STIVEN CATALYST

DRITTE SERIE · NOVEMBER 2026

MANAGEMENT REVIEW

Management ohne Theater.



KI

KI, DIE LEBENSLÄUFE SICHTET: DER TEST UND DAS GESETZ

Derselbe Lebenslauf unter anderen Namen, die Gruppe, die am meisten verlor, was die KI-Verordnung von Anbietern und Arbeitgebern verlangt, ein Ergebnis, das eine Designentscheidung umkehrte, ein Test für das eigene System und eine Karte.

85,1 %

der Tests bevorzugten Lebensläufe mit weiß konnotierten Namen gegenüber denselben mit Schwarz konnotierten; 8,6 % umgekehrt.

WILSON & CALISKAN, 2024

IN DIESER AUSGABE

TITELTHEMA
Derselbe Lebenslauf, ein anderer Name
SEITE 03

DAS MODELL
Was die KI-Verordnung verlangt
SEITE 05

WERKZEUG DER AUSGABE
Eine Karte für ein Sichtungssystem
SEITE 08

NR. 106 · KI

In dieser Ausgabe

Viele Bewerbungen liest eine Software, bevor ein Mensch sie sieht, und Sprachmodelle machen einen weiteren Filter billig. Ein Filter, der Menschen reiht, entscheidet, wer angesehen wird. Diese Ausgabe fragt, was ein sorgfältiger Test fand, als sich nur der Name änderte, und was europäisches Recht verlangt, bevor ein solches System jemanden sichtet.

In einer Studie von 2024 reihten drei KI-Modelle dieselben Lebensläufe unter verschiedenen Namen gegen echte Stellenanzeigen. Weiß konnotierte Namen wurden in 85,1 % der Tests bevorzugt, Namen Schwarzer Männer schlugen die weißer Männer nie. Eine Korrektur von 2026 und ein Test mit anderen Namen zeigen, wie viel am Testdesign hängt. Die KI-Verordnung stuft solche Systeme als hochriskant ein und verpflichtet Anbieter und Arbeitgeber. Ein Namenstausch-Test und eine Karte schließen die Ausgabe.

Stiven Janaqi HERAUSGEBER · STIVEN CATALYST

03 TITELTHEMA
Derselbe Lebenslauf, ein anderer Name

04 DIE ZAHLEN
Die Gruppe, die am meisten verlor

05 DAS MODELL
Was die KI-Verordnung verlangt

06 WAS DIE FORSCHUNG SAGT
Ein Ergebnis, das sich umkehrte

07 WIE ES GEMESSEN WIRD
Das eigene System testen

08 WERKZEUG DER AUSGABE
Eine Karte für ein Sichtungssystem

09 QUELLEN
Quellen und Methode

SO LESEN SIE DIESE AUSGABE

Zahl
Jede Zahl hat Quelle und Jahr unten auf ihrer Seite.

Unsere Lesart
Wo der Herausgeber deutet und nicht die Forschung, steht es dabei.

Praxis
Schritte und Karte sind Vorschläge zum Ausprobieren, keine Forschungsergebnisse.

TITELTHEMA

Derselbe Lebenslauf, ein anderer Name

Kyra Wilson und Aylin Caliskan nahmen 554 öffentliche Lebensläufe und 571 Stellenanzeigen aus neun Berufen, von Geschäftsführern bis zu Lehrkräften. Jeder Lebenslauf bekam einen Vornamen, der mit Schwarzen oder weißen Männern oder Frauen verbunden wird, und den Nachnamen Williams. Drei offene KI-Modelle für den Textabgleich reihten die Lebensläufe je Anzeige, die besten 10 % galten als ausgewählt.

TESTS NUR NACH HERKUNFT, DREI MODELLE × NEUN BERUFE: WELCHE LEBENSLÄUFE SIGNIFIKANT ÖFTER AUSGEWÄHLT WURDEN (2024)



Der Aufbau funktionierte als Filter: Ohne Namen schnitten Lebensläufe aus dem ausgeschriebenen Beruf in jedem Beruf besser ab als andere. Die einzigen Tests, in denen Schwarz konnotierte Namen gewannen, stammten von einem der drei Modelle.

