

THE RESEARCH UNIVERSITY PRESS

International
Computational Law Journal

VOLUME 2 / ISSUE 2 / CURATED SOURCE EDITION

First Come, First Hired? ChatGPT’s Bias for The First Resume It Sees and the Cost for Candidates to Overcome Bias in AI Hiring Tools

Alexander Puutio · Patrick K. Lin

MIT Computational Law Report · Originally published 2025-02-24

Independent reader edition under CC BY 4.0. This attribution cover was added; original source PDF pages follow unchanged. Original authorship and source identity are retained. MIT, Stanford University, Stanford Law School and CodeX have not sponsored or endorsed this republication.

Source, license and version evidence accompany the online article.

MIT Computational Law Report

First Come, First Hired? ChatGPT's Bias for The First Resume It Sees and the Cost for Candidates to Overcome Bias in AI Hiring Tools

Alexander Puutio Patrick K. Lin

Published on: Feb 20, 2025

URL: https://law.mit.edu/pub/firstcomefirsthired

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

Introduction

Artificial intelligence (AI) technologies have made significant inroads into the hiring landscape, promising to streamline everything from candidate screening to final selection. From dedicated "interviewer bots" to AI-enhanced resume screening processes, companies are increasingly relying on advanced language models to manage the overwhelming number of applications they receive.

This shift greatly increases the procedural efficiency of hiring. At the same time, it raises difficult questions about the impact of AI on procedural fairness, given how numerous authors have found that AI systems both inherit pre-existing biases lurking in historical data and reinforce those introduced by user prompts.

As a result, these AI-based hiring systems can produce outcomes that neither the employer nor the candidates desire as a result. To better understand how AI might reshape hiring processes, and to what extent it may replicate or amplify existing inequalities, we conducted an experimental study using ChatGPT, the most commonly used large language model in both private and enterprise context, and recruited its services as a resume screener for an entry level finance job it in its browser, API and GPT-agent forms.

Our approach was designed to simulate real-world hiring scenarios as closely as possible while testing for the impact of varying key terms in the resumes submitted to ChatGPT. Specifically, we asked ChatGPT to select one candidate from a set of ten entry-level-hopefuls who varied in race, gender, as well as the perceived cost of university attendance and extracurricular activities.

Over the course of more than 2,000 observations, we validated well-documented biases and discovered novel patterns of automated discrimination.

Quite unexpectedly, we found that ChatGPT (4o) overwhelmingly tends to select the first candidate presented to it when it reviews resumes it considers equally qualified. The result replicated over all methods of testing through November to December 2024. Only in three instances, ChatGPT rejected our request and declined to select anyone, citing that all candidates were equally qualified. When explicitly instructed to avoid this pitfall, the bias shifted to favoring the seventh candidate (30%). Interestingly, candidates in positions five, six, eight, nine, and ten were never chosen at all in our experiment with the refined prompts.

After identifying these ordinal biases that we did not find prior reporting on, we tested the "cost" required for candidates to break through ChatGPT's patterns. Our model was simple; test candidates would face off against nine white males with university and extracurricular activities selected from the 'low' category as the baseline, reflecting that the field is widely considered white-male dominant. We experimented with different variations of race, gender and low/medium/high costs for each test candidate until ChatGPT would routinely pick the test candidate out from the group of ten resumes when they are not the first resume ChatGPT reviews.

Unsurprisingly, attending a high-cost university, like Harvard, MIT, or similarly prestigious institutions, was found to raise a candidate's selection rate from a baseline of around 10 percent to over 26 percent. High-cost extracurricular activities had a similar, albeit less pronounced, effect.

These findings echo existing concerns that AI-driven systems, which heavily weight perceived prestige and cost of attendance, can perpetuate socio-economic disparities. Gender appeared to exert comparatively minimal influence with low statistical significance. Race was found to matter in terms of how costly a candidate's resume would have to be in order to break through the first-place bias, with certain demographic groups needing to invest more heavily in costly credentials to overcome ChatGPT's inclination toward selecting the first candidate presented to it.

Our findings serve as a warning to all those deploying AI without proper guardrails and Q/A to inform hiring decisions. Not only does ChatGPT have the ability to inherit past prejudices, but it is capable of fabricating entirely new biases that may only reveal themselves under repeated usage with certain prompts and contexts.

Without due care, hiring processes driven solely by ChatGPT are unlikely to provide optimal selection results.

