STIVEN CATALYST

THIRD SERIES · NOVEMBER 2026

MANAGEMENT REVIEW

Management without theatre.



AI

AI THAT SCREENS CVS: THE TEST AND THE LAW

The same CV under different names, the group that lost most, what the AI Act asks of providers and employers, a result that turned around with one design choice, a way to test your own screen, and a card.

85.1%

of tests favoured CVs with White-associated names over the same CVs with Black-associated names; 8.6% the reverse.

WILSON & CALISKAN, 2024

INSIDE

COVER STORY
Same CV, another name
PAGE 03

THE MODEL
What the AI Act asks
PAGE 05

TOOL OF THE ISSUE
A card for a CV screen
PAGE 08

NO. 106 · AI

In this issue

Many applications are read by software before a person sees them, and language models make it cheap to add one more filter. A filter that ranks people decides who gets a look. This issue asks what a careful test found when only the name changed, and what European law now asks before such a system screens anyone.

In a 2024 study the same CVs, under different names, were ranked against real job ads by three AI models. Names associated with White people were favoured in 85.1% of tests, and Black men's names never beat White men's. A correction in 2026 and a test with other names show how much the design matters. The AI Act counts such systems as high-risk and sets duties for providers and employers. A name-swap test and a card close the issue.

Stiven Janaqi EDITOR · STIVEN CATALYST

03 COVER STORY
Same CV, another name

04 THE NUMBERS
The group that lost most

05 THE MODEL
What the AI Act asks

06 WHAT THE RESEARCH SAYS
A result that turned around

07 HOW IT IS MEASURED
Test your own screen

08 TOOL OF THE ISSUE
A card for a CV screen

09 SOURCES
Sources and method

HOW TO READ THIS ISSUE

Figure
Every figure has its source and year at the foot of its page.

Our reading
Where the editor interprets rather than the research, it says so.

Practice
The steps and the card are proposals to try, not research results.

COVER STORY

Same CV, another name

Kyra Wilson and Aylin Caliskan took 554 public CVs and 571 job ads in nine occupations, from chief executives to teachers. Each CV got a first name linked to Black or White men or women, and the surname Williams. Three open AI models, built to match texts, ranked the CVs for each ad, and the top 10% counted as selected.

TESTS BY RACE ONLY, THREE MODELS × NINE OCCUPATIONS: WHICH CVS WERE SELECTED SIGNIFICANTLY MORE OFTEN (2024)



The set-up worked as a screen: without names, CVs from the advertised occupation scored higher than others for every occupation. The only tests where Black-associated names won came from one of the three models.

OUR READING

Nothing in these CVs differed but the name. Whatever the models preferred, it was not the work.

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

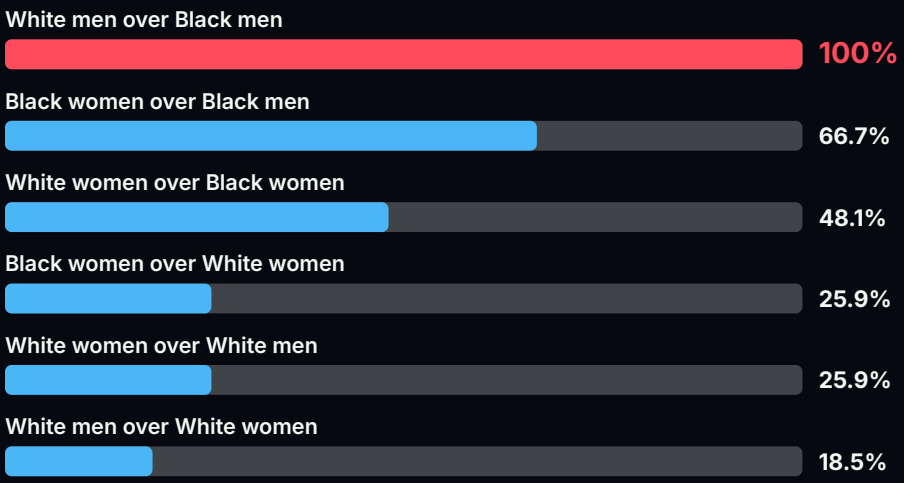
Wilson & Caliskan (2024), figure 4. A simulation with US names and occupations, not an audit of a tool in use; CVs were cut at 1,300 tokens. Shares are of 27 chi-square tests at $p < 0.05$. Names signal race and gender only roughly, as the authors note.

THE NUMBERS

The group that lost most

The researchers then compared the four groups pair by pair, again 27 tests for each pair. The differences between White men's and White women's names were the smallest; comparisons that included Black names showed larger ones.

SHARE OF 27 TESTS IN WHICH THE FIRST GROUP'S CVS WERE SELECTED SIGNIFICANTLY MORE OFTEN (2024)



Black men's names were preferred to Black women's in 14.8% of tests and to White men's in none. The authors note that this follows what field studies of hiring have long found, and that the bias against one group is not the sum of a race effect and a gender effect.

OUR READING

A check by race alone and by gender alone can miss the group that loses most. Test the combinations.

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

Wilson & Caliskan (2024), figures 6–9; 20 names per group. The rest of each 100% is tests with no significant difference. These intersectional results were not affected by the 2026 correction and were replicated by a team at the University of Salzburg.

THE MODEL

What the AI Act asks

Annex III, point 4(a) lists AI meant for recruitment or selection, in particular to target job ads, to analyse and filter applications and to evaluate candidates. Such systems are high-risk; the rules apply from 2 December 2027. Here is what falls on each side.