UNSERE LESART

In diesen Lebensläufen unterschied sich nur der Name. Was immer die Modelle bevorzugten, die Arbeit war es nicht.

Quelle: Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024

Wilson & Caliskan (2024), Abbildung 4. Eine Simulation mit US-Namen und -Berufen, keine Prüfung eines eingesetzten Werkzeugs; Lebensläufe wurden bei 1.300 Tokens gekürzt. Anteile an 27 Chi-Quadrat-Tests bei p < 0,05. Namen zeigen Herkunft und Geschlecht nur grob an, wie die Forschenden anmerken.

DIE ZAHLEN

Die Gruppe, die am meisten verlor

Dann verglichen die Forschenden die vier Gruppen paarweise, wieder 27 Tests je Paar. Die Unterschiede zwischen Namen weißer Männer und weißer Frauen waren am kleinsten; Vergleiche mit Schwarz konnotierten Namen zeigten größere.

ANTEIL VON 27 TESTS, IN DENEN LEBENSLÄUFE DER ERSTEN GRUPPE SIGNIFIKANT ÖFTER AUSGEWÄHLT WURDEN (2024)



Namen Schwarzer Männer wurden in 14,8 % der Tests denen Schwarzer Frauen vorgezogen, denen weißer Männer in keinem. Die Forschenden merken an, dass dies dem entspricht, was Feldstudien zur Einstellung seit Langem finden, und dass die Verzerrung gegen eine Gruppe nicht die Summe eines Herkunfts- und eines Geschlechtseffekts ist.

UNSERE LESART

Eine Prüfung nur nach Herkunft und nur nach Geschlecht kann die Gruppe übersehen, die am meisten verliert. Prüfen Sie die Kombinationen.

Quelle: Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024

Wilson & Caliskan (2024), Abbildungen 6–9; 20 Namen je Gruppe. Der Rest zu 100 % sind Tests ohne signifikanten Unterschied. Diese intersektionalen Ergebnisse betraf die Korrektur von 2026 nicht; ein Team der Universität Salzburg hat sie repliziert.

DAS MODELL

Was die KI-Verordnung verlangt

Anhang III Nummer 4 Buchstabe a nennt KI für Einstellung oder Auswahl, insbesondere für gezielte Stellenanzeigen, zum Analysieren und Filtern von Bewerbungen und zum Bewerten von Bewerbern. Solche Systeme sind hochriskant; die Regeln gelten ab dem 2. Dezember 2027. Das trifft jede Seite.

Der Anbieter, der es baut

- Prüft seine Daten auf mögliche Verzerrungen, erkennt, verhindert und mindert sie (Art. 10 Abs. 2 lit. f, g)
- Nennt in der Anleitung die Genauigkeit und, wo angebracht, Ergebnisse für bestimmte Gruppen (Art. 13 Abs. 3 lit. b)
- Lässt die Aufsicht übermäßiges Vertrauen im Blick behalten und das Ergebnis übergehen (Art. 14 Abs. 4 lit. b, d)

Der Arbeitgeber, der es einsetzt

- Hält die Eingabedaten, die er steuert, relevant und hinreichend repräsentativ (Art. 26 Abs. 4)
- Erklärt Betroffenen auf Verlangen die Rolle der KI und die wesentlichen Elemente der Entscheidung (Art. 86)
- Kann selbst Anbieter werden, wenn er allgemeine KI zur Sichtung von Lebensläufen umwidmet (Art. 25 Abs. 1 lit. c)

Die letzte Zeile zählt vor allem für kleine Firmen: Ein allgemeiner Chatbot, der Bewerbungen sortieren soll, dient einem hochriskanten Zweck, für den er nicht in Verkehr gebracht wurde. Artikel 25 gibt der Firma, die das tut, dann die Pflichten des Anbieters nach Artikel 16.

UNSERE LESART

„Ein Mensch prüft das“ ist keine Aufsicht, wenn dieser Mensch nur die Auswahlliste sieht. Aufsicht beginnt bei den Abgelehnten.