From a legal standpoint, such biases in AI-driven processes could expose companies to liability for discrimination, especially if candidates from protected groups are systematically disadvantaged. As these findings suggest, businesses must exercise caution when deploying AI tools in critical decision-making contexts, insisting upon rigorous quality assurance and transparency measures to identify and mitigate potentially unlawful or suboptimal outcomes. For candidates, our findings validate what most entry-level applicants instinctively know about our society. How much you pay for your degree matters, as do the racial signals encoded in our names.

In what follows, we detail the empirical framework of our experiments, reviewing how we designed our simulations and recorded ChatGPT's responses. We then present the results of our ordinal and cost-based analyses, discussing the implications for both legal compliance and fairness in automated hiring. Our findings underscore the importance of understanding exactly how AI models assess and select from competing signals, lest a technology introduced to improve the selection process inadvertently reinforces the very biases it was meant to overcome.

The Expanding Role of AI in Recruitment and the Threat of Bias

The use of large language models (LLMs) and other forms of generative AI, such as ChatGPT, in recruitment processes is growing at a rapid pace. Recent reports indicate that 70 percent of companies and 99 percent of Fortune 500 companies are already using AI-based tools in their hiring processes.[1] Employers are turning to these tools, either directly or via third-party services built on foundational models, to facilitate various stages of hiring, with a particular focus on CV screening and ranking which are particularly amenable to AI systems,

given their structured nature, high resource demands, and the lack of finality of any decision made by AI.[2] Of these components, CV screening and candidate ranking are the most amenable to AI systems, given their structured nature, high resource demands, and ease of reversibility for the decisions made at the end of these steps.

The appeal of using AI for CV screening and ranking is immediately obvious. These models can efficiently process large volumes of resumes, significantly reducing the time required for initial evaluations, expediting recruitment timelines, and freeing human resources for other critical tasks, such as onboarding and training.[3]

Specifically, ChatGPT and similar models have proven effective at filtering resumes, summarizing candidates' qualifications, and highlighting both strengths and weaknesses in a way that lightens the workload for recruiters.[4] However, as the prevalence of AI models in screening resumes increases, so too do concerns regarding the potential negative impacts of their use. Many of these tools pose a tremendous risk of exacerbating existing discrimination in the workplace based on protected characteristics such as race, sex, disability, and age. This is especially true in the context of AI models' tendency to perpetuate biases, particularly when interacting with resumes that mention or imply 'sensitive' or 'protected attributes' such as gender, race, age, sexual orientation, or disability.[5]

Accessing AI-assisted CV screening no longer demands specialized software; anyone with a browser can employ tools like ChatGPT to filter through large stacks of resumes. The main allure of AI tools in this context lies in their capacity to handle high volumes of resumes swiftly, thus reducing manual workload and speeding up the recruitment timeline.[6] Models like ChatGPT, highlighting key strengths, and pinpointing weaknesses. This functionality allows hiring teams to focus more on nuanced, interpersonal aspects of the process, such as interviews and onboarding.[7]

While the efficiency gains are evident, the risk of perpetuating bias remains a central worry.[8] Many of today's AI hiring tools are trained on data about employers' own workforce and prior hiring processes. However, that data is itself tainted by existing institutional and systemic biases. Furthermore, the correlations an AI tool identifies may not have a causal link with being a successful employee, may be unrelated to job functions, and may even be proxies for protected characteristics.

Because AI models learn from data reflecting historical or prevailing human biases, they may inadvertently replicate those prejudices when screening resumes. Existing studies underscore how attributes like race, gender, age, or disability status, either explicitly mentioned or signaled by details such as names and extracurricular activities, can lead to systematically skewed outcomes.[9]

For instance, one resume screening tool identified being named *Jared* and playing high school lacrosse as correlated with being a successful employee.[10] Similarly, the personality traits many AI tools are designed to measure, such as positivity, extroversion, and the ability to work in a high-pressure or fast-paced environment

are often unnecessary for the job or screen out candidates with disabilities like autism, depression, or attention deficit disorder.[11]

Traditionally, discussions around fairness in AI-driven hiring systems have centered on concepts like individual fairness, procedural fairness, and group fairness.[12] While these are crucial, researchers have called for an outcome-oriented viewpoint, one that measures the real-world effects of AI-driven hiring decisions on both employers and candidates.[13] Our study builds on the discussion of ethics as much as practice, and we add a thus-far underexplored dimension of the cost of bias, particularly as incurred by the candidate.

