes.com/wp-content/uploads/2010/09/CEO...i&w=600&h=577&bih=804&biw=1320&q=ceo&ved=0ahUKEwjCs4XqoMLUAhXJCpoKHSIHdyEQMwg4KAiWAg&iact=mr&uact=8>



Google Bild-Suche „CEO“ Seite 2



Wer erhält überhaupt Stellenausschreibungen ?

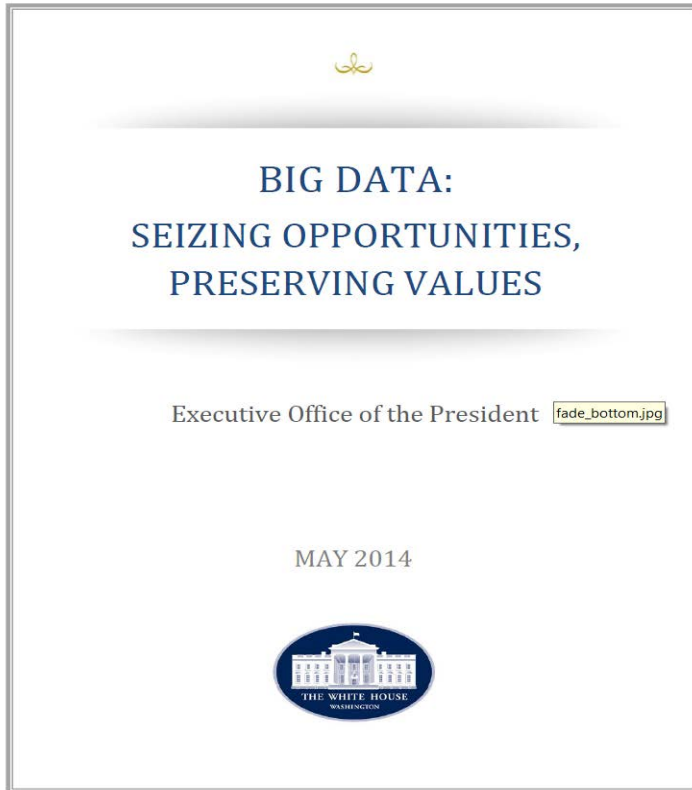
- Personen, die sich in ihren Google-Werbееinstellungen als weiblich identifizieren und dann per Google nach Jobs suchen, erhalten seltener Anzeigen für den beruflichen Aufstieg als männliche Nutzer (Datta et al. 2015)
- Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies*, 2015(1), 92-112.
<https://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf>

Ähnliches gilt für Xing, LinkedIn etc.

Literatur-Grundlage

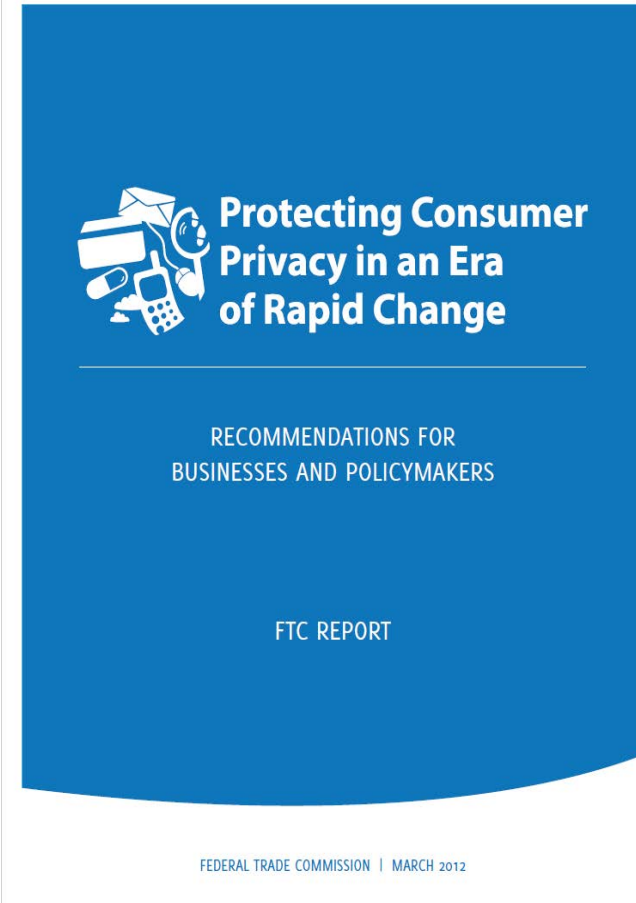
- Im deutschsprachigen Raum: Profilbildung und Erkenntnisse über sensible personenbezogene Informationen v. a. unter Privatsphäre und Datenschutz oder Überwachung diskutiert
- In den USA Debatte zu gesellschaftlichen Ein- und Ausschlusspotentialen von Big Data Analytik
- Zwei Berichte des Weißen Hauses 2014 (White House 2014) und 2016 (White House 2016) und der Federal Trade Commission (2016)
- → Analysen und Ergebnisse, soweit sie sich auf Diskriminierungspotentiale in der Arbeitswelt beziehen