The provider, who builds it

- Examines its data for possible biases and takes measures to detect, prevent and reduce them (Art. 10(2)(f), (g))
- States accuracy and, where appropriate, results for specific groups in the instructions (13(3)(b))
- Is built so overseers stay aware of over-reliance and can override the output (14(4)(b), (d))

The employer, who uses it

- Keeps the input data it controls relevant and sufficiently representative (26(4))
- Explains, when an affected candidate asks, the AI's role and the main elements of the decision (86)
- Can become the provider by turning a general AI system to screening CVs (25(1)(c))

The last line matters most for small firms: a general chatbot set to sort applicants is put to a high-risk purpose it was not placed on the market for. Article 25 then gives the firm that did so the provider's duties under Article 16.

OUR READING

“A person checks it” is not oversight if that person only sees the shortlist. Oversight starts with the rejected.

Sources: 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

Regulation (EU) 2024/1689, Articles 10, 13, 14, 16, 25, 26 and 86 and Annex III; date from Article 113 as amended by Regulation (EU) 2026/1744, which left the provisions cited here unchanged. The wording and the callout are the editors'; not legal advice. Dates, the employer's other duties and Article 6(3) are in No. 41. The linked essay is the author's view of who decides when a system decides first.

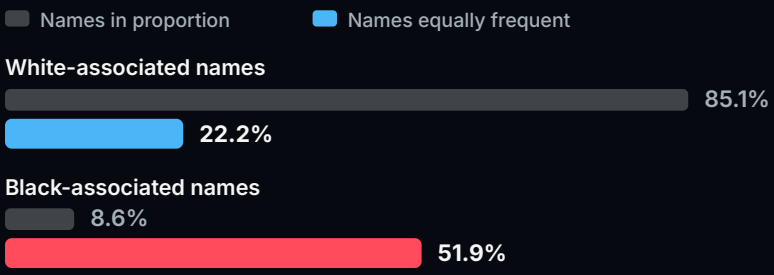
More in the essay The system is faster than you

WHAT THE RESEARCH SAYS

A result that turned around

The main test used White names that occur about 5.5 times as often in written text as the Black names, matching the population. In a second run the White names were about as frequent as the Black ones. Nothing else changed.

SHARE OF 27 RACE TESTS WON BY EACH GROUP, BY HOW THE NAMES WERE CHOSEN (2024)



In August 2026 the authors added a correction: after a team at the University of Salzburg tried to replicate the study, a bug was found that had swapped the labels in the gender-only tests. Women's names, not men's, were the ones preferred in 51.9% of those tests. The race and intersectional results stood.

OUR READING

One audit is a reading, not a verdict. Ask how the test was built, and whether anyone repeated it.

Source: 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) with the correction note; section “Frequency effects”. The authors read the reversal as the models' sensitivity to how often a name occurs. The callout is the editors'.

HOW IT IS MEASURED

Test your own screen

The study's method can be turned into a test for a screen a team already uses: identical CVs that differ in one signal, run through the same step, with the selection rates compared. The steps below are the editors' proposal, not the study's.

1. **Take real ads, invented CVs**
   No real applicant's data; several occupations.
2. **Change one signal only**
   A name, a women's club, a school, an age; and their combinations.
3. **Keep your real cut-off**
   The share that a person actually reads, such as the top 10%.
4. **Compare selection rates**
   Per group and per combination; then vary length and wording.
5. **Report it**
   Keep the record; tell the provider of any risk (Art. 26(5)).

OUR READING

A screen that was never tested with a swapped name has no known bias, which is not the same as none.

Sources: 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

Proposal of the editors, modelled on Wilson & Caliskan (2024), who also changed CV length and found it altered results. Since July 2026, Article 4a lets employers using high-risk AI process sensitive data to detect bias, under strict conditions; it creates no duty to do so. The structured interview is in No. 65.

TOOL OF THE ISSUE

A card for a CV screen

One card for each system that reads applications before a person does. The ATS CV tool shows the other side: how application software reads a CV, and what a candidate can do so it is read correctly.

01 System and its intended purpose built for recruitment, or a general tool we turned to it?

02 What it does in the process targets ads, filters, ranks or scores; where a person comes in

03 Who sees the rejected a name, how many are reviewed, who may override

04 Last swapped-name test date, signals and combinations tested, result, sent to provider

05 Our input data ads and criteria we feed it: relevant to the job?

06 Candidates told and explained where they learn AI is used; who answers a request

Sources: 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

The card is the editors' proposal, built on Articles 14, 25, 26 and 86 of the AI Act and on the method of Wilson & Caliskan (2024); it is not a tested instrument and not legal advice. The employer's duties before use, such as informing workers' representatives, are in No. 41.

Open the tool: ATS CV stivencatalyst.com/tools/ats-cv/

SOURCES AND METHOD

Every figure has a source.

The figures in this issue come from the sources below. The year shows how recent each one is.

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

EDITORIAL METHOD

Each figure was checked for its year, its publisher and what it measures, in the full text where it could be opened. A source read only in a summary, an abstract or excerpts is marked “via”. A figure not confirmed in what was read is marked so on its page, or left out. The editor's interpretation is marked “Our reading”.



MANAGEMENT REVIEW

Management without theatre.

Every issue, one management question, checked against the best research.

ALL ISSUES

stivencatalyst.com/magazine/management-review.html

Management Review · No. 106 · November 2026 · Stiven Catalyst
Edited by Stiven Janaqi, who answers for it; “the editors” in this publication means him. Its texts, translations, layout and the checking of figures against their sources are done in working sessions with Claude, an AI model by Anthropic.
stivencatalyst.com/how-this-is-made.html