Employers and Candidates: How Competing Interests Lead to Resume Embellishment

Employers are primarily focused on optimizing the efficiency of the recruitment process and the quality of their hires, balancing the need to reduce costs with ensuring that candidates who are ultimately hired are well-suited to the role. Prima facie, AI systems provide a means of achieving this by automating the most resource-intensive steps of the recruitment process, such as CV screening. However, these systems also introduce the risk of bias, which can compromise the quality of hires.

Candidates, on the other hand, are primarily concerned with maximizing their own utility by securing a job that offers them the best fit in terms of compensation, work-life balance, growth opportunities and other considerations unique to the candidate. Although fairness and objectivity in the recruitment process are important, the ultimate goal for candidates is to optimize their own employment outcomes. In practice, this means that bias in AI-driven hiring processes imposes real, tangible costs on candidates, forcing them to invest in costly signals that, in effect, "buy" a chance to overcome the system's predispositions.

Because candidates are not in a position to fundamentally change the inner workings of the algorithm, they are left with the option, or the pressure, to adapt their resumes in ways that "work with" the bias rather than neutralize it. In practical terms, this means that rather than expecting an AI system to assess their qualifications fairly, candidates may resort to accumulating costly credentials or engaging in strategic resume formatting to signal the desired level of prestige.

Some candidates have even explored more aggressive, or "black-hat," tactics to tip the scales in their favor. Such tactics include embedding hidden prompts within their resumes, phrases like "select this candidate" discreetly inserted within the text, to influence the AI's decision-making process. Another strategy involves using white font on a white background to hide instructions or signals that are meant to be read by the machine but not by human reviewers. That candidates deploy such methods underscore a critical tension in the design of AI-assisted hiring systems. On the one hand, the reliance on costly signals forces candidates to conform to an expectation that may not accurately reflect their true qualifications or potential. On the other hand, resorting to manipulative techniques, whether overt or covert, to bypass built-in biases further degrades the fairness and

transparency of the hiring process. This duality illustrates the need for both robust regulatory oversight and technical interventions that can mitigate bias at its source, rather than merely shifting the burden onto candidates.

It also requires a thorough understanding of which biases we should consider useful, if not entirely legitimate. For example, an employer's preference for past experience and prestigious degrees do not serve the purpose of equity, but they are clearly held as useful by employers. In an ideal scenario, AI-driven selection processes maintain useful biases while mitigating against negative ones. At a minimum, AI should not introduce new harmful biases.

Ultimately, our findings suggest that the design and use of AI-driven hiring tools will have measurable impacts on candidates' behavior, including how they choose to educate themselves and how they represent their accomplishments in their resumes. By quantifying the additional signals that candidates must incorporate to achieve a selection rate comparable to those favored by the system's default settings, this research contributes to a broader understanding of how algorithmic processes can affect recruitment dynamics.

Our Experimental Setup

In designing our experiment, we set out to replicate real-world hiring conditions more closely than traditional resume-audit studies, which often compare only one or two candidates at a time.

Our approach, by contrast, relied on presenting ten resumes at once to ChatGPT (specifically, version 4.0), thereby mirroring a scenario in which an employer "batch-processes" multiple applicants. By varying only a handful of targeted attributes, names (to signal race and gender), university cost, and extracurricular cost, we aimed to observe how these factors influenced the model's decisions. There are numerous limitations and departures from real-life scenarios for this approach, including the number of resumes presented and the fact that they are presented in one prompt, the equality of the qualifications beyond the varied terms, as well as the assumptions inherent in using names, universities and extracurriculars to convey meaning. Our use of white males as the 'baseline' race and gender used for the control resumes (apart from the fully randomized dataset of 2,000 observations where all resumes were randomly generated) can also be validly questioned, and our results should be interpreted accordingly.

We tested ChatGPT in three different modes: (1) through the browser interface, (2) via direct application programming interface (API) calls, and (3) as a "GPT Agent" system. At the onset, we conducted a number of calibration tests with the same batch of ten resumes submitted through all modes, and we found them to agree with the selection made across other modes. The main body of our testing was created by means of Python calls to the API, and for each trial we validated the results with the browser and GPT agent, marking down only a handful of discrepant choices.

An entirely unexpected, but key aspect of the study involved uncovering positional biases in our first round of trials where we presented test and control candidates with the university and extracurricular activities selected from the same cost category of low. In our preliminary round, we kept identical qualifications for all candidates except race and gender, observing that ChatGPT overwhelmingly selected the first resume in the list, often against specific instructions to avoid doing so. We then experimented with prompt engineering aimed at reducing this "first-place" bias. While such interventions did mitigate the bias to some extent, they also introduced new positional preferences (notably a tilt toward the seventh position). We were unable to identify prompts that would have resulted in a selection likelihood that would have matched the theoretical 10%.