White House Reports 2014 und 2016 („Podesta-Reports“)



“big data analytics have the potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, **employment**, health, education, and the marketplace.”
(White House - Podesta Report 2014)

Federal Trade Commission 2012 und 2016



Diskriminierung - Definition

- illegitime Ungleichbehandlung von Menschen(gruppen) auf der Grundlage eines oder mehrerer Kriterien
- “This report uses the term “discrimination” in a very broad sense to refer to **outsized harmful impacts—whether intended or otherwise**—that the design, implementation, and utilization of algorithmic systems can have on discrete communities and other groups that share certain characteristics.” (White House 2016: 25).
 - Nicht: (breitere) statistische Definition, die unter Diskriminierung jede Form von Unterschied zwischen Merkmalsträgern versteht
 - Nicht: (engere) juristische Definition, wie im Allgemeinen Gleichbehandlungsgesetz (AGG) festgelegt (Benachteiligungen aufgrund von Rasse, ethnischer Herkunft, Geschlecht, Religion, Behinderung, Alter oder sexuelle Identität. D.h. sozioökonomischer Status, politische Aktivität, Elternschaft u.a. fallen nicht hinein)

Differenzierung vs. Diskriminierung

- Bewertungen in Form von Rangfolgen, Präferenzen, Ein- und Ausschlüssen
- legitime *Differenzierungen* (Bsp. Kreditwürdigkeit)
- illegitime *Diskriminierung*
- Grenzen sind kontextspezifisch und Resultat von gesellschaftlichen Aushandlungsprozessen

Unterscheidung zwischen

- beabsichtigter
- unbeabsichtigter Diskriminierung

Personenprofile auf der Grundlage von Social Media Daten

- Arbeitgeber „googlen“ Bewerber
- Social Media Profile oft unzureichend geschützt
- Z.T. willkürliche oder veraltete Informationen (Stereotypisierung – Bsp. „Scheidungskind“)
- Einfallstore für Abwertung, Stigmatisierung
- Studie kombinierte Facebook -"Likes" mit demografischen Profilen und psychometrischen Tests:
 - sexuelle Orientierung der Männer: 88%
 - Religionszugehörigkeit (Christ oder Muslim, 82%)
 - Hautfarbe (weiße oder Afro-Amerikaner, 95%)
 - politische Ausrichtung (Demokrat or Republikaner, 85%)
 - Alkohol und Zigarettengebrauch (zwischen 65% und 75%)

richtig zugeordnet (Kosinski/Stillwell/Graepel 2013)

Profilbildung: Bsp. Acxiom

- US-Unternehmen Acxiom: bis zu 3.000 einzelne Eigenschaften von etwa 700 Millionen Menschen; über 2,5 Milliarden Kundenbeziehungen, 15.000 Kundendatenbanken von 7.000 Unternehmen aus Bereichen der Finanzwirtschaft, Versicherung, Handel, Gesundheit, Technologie oder Autoindustrie
- In Deutschland ist Acxiom seit 2004 mit einer Tochterfirma aktiv und hat Daten über 44 Millionen Bundesbürger/innen gesammelt (Christl 2014: 62)
- → „Personalisierte“ Werbung
- Aber auch → Zugang zu Arbeitsplätzen, Wohnungen, Gesundheitsdiensten, Krediten, Versicherungen oder Bildung?

Alter, Geschlecht, ethnische Herkunft: Verdeckte Diskriminierung

- Alters- oder Geschlechtsdiskriminierung ist bei Personalauswahl laut Antidiskriminierungsgesetzen verboten
- Indikatoren, die Aussagen über Geschlecht und Alter zulassen, können aber stellvertretend eingesetzt werden (proxy) und damit diskriminierende Präferenzen kaschieren (Hofstetter 2016: 380)
- etwa zu ethnischer Herkunft, Geschlecht oder sexueller Orientierung (Barocas/Selbst 2016: 692f)

Assessment-Center – algorithmengestützte Entscheidungen

- Führen diese zu "objektiveren" Auswahlprozessen oder rationalisieren und „versachlichen“ sie Diskriminierung?
- Sentiment-Analysen (Tonfall, Gestimmtheit, Wortschatz; Emotionen)
- Statistische Wahrscheinlichkeiten für Depressionen und andere physische und psychische Erkrankungen berechnet (Individuum \neq Gruppe!)
- Wahrscheinlichkeit der Verbleibdauer am Arbeitsplatz (u.a. durch Wohnort-Entfernung) (White House 2016, Barocas/Selbst 2016: 680)(Korrelation \neq Kausalität!)
- Gewichtung von Faktoren: Bsp. Reputation der Universität: Studium an einer Elite-Uni sagt wenig aus über die Fähigkeiten und Kompetenzen einer Person für eine bestimmte Tätigkeit