Quellen: European Parliament and Council of the European Union, Official Journal of the EU, 2024; European Parliament and Council of the European Union, Official Journal of the EU, 2026

Verordnung (EU) 2024/1689, Artikel 10, 13, 14, 16, 25, 26 und 86 und Anhang III; Datum nach Artikel 113 in der Fassung der Verordnung (EU) 2026/1744, die die hier zitierten Bestimmungen unverändert ließ. Formulierung und Lesehinweis: Redaktion; keine Rechtsberatung. Fristen, weitere Arbeitgeberpflichten und Art. 6 Abs. 3 in Nr. 41. Der Essay: die Sicht des Autors, wer entscheidet, wenn ein System zuerst entscheidet.

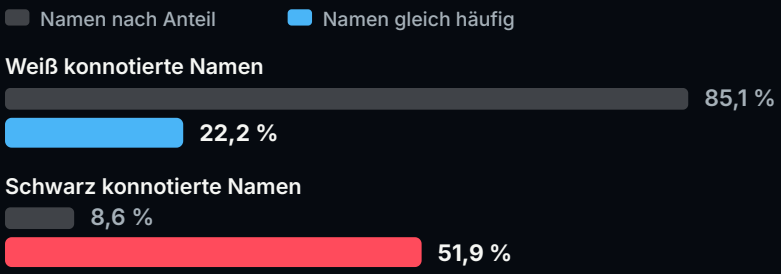
Mehr im Essay Das System ist schneller als Sie

WAS DIE FORSCHUNG SAGT

Ein Ergebnis, das sich umkehrte

Der Haupttest nutzte weiße Namen, die in Texten etwa 5,5-mal so häufig vorkommen wie die Schwarzen, passend zur Bevölkerung. In einem zweiten Durchlauf waren die weißen Namen etwa so häufig wie die Schwarzen. Sonst änderte sich nichts.

ANTEIL DER 27 HERKUNFTSTESTS, DEN JEDE GRUPPE GEWANN, NACH WAHL DER NAMEN (2024)



Im August 2026 ergänzten die Forschenden eine Korrektur: Nachdem ein Team der Universität Salzburg die Studie zu wiederholen versucht hatte, fand sich ein Programmfehler, der in den reinen Geschlechtertests die Bezeichnungen vertauscht hatte. Frauennamen, nicht Männernamen, wurden in 51,9 % dieser Tests bevorzugt. Die Ergebnisse zu Herkunft und Überschneidungen blieben bestehen.

UNSERE LESART

Ein einzelnes Audit ist eine Messung, kein Urteil. Fragen Sie, wie der Test gebaut war und ob ihn jemand wiederholt hat.

Quelle: Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024

Wilson & Caliskan (2024), arXiv-Version 3 (30. August 2026) mit dem Korrekturhinweis; Abschnitt zu Häufigkeitseffekten. Die Forschenden deuten die Umkehr als Empfindlichkeit der Modelle für die Häufigkeit eines Namens. Der Lesehinweis stammt von der Redaktion.

WIE ES GEMESSEN WIRD

Das eigene System testen

Die Methode der Studie lässt sich in einen Test für ein bereits genutztes System verwandeln: identische Lebensläufe, die sich in einem Signal unterscheiden, durch denselben Schritt geschickt, die Auswahlquoten verglichen. Die Schritte unten sind ein Vorschlag der Redaktion, nicht der Studie.

1. **Echte Anzeigen, erfundene Lebensläufe**
   Keine Daten echter Bewerber; mehrere Berufe.

2. **Nur ein Signal ändern**
   Ein Name, ein Frauenverein, eine Schule, ein Alter; und Kombinationen.

3. **Die echte Grenze behalten**
   Den Anteil, den wirklich ein Mensch liest, etwa die besten 10 %.

4. **Auswahlquoten vergleichen**
   Je Gruppe und Kombination; dann Länge und Wortwahl variieren.

5. **Melden**
   Protokoll aufbewahren; den Anbieter über Risiken informieren (Art. 26 Abs. 5).

UNSERE LESART

Ein System, das nie mit getauschtem Namen getestet wurde, hat keine bekannte Verzerrung, was nicht dasselbe ist wie keine.

