

## **ABIDA – Assessing Big Data**

Vertiefungsstudie Arbeit – Fokusgruppenbericht

Münster, 27. Juli 2018

### **„Diskriminierung durch Algorithmen und KI im eRecruiting“**

#### **Fokusgruppenbericht**

---

Im Rahmen des ABIDA-Projekts fand am 27. Juli 2018 in Münster die Fokusgruppe zur Vertiefungsstudie Arbeit mit dem Titel „Diskriminierung durch Algorithmen und KI im eRecruiting“ statt. Hierbei kam es zu einem intensiven Austausch zwischen Vertretern aus dem Kontext von Arbeitsrecht, Personalberatung, Behörden und Wissenschaft.

#### **Aktueller Forschungsstand und Fokusgruppe**

Nachdem bereits der interdisziplinäre ABIDA-Expertenworkshop am 6. Juni 2018 zum Thema „People Analytics – Arbeit 4.0“<sup>1</sup> Forschungsbedarf beim Thema Gleichbehandlungsrecht und Arbeit 4.0 aufgezeigt hatte, wollte die Fokusgruppe diesen Bereich im Hinblick auf Rekrutierungsstrategien unter Beteiligung von Big-Data-Algorithmen näher ausleuchten.

Einleitend wurde der Umstand aufgegriffen, dass Bürger<sup>2</sup> die Validität von Big-Data-Algorithmen im Hinblick darauf anzweifeln, dass sie Eigenschaften von Bewerbern wie Teamfähigkeit und Sozialverhalten nicht angemessen abbilden können. Hierbei ist deutlich geworden, dass es entsprechender Software im Rekrutierungsbereich, die derartige Eigenschaften bewertet und die den Teilnehmenden bekannt ist, an Nachvollziehbarkeit mangelt. Um mehr Akzeptanz bei Betroffenen zu erreichen, erschien es den Teilnehmenden der Fokusgruppe sinnvoll, über die verwendeten Algorithmen bis zu einem gewissen Maße Transparenz herzustellen. Gefordert wurde hierbei aber vorrangig nicht die vollständige Offenlegung des Algorithmus, weil mit einer solchen Maßnahme nur relativ begrenzte Adressatenkreise einen Mehrwert gewannen, sondern Transparenz hinsichtlich der wesentlichen, für die letztlich getroffene Rekrutierungsentscheidung relevanten Parameter und ihrer Gewichtung. Hierdurch würde deutlich werden, dass der Einsatz von Big-Data-Algorithmen bei der Rekrutierung nicht dazu diene, zufällig ausgewählten Umständen nunmehr Relevanz für die Entscheidungsfindung zuzusprechen, obwohl diese nicht signifikant mit erheblichen Eigenschaften korrelierten. Vielmehr gehe es darum, plausible Zusammenhänge herzustellen. Für Bewerber sei dies im Hinblick auf die Nachvollziehbarkeit der Rekrutierungsentscheidung der Fall. Für Arbeitgeber könnte der unterstützende Einsatz von Big-Data-Algorithmen auch im Rahmen von Rekrutierungsentscheidungen sinnvoll sein. Es wurde hervorgehoben, dass Big-Data-Algorithmen auf diese Weise möglicherweise zur besseren Durchdringung von Wirkungszusammenhängen führen könnten. Dies gelte zum Beispiel für die Offenlegung struktureller Benachteiligung in Rekrutierungsprozessen. Zur Verdeutlichung wurde der Fall eines Arbeitgebers diskutiert, der eine ganz überwiegend männliche Belegschaft hatte. Das Ergebnis der

---

<sup>1</sup> Workshopbericht online abrufbar unter: <http://www.abida.de/de/blog-category/workshopberichte>.

<sup>2</sup> Die Verwendung des generischen Maskulinums erfolgt aus Gründen der besseren Lesbarkeit. Hiermit möchten die Verfasser keine Benachteiligung wegen des Geschlechts verstanden wissen.