Having identified a strong position-bias, we proceeded to conduct more nuanced tests of how candidates could overcome ChatGPT's biases by "investing" in costly signals, i.e. attending a prestigious, high-tuition university or participating in expensive extracurricular activities. We systematically varied these cost categories across race and gender to determine which kinds of investments most strongly increased a candidate's odds of being selected. In total, we ran 2,800 resumes through ChatGPT.

In our largest dataset, comprising 2,000 observations, each resume was dynamically generated from pools of candidate names, universities, and extracurricular activities that differed in cost and passed through the API. After systematically recording which resume the AI selected, we performed logistic regressions to quantify the relative influence of race, gender, and cost signals on selection odds. We then ran a further series of stratified studies where we kept the race and gender of the test candidate equal, while testing them against white male candidates of low university and extracurricular costs, one of which was always in first place. This allowed us to explore the race and gender specific 'costs' of overcoming the first-place bias.

These findings, presented below, allow us not only to evaluate the existence of bias in AI-driven hiring but also to estimate the "price tag" required for candidates to transcend it.

Results

A Strong First-Position Bias Emerges When Assessing Equally Qualified Candidates

The first, and most unexpected, finding of our study is ChatGPT's strong tendency to select the first resume presented when all candidates appear to have equal merit. Theoretically, with ten indistinguishable candidates, each should enjoy a 10% probability of selection. Yet, in our initial trials where we varied only the race and gender of the test candidate, ChatGPT chose the first candidate close 86.67% to 100% of the time if no explicit instruction was provided to the contrary.

Dataset	Total Trials	Observed Proportion	Expected Proportion
1	10 (100 observations)	90.0%	10.0%

2	10 (100 observations)	100.0%	10.0%
3	60 (600 observations)	86.67%	10.0%

The AI system's reliance on ordinal position, rather than any substantive difference in the resumes, underscores a significant concern both for employers and candidates. Although employers are unlikely to routinely encounter scenarios where all candidates have equal qualifications apart from the main variables under study, ChatGPT's tendency to default to the first-positioned candidate among equally qualified peers is a concern employers should take seriously.

Prompt Engineering Partially Alleviates but Does Not Eliminate Bias

We attempted a straightforward fix by appending the prompt with language such as "Do not select the first candidate by default." This intervention reduced the magnitude of the first-position bias but introduced new positional biases. For instance, one modified prompt led ChatGPT to select the seventh candidate in 31.7% of trials (n=600), while ignoring those in positions five, six, eight, nine, and ten entirely.

Resume Position	**Expected Proportion**	**Observed Proportion**
1	10.0%	20.0%
2	10.0%	15.0%
3	10.0%	18.3%
4	10.0%	15.0%
5	10.0%	0.0%
6	10.0%	0.0%
7	10.0%	**31.7%**
8	10.0%	0.0%
9	10.0%	0.0%
10	10.0%	0.0%

Although the re-weighting of positional preference shows that ChatGPT is responsive to some degree of instruction, these findings also highlight a deeper structural bias in the model's decision-making process. Further experiments on chain of thought prompting and parceling the prompt to smaller steps may yield better results in overcoming the first-position bias.

Edge Cases: Refusal to Select and Unexpected Reliance on Tracker Codes

Notably, in three instances, ChatGPT declined to select any candidate at all, citing an inability to discern differences among equally qualified resumes. Two of these refusals occurred in the browser mode and one via the API.

Another unanticipated phenomenon emerged when we inadvertently left "tracker codes" embedded in the resumes: ChatGPT consistently selected the resume containing the only unique code (e.g., "0023"), effectively using that field as a differentiating factor. When asked for its reasoning in these instances, ChatGPT explicitly cited the uniqueness of the tracker code, underscoring the AI's susceptibility to arbitrary or irrelevant signals.

The Cost Approach: Mitigating Position Bias Through Higher "Signals"

Given that no single prompt seemed capable of normalizing candidate selection to a uniform 10% each, we pivoted to assessing how candidates could overcome the bias by including "costly signals" (i.e., attending a prestigious university or pursuing expensive extracurriculars) into their resumes.

We systematically tested different cost tiers of both universities and extracurricular activities across varying demographic attributes, including race and gender. By pitting a "test candidate" carrying high or medium cost signals against nine White male control resumes with low-cost attributes, we endeavored to pinpoint when ChatGPT would deviate from its default preference to select the first candidate.