Big-Data-Anwendungen liefern Einfallstore für *bias*

- Verzerrungen in Datensätzen können zur Fortsetzung von Diskriminierung führen
- In Datensätzen schreiben sich *unbeabsichtigt* historische Vorurteile oder Ungleichgewichte fort: Eingaben oder Ergebnisse aus der Vergangenheit reproduzieren sich in den algorithmischen Outputs
 - Bsp.: Algorithmus korreliert Kinderzahl und Geschlecht mit Daten der bisherigen Einstellungspolitik: Wenn Mütter bisher kaum oder nur in Teilzeit bei Personalauswahl berücksichtigt, wird dies fortgeschrieben (Barocas und Selbst 2016: 689)

Gesellschaftliche Folgenabschätzung

- Diskriminierungen durch Big Data gehen nicht nur zu Lasten von Individuen
- Gesamtgesellschaftliche Implikationen: Zukunft des Solidaritätsprinzips
- Warnung vor weiterer "Benachteiligung der bereits Benachteiligten" (Barocas und Selbst 2016: 675)

FRANK PASQUALE

THE BLACK BOX SOCIETY

The Secret Algorithms
That Control Money
and Information

- Datensätze, Algorithmen und Entscheidungssysteme sind intransparent
 - Betriebs-, Geschäfts-geheimnisse oder geistiges Eigentum
 - Fehler oder rechtswidrige Praktiken?
- Transparenz, Rechenschaftspflicht und Verfahrensgerechtigkeit
 - „Algorithmenethik“
 - „Algorithmen- TÜV“

Nötige gesellschaftliche Auseinandersetzungen

- Welche Daten & Informationen sollten für Big Data-basierte Entscheidungen zugrunde gelegt werden?
- Können alle Indikatoren, die durch Mustererkennung identifiziert und als gewichtig angesehen werden, die Grundlage für Entscheidungen sein oder müssen rein „sachbezogene“ Informationen den Ausschlag geben?
- Sollten Kompetenzen der zentrale Faktor für die Bewertung von BewerberInnen sein oder dürfen auch Wohnortnähe und die Nutzung sozialer Medien zu Rate gezogen werden?
- Welche *proxies* sind in Datensätzen versteckt, die zur Benachteiligung bestimmter Gruppen führen können?
- In welchen Bereichen darf Big Data für Entscheidungen über Ein- und Ausschluss eingesetzt werden?

Big Data - Chancen und Risiken: normative Fragen

- Diskriminierungen offenlegen
- Egalitären Zugang zu Arbeitsplätzen gewährleisten
- Cave: Potential, Diskriminierungen zu verschärfen:
 - Ungleichbehandlung in immer mehr gesellschaftliche Bereiche hineinzutragen
 - Anspruch, Persönlichkeitsbewertungen berechenbar zu machen
 - Streben nach Effizienz und Effektivität bei Personalrekrutierung und zur Bewertung der Arbeitsleistung
- Versus ethische Ansprüche an die Wahrung menschlicher Autonomie (Zuboff 2015), gesellschaftlicher Solidarität (Rouvroy 2016) und Fairness (Ulbricht/Schneider 2017)

Forschungsfragen

- Nachweis von *bias* und Diskriminierung schwierig, kaum empirische Studien
- Wenig Zugang zu Datenmaterial
- In Europa bestimmte Parameter nicht (offiziell) abgebildet
- Wie „*equal opportunity by design*“ und „*bias mitigation*“ (Reflektion der individuellen Vorurteile) ausgestalten?
- Gruppenspezifische impact assessments (Sweeney 2013)?

Literatur

- ❑ Barocas, S./Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671–732.
- ❑ Christl, W. (2014). Kommerzielle Digitale Überwachung im Alltag. <http://crackedlabs.org/>
- ❑ Hofstetter, Y. (2016). *Das Ende der Demokratie*. Gütersloh.
- ❑ Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *NAS-Proceedings, USA*, 110(15), 5802–5805.
- ❑ Rouvroy, A. (2016). Council of Europe, “Of Data and Men”. *Fundamental Rights and Freedoms in a World of Big Data*.
- ❑ Sweeney, L. (2013). Discrimination in Online Ad Delivery. *Communications of the Association of Computing Machinery*, 56(5), 44–54.
- ❑ Ulbricht, L./Schneider, I.(2017): Ist Big Data fair? Normativ hergestellte Erwartungen an Big Data, in: Heil, R. et al.: *Big Data und Gesellschaft*. Wiesbaden (i.E.).
- ❑ White House (2014). *Big Data: Seizing Opportunities, Preserving Values*.
- ❑ White House (2016). *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*.
- ❑ Zuboff, S. (2015): Big other. Surveillance capitalism and the prospects of an information civilization. *J. of Information Technology*, 30 (1), 75–89.