Analyse einer Big-Data-Software, die der Arbeitgeber zur Suche nach Bewerbern eingesetzt hatte, war eine ungewöhnlich hohe Anzahl von männlichen Bewerbenden für die vom Arbeitgeber zu besetzende Stelle. Bei der Ursachenforschung hinsichtlich dieses Befundes wurde deutlich: Entscheidend war die große Zahl männlicher Bewerber, die sich auf die für Männer attraktive Arbeitgebermarke zurückführen ließ.

Die Teilnehmer kamen insgesamt überein, dass ein Einsatz automatisierter Instrumente wahrscheinlich und auch nicht zu verhindern sei, weil ein großes Interesse seitens der Beteiligten, insbesondere der Arbeitgeber hieran bestehe. Arbeitgebern komme es auf Aufwand, Kosten und Nutzen an und es sei so, dass algorithmische Instrumente dem Einsatz menschlicher Arbeitnehmer bezüglich Schnelligkeit überlegen seien. Zudem würde auch die scheinbare Objektivität als Überlegenheit gegenüber menschlicher Entscheidungsfindung wahrgenommen. Dies gelte in Rekrutierungsverfahren in erster Linie für den Abschnitt der Negativ-Selektion bei hohen Bewerberzahlen, d.h. beim Aussortieren von Bewerbern, um ein erstes Filtern von Bewerbungseingängen zu erreichen. Bei der Positiv-Selektion, also der Auswahl von bestimmten Bewerbern in einen engeren Kandidatenkreis, wurde eine automatisierte, rein algorithmische Entscheidung hingegen abgelehnt und auch als unwahrscheinlich eingestuft. Hier werde auch in Zukunft weiterhin auf übliche Elemente wie persönliche Vorstellungsgespräche zurückgegriffen.

Betont wurde seitens der Teilnehmer, dass die potentiellen Arbeitnehmer selbst auf Gefährdungen durch Algorithmen reagieren könnten, indem sie versuchten, Texte zu verfassen, um die von dem Algorithmus gesuchten Eigenschaften widerzuspiegeln. Es wurde hierbei aber auch klar, dass für eine solche Vorgehensweise ein gewisses Maß an Digitalkompetenz vonnöten sei, was nicht bei allen Bewerbern in ausreichendem Maße vorhanden sei. Entsprechende Software als Antwort auf die Verwendung von Algorithmen durch die Arbeitgeber würde allerdings schon entwickelt.

Die Fokusgruppe folgte im weiteren Verlauf einem sehr strukturierten Ansatz: Die Mitarbeiter des ABIDA-Projekts hatten vorgegeben, dass anhand des „Lebenszyklus“ eines Algorithmus nachvollzogen werden sollte, in welchen Phasen seiner Entstehung und seines Einsatzes besondere Potentiale für Benachteiligungen i.S.v. § 3 AGG bestehen. Hierzu wurde der „Lebenszyklus“ des Algorithmus in die Phasen der Zielfestlegung, der Operationalisierung und der Phase der Anwendung aufgeteilt.

Im Rahmen der Zielfestlegung wurde zunächst die grundlegende Entscheidung des Arbeitgebers beim Einsatz einer Big-Data-Software im Recruiting näher thematisiert. Bereits bei der Frage, als welche Art von Arbeitgeber sich ein Unternehmen präsentieren möchte, werden die ersten Weichen dahingehend gestellt, welche potentiellen Arbeitnehmer sich bewerben („Selbstselektion“). Auf der nächsten Stufe stellt sich für den Arbeitgeber die Frage „Welcher Arbeitnehmer wird eigentlich gesucht?“. Die Teilnehmer der Fokusgruppe gaben zu bedenken, dass Arbeitgeber sich oft nicht bewusst sind, welche Voraussetzungen die Bewerber mitbringen müssen bzw. für die angebotene Tätigkeit erforderlich sind. Hier besteht ein Benachteiligungspotential für die Bewerber über Vorurteile, aber auch über Kriterien, die sowohl mit geschützten Merkmalen i.S.v. § 1 AGG (ethnische Herkunft, Geschlecht, Religion oder Weltanschauung, Alter, Behinderung, sexuelle Orientierung) als auch mit relevanten Eigenschaften korrelieren. Die wirtschaftliche Zielfestlegung des Arbeitnehmers bestünde oftmals aus „weichen“ Kriterien (etwa bei Frage: Was ist ein „guter Arbeitnehmer“?), die die Gefahr einer zumindest mittelbaren Benachteiligung im Sinne des AGG bergen. Auf dieser Ebene würden bereits viele Fehler gemacht, weil die Unternehmen der Frage nach dem geeigneten Bewerber nicht intensiv genug nachgingen. Hierbei waren für die Teilnehmer die Fragen von Relevanz, wer die zur Beurteilung von Bewerbern maßgeblichen Kriterien festlegte und was dabei von wem erfasst werde. Wichtig sei auch, dass die Kriterien den Bezug zu den nötigen beruflichen Anforderungen aufwiesen. Gegebenenfalls sei demnach die Rechtfertigung einer unmittelbaren Benachteiligung möglich bzw. es läge schon gar keine mittelbare Benachteiligung vor.