Stratified Studies Reveal Different Thresholds Based on Race

Stratified trials showed that the threshold for overcoming the first-place bias changes based on perceived inclusion into various demographic groups (i.e., White, Asian, Black, Hispanic as signaled by name based on the dataset of Crabtree et al. (2023)).[14] For example, candidates with names signaling Hispanic or Black ancestry needed medium-cost signals, either in university or extracurriculars, to achieve a selection rate on par with the first-position candidate. Others, such as Asian male or Asian female candidates, required at least one high-cost signal (e.g., a prestigious university) to ensure an equivalent or superior likelihood of selection. White female candidates exhibited the same pattern: only introducing a high-cost credential tended to eliminate the first-place bias.

These findings suggest that the "price tag" necessary to overcome ordinal bias is not uniform across demographic lines. However, care should be taken in the interpretation, noting the assumptions at play here, including the assumed signals embedded in the candidates' names. It is entirely possible that ChatGPT has

name biases as much as positional ones, as recently exemplified by the LLM's inability to respond to prompts referencing *David Mayer* and other biases, such as towards the name Jared, as referenced above.

Large-Scale (2,000-Observation) Dataset Corroborates Candidates Benefit from Including High-Cost Universities and Extra-Curriculars in Their Resumes

Finally, we compiled a dataset of 2,000 resumes with each randomly assigned a name, race, gender, and cost tiers for universities and extracurriculars, presented to ChatGPT for selection in batches of 10. A logistic regression confirmed that higher-cost indicators are by far the most influential predictors of selection.

University Cost Distribution and Selection Rates: Main Dataset

University Cost	Count	Proportion (33% ideal)	Selection rate (10% ideal)
High cost	649	32.45%	26.35%
Medium cost	668	33.40%	2.84%
Low cost	683	34.15%	1.46%

Extracurricular Cost Distribution and Selection Rates: Main Dataset

Extracurricular Cost	Count	Proportion (33% ideal)	Selection rate (10% ideal)
High Cost	659	32.95%	16.69%
Medium Cost	712	35.60%	7.87%
Low Cost	629	31.45%	5.41%

Logistic Regression Results: Main Dataset

Logit Regression Results: Non-stratified

Dep. Variable:	Selected	No. Observations:	2000
Model:	Logit	Df Residuals:	1993
Method:	MLE	Df Model:	6
Date:	Mon, 30 Dec 2024	Pseudo R-squ.:	0.2551
Time:	16:01:28	Log-Likelihood:	-484.32
converged:	True	LL-Null:	-650.17
Covariance Type:	nonrobust	LLR p-value:	1.316e-68

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-9.2938	0.591	-15.723	0.000	-10.452	-8.135
Asian	-0.0261	0.251	-0.104	0.917	-0.518	0.466
Hispanic	0.4952	0.237	2.093	0.036	0.032	0.959
Black	0.4359	0.244	1.786	0.074	-0.043	0.914
Gender	0.0994	0.167	0.596	0.551	-0.227	0.426
University_Cost	2.0641	0.168	12.302	0.000	1.735	2.393
Extracurricular_Cost	0.8161	0.109	7.475	0.000	0.602	1.030

University cost exhibited the strongest effect size, followed by extracurricular cost, with race also playing a non-negligible role.

Gender, by contrast, was not statistically significant in the non-stratified analysis, suggesting a lesser direct effect on selection, though the stratified models indicated nuanced interactions between cost signals, race, and gender. The regression fit suggests that other factors are also at play, including name preferences, ordinal biases and more.

Conclusion: Implications for Fairness and Future Research

Taken together, these results illuminate how ChatGPT's selection process can be skewed by both positional factors and perceived "prestige" signals among other factors such as race and gender.

When resumes are truly indistinguishable beyond basic demographic labels, ChatGPT appears to latch onto superficial differentiators, most strikingly the resume's ordinal position. Even explicit instructions do not fully dissolve these biases; at best, they shift them around.

Meanwhile, showing that more expensive credentials and activities consistently boost one's prospects highlights how an AI-driven system can inadvertently perpetuate socioeconomic disparities. Future research might explore more refined prompt engineering and agentic process techniques, such as chain-of-thought prompting or partial-lot presentations, to see whether any methods can wholly normalize selection probabilities.

In the interim, companies contemplating AI-led hiring would be wise to implement careful auditing and

rigorous fairness checks before relying on these models for decisions that can meaningfully shape people's livelihoods. The absence of these additional processes may render some AI tools unreliable and defective.