Quellen: Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024; European Parliament and Council of the European Union, Official Journal of the EU, 2024; European Parliament and Council of the European Union, Official Journal of the EU, 2026

Vorschlag der Redaktion nach dem Vorbild von Wilson & Caliskan (2024), die auch die Länge der Lebensläufe änderten und damit die Ergebnisse veränderten. Seit Juli 2026 erlaubt Artikel 4a Arbeitgebern mit Hochrisiko-KI, sensible Daten zur Erkennung von Verzerrungen unter strengen Bedingungen zu verarbeiten; eine Pflicht dazu entsteht nicht. Das strukturierte Interview steht in Nr. 65.

WERKZEUG DER AUSGABE

Eine Karte für ein Sichtungssystem

Eine Karte für jedes System, das Bewerbungen vor einem Menschen liest. Das Werkzeug ATS CV zeigt die andere Seite: wie Bewerbungssoftware einen Lebenslauf liest und was Bewerber tun können, damit er richtig gelesen wird.

01	**System und Zweckbestimmung**	für Recruiting gebaut oder ein allgemeines Werkzeug, das wir umgewidmet haben?
02	**Was es im Ablauf tut**	Anzeigen steuern, filtern, reihen oder bewerten; wo ein Mensch eingreift
03	**Wer die Abgelehnten sieht**	ein Name, wie viele geprüft werden, wer übergehen darf
04	**Letzter Test mit getauschtem Namen**	Datum, Signale und Kombinationen, Ergebnis, an den Anbieter
05	**Unsere Eingabedaten**	Anzeigen und Kriterien, die wir eingeben: für die Stelle relevant?
06	**Bewerber informiert, Erklärung**	wo sie vom KI-Einsatz erfahren; wer Anfragen beantwortet

Quellen: European Parliament and Council of the European Union, Official Journal of the EU, 2024; Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024

Die Karte ist ein Vorschlag der Redaktion, gestützt auf die Artikel 14, 25, 26 und 86 der KI-Verordnung und die Methode von Wilson & Caliskan (2024); sie ist kein geprüftes Instrument und keine Rechtsberatung. Pflichten des Arbeitgebers vor dem Einsatz, etwa die Information der Arbeitnehmervertretung, stehen in Nr. 41.

Tool öffnen: ATS-Lebenslauf stivencatalyst.com/de/tools/ats-cv/

QUELLEN UND METHODE

Jede Zahl hat eine Quelle.

Die Zahlen dieser Ausgabe stammen aus den folgenden Quellen. Das Jahr zeigt, wie aktuell jede ist.

Gender, Race, and Intersectional Bias in Resume Screening via Language Model Retrieval (arXiv version 3, with the authors' correction of 29 August 2026)
Kyra Wilson & Aylin Caliskan, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, 2024
https://arxiv.org/abs/2407.20371

Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)
European Parliament and Council of the European Union, Official Journal of the EU, 2024
https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Regulation (EU) 2026/1744 amending Regulations (EU) 2024/1689, (EU) 2018/1139 and (EU) 2023/1230 (Digital Omnibus on AI)
European Parliament and Council of the European Union, Official Journal of the EU, 2026
https://eur-lex.europa.eu/eli/reg/2026/1744/oj

REDAKTIONELLE METHODE

Jede Zahl wurde auf Jahr, Herausgeber und Bedeutung geprüft, am Volltext, wo er sich öffnen ließ. Eine nur in einer Zusammenfassung, einem Abstract oder Auszügen gelesene Quelle ist mit „über“ markiert. Eine Zahl, die sich im Gelesenen nicht bestätigen ließ, ist auf ihrer Seite so markiert oder fehlt. Die Deutung des Herausgebers ist als „Unsere Lesart“ markiert.



MANAGEMENT REVIEW

Management ohne Theater.

Jede Ausgabe eine Führungsfrage, geprüft an der besten Forschung.

ALLE AUSGABEN

stivencatalyst.com/de/magazine/management-review.html

Management Review · Nr. 106 · November 2026 · Stiven Catalyst
Herausgegeben von Stiven Janaqi, der sie verantwortet; „die Redaktion" ist in dieser Publikation er. Texte, Übersetzungen, Layout und die Prüfung der Zahlen an ihren Quellen entstehen in Arbeitssitzungen mit Claude, einem KI-Modell von Anthropic.
stivencatalyst.com/de/how-this-is-made.html