Hierzu gaben die Teilnehmer einen Einblick in die Praxis. Bei größeren Unternehmen sei es üblich, dass diese sich bis zu zehn relevante Kriterien für zu besetzende Stelle suchten. Hierbei seien mehrere Fachbe-

reiche beteiligt. In den Unternehmen seien es Psychologen, die Tests und Kriterien zur Rekrutierung entwickelten, nicht die Programmierer. Diese brächten häufig 10 bis 15 Jahre Berufserfahrung mit. Entscheidend sei, dass bei der Auswahl von Kriterien ein relevanter Ursache-Wirkungs-Zusammenhang zwischen dem verwendeten Kriterium und der Performance des Bewerbers bestünde. Gesucht werde nach der bestmöglichen Besetzung einer Position.

Es wurde weiter darauf aufmerksam gemacht, dass nach derzeitiger Rechtslage ein abgelehnter Bewerber keinen Anspruch darauf habe, zu erfahren, warum er abgelehnt worden sei. Insofern böten Algorithmen auch die Chance, einen solchen Anspruch zu realisieren, wenn die notwendige Transparenz hergestellt werde und die „black box“ so nachvollziehbar wie möglich geöffnet werde. Dies leitete eine kurze Diskussion über Sinn und Unsinn von Feedback ein, bei der zum einen deutlich gemacht wurde, dass Feedback zu einer positiven „candidate experience“ führen könne. Andererseits wurde angemerkt, dass Feedback auch einen möglicherweise gar nicht intendierten Anpassungsdruck auslösen könnte. Generell wurde hervorgehoben, dass nicht mit Abschluss eines Arbeitsvertrags endende Verhandlungen zwischen Arbeitgeber und Bewerber etwas Positives für beide Seiten sein könnten, wenn der Arbeitgeber und der Bewerber im Ergebnis nicht zusammenpassten.

Es wurde ebenfalls deutlich, dass die verwendete Datengrundlage ein erhebliches Benachteiligungspotential aufweisen kann. Als potentiell diskriminierendes Phänomen wurde sowohl die Validität als auch die Varietät von Datengrundlagen angesprochen. Eine Datengrundlage könnte etwa wegen einer nicht ausreichend großen Datenmenge nicht hinreichend valide sein.

Auch die Varietät der Datengrundlage spiele eine entscheidende Rolle, um potentiell diskriminierenden Effekten vorzubeugen. So entstehe aufgrund einseitiger Unternehmensdaten durch die Über- bzw. Unterrepräsentation von geschützten Gruppen (z.B. überwiegender Anteil von Männern im Unternehmen) ein Echo-Effekt.