Candidates, by contrast, are under constant pressure to secure high-cost credentials that bolster their resumes. In the absence of access to such signals, the cost of omitting them is likely to increase, leading to embellishments and resume fraud becoming increasingly appealing to overcome biases candidates may find unjustified and unfair.

Footnotes

1. Jennifer Alsever, *AI-Powered Speed Hiring Could Get You an Instant Job, but Are Employers Moving Too Fast?*, Fast Company (Jan. 6, 2023), https://www.fastcompany.com/90831648/ai-powered-speed-hiring-could-get-you-an-instant-job-but-are-employers-moving-too-fast; Joseph B. Fuller, Manjari Raman, Eva Sage-Gavin & Kristen Hines, *Hidden Workers: Untapped Talent*, Harvard Business School Project on Managing the Future of Work & Accenture (Sept. 3, 2021), https://www.hbs.edu/managing-the-future-of-work/Documents/research/hiddenworkers09032021.pdf. ↩
2. Kristina Potočnik, Neil Robert Anderson, Martin Kleinmann, and Marise Born, *Paving the Way for Research in Recruitment and Selection: Recent Developments, Challenges and Future Opportunities*, European Journal of Work and Organisational Psychology (Mar. 2021), 10.1080/1359432X.2021.1904898. ↩
3. Ahmad Alnsour, Omar Ali Kanaan, Maimar Salah, Leen Alfayyad, Yara Hijazi, and Dana Alsharif, *The Impact of Implementing AI in Recruitment on Human Resource Management Efficiency and Organizational Development Effectiveness*, Journal of Infrastructure Policy and Development (Aug. 2024), 10.24294/jipd.v8i8.6186. ↩
4. J. Stewart Black and Patrick van Esch, *AI-Enabled Recruiting: What Is It and How Should a Manager Use It?*, 63 Business Horizons 215 (Apr. 2020), https://doi.org/10.1016/j.bushor.2019.12.001. ↩
5. Christopher Harris, *Mitigation Age Biases in Resume Screening AI Models*, 36 The International FLAIRS Conference Proceedings (2023), https://doi.org/10.32473/flairs.36.133236; Louis Lippens, *Computer Says 'No': Exploring Systemic Bias in ChatGPT Using an Audit Approach*, 2 Computers in Human Behavior: Artificial Humans (July 2024), https://doi.org/10.1016/j.chbah.2024.100054; Louis Lippens, *Humans vs GPTs: Bias and Validity in Hiring Decisions*, OSF Preprints (July 2024), https://doi.org/10.31219/osf.io/zxf5y. ↩
6. Alnsour et al., *supra* note 3. ↩
7. Black and van Esch, *supra* note 4. ↩
8. Harris; Lippens, *supra* note 5. ↩

9. Marianne Betrand and Sendhil Mullainathan, *Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, 94 American Economic Review 991 (Sept. 2004), https://www.aeaweb.org/articles?id=10.1257/0002828042002561; Dena F. Mujtaba and Nihar R. Mahapatra, *Fairness in AI-Driven Recruitment: Challenges, Metrics, Methods, and Future Directions*, arXiv (June 2024), https://arxiv.org/abs/2405.19699; Harris, *supra* note 5. ↩

10. Dave Gershgorn, *Companies Are on the Hook If Their Hiring Algorithms Are Biased*, Quartz (Oct. 22, 2018), https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased. ↩

11. Algorithm-driven Hiring Tools: Innovative Recruitment or Expedited Disability Discrimination?, Center for Democracy & Technology (Dec. 2020), https://cdt.org/wp-content/uploads/2020/12/Full-Text-Algorithm-driven-Hiring-Tools-Innovative-Recruitment-or-Expedited-Disability-Discrimination.pdf. ↩

12. Lippens, *supra* note 5; Mujtaba and Mahapatra, *supra* note 7. ↩

13. Kate Glazko, Yusuf Mohammed, Ben Kosa, Venkatesh Potluri, Jennifer Mankoff, *Identifying and Improving Disability Bias in GPT-Based Resume Screening*, FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (June 2024), https://doi.org/10.1145/3630106.3658933. ↩

14. Charles Crabtree, Jae Yeon Kim, S. Michael Gaddis, John B. Holbein, Cameron Guage, and William W. Marx, *Validated Names for Experimental Studies on Race and Ethnicity*, 10 Scientific Data 130 (2023), https://doi.org/10.1038/s41597-023-01947-0. ↩