Darüber hinaus sei auch eine Benachteiligung auf der Anwenderebene möglich. Hierzu setzten sich die Teilnehmer vertieft mit der Mensch-Maschine-Schnittstelle, also dem Umgang des menschlichen Entscheiders mit dem Ergebnis einer entscheidungsunterstützenden Software, auseinander. Es sei wichtig, dass der Anwender das Ergebnis des Big-Data-Algorithmus nicht als „Vorgabe“ oder als „Empfehlung“ auffasse, sondern es kritisch hinterfrage. Hier bestehe die Möglichkeit, möglicherweise benachteiligende Ergebnisse zu hinterfragen. Die Teilnehmer waren sich einig, dass in der Regel Ergebnisse vermutlich erst dann hinterfragt werden, wenn durch den Algorithmus nicht genügend passende bzw. qualifizierte Bewerbende vorgeschlagen werden. Es würde mehr Kraft benötigt, von der Entscheidung des Algorithmus abzuweichen. Es wurde auch angesprochen, wie ein Ergebnis sinnvoll zu visualisieren sei, um Bewusstsein für das Problem von Benachteiligungen zu erzeugen, ohne Anwenderfreundlichkeit einzubüßen. Die Teilnehmer befürworteten ein Dashboard-Format und betonten, dass das Ergebnis möglichst unmittelbar wahrnehmbar sein müsste.

Abschließend wurden Lösungsansätze entworfen, um den herausgearbeiteten Diskriminierungspotentialen zu begegnen.

Besprochen wurde etwa die Möglichkeit der Einführung von sog. Algorithmic Audits, die in den USA als Mittel der Selbstregulierung gegen Benachteiligungsrisiken angeboten werden. Diesbezüglich waren die Teilnehmer grundsätzlich skeptisch, da die Erfahrungen mit ähnlichen bestehenden Regulierungen wie Zertifizierungen im datenschutzrechtlichen Bereich sie hinsichtlich ihrer Effektivität nicht durchgehend überzeugt hätten.

Ein weiterer Vorschlag war die Selbstzertifizierung mittels DIN-Normen. Hier wurde von den Teilnehmern die bereits bestehende DIN 33430 des Berufsverbands Deutscher Psychologinnen und Psychologen (BDP) für Eignungsbeurteilung genannt, da ein Teilnehmer danach arbeitet.

Maßnahmen des kollektiven Rechtsschutzes wurden kontrovers beurteilt. Einerseits sahen Teilnehmer das Benachteiligungspotential von Algorithmen bei der Rekrutierung von Arbeitnehmern aufgrund der techno-

logischen Komplexität des Sachverhalts als geradezu prädestiniert für Maßnahmen des kollektiven Rechtsschutzes an, weil die größte Problematik von Verstößen gegen das Benachteiligungsverbot des AGG in Nachweisschwierigkeiten bestehe. Vorrangig sei daher neben dem Verbandsklagerecht der Auskunftsanspruch über die Verwendung von Algorithmen, bestenfalls eine Kennzeichnungspflicht, sowie über den nachvollziehbaren Inhalt der „black box“. Andere Teilnehmer wiederum hinterfragten die Effektivität eines solchen Vorgehens und verwiesen auf die bereits umfangreich und aufwändig ausgestaltete Mitbestimmung bei IT-Sachverhalten in den Unternehmen.

## Ausblick

Als Ergebnis der Fokusgruppe wurde deutlich, dass in bestimmten Teilbereichen der Rekrutierung von Arbeitnehmern (größere Unternehmen, Negativ-Selektion bei hohen Bewerberzahlen) der Einsatz von Big-Data-Algorithmen bereits Realität ist bzw. in näherer Zukunft wohl bevorsteht. Angesichts einer systematischen Analyse des „Lebenszyklus“ des Algorithmus wurde deutlich, dass auf jeder Ebene Potential für Benachteiligungen besteht. Als Gegenmaßnahme erscheint im Bereich der Selbstregulierung die Förderung von Algorithmic Audits möglich. Weiter könnten möglicherweise Maßnahmen des kollektiven Rechtsschutzes angesichts der Komplexität der Thematik und der auch hieraus resultierenden Nachweisschwierigkeiten von Betroffenen helfen der Problematik zu begegnen, sowie ein Auskunftsanspruch der Bewerbenden über die Gründe für die Ablehnung und das Verfahren für die Entscheidungsfindung, der teilweise auch schon für analoge Bewerbungsverfahren gefordert wird.

Das ABIDA-Team bedankt sich nochmals recht herzlich bei den Teilnehmern für die gewährten Einblicke.

Weiterführende Links:

*Dzida/Groh, NJW 2018, 1917, Hoeren/Niehoff, RW 2018, 47 Brockmeyer, ZD 2018, 258; Hoeren, ZD 2015, 102.*