

**New York City Council
Committee on Technology**

Ethical Implications of Using Artificial Intelligence and Automated Decision Systems

November 13, 2020

Written Testimony of

Dr. Sarah Myers West, Postdoctoral Researcher, AI Now Institute.

Good afternoon Chairman Holden and members of the Committee on Technology. My name is Dr. Sarah Myers West and I am a Postdoctoral Researcher at the AI Now Institute, an interdisciplinary research institute at New York University that focuses on the social implications of artificial intelligence. I respectfully submit the following testimony on Int. 1894-2020, the “Fair Shot Act”, a law that would regulate the sale of automated employment decision tools.

The City Council’s scrutiny of this space is particularly important in a moment where the pandemic has introduced unprecedented challenges for workers in the city. As a growing number of New Yorkers are grappling with the prospect of unemployment, it is critical that employers and technology companies be held accountable for ensuring that workers’ rights are protected, both in their hiring practices and in the development, acquisition and use of technologies that mediate their relationships with current and future workers.

The automated employment decision tools that the Fair Shot Act aims to regulate are already in wide use across a range of industries and job categories, from tools used to source potential candidates, to the software used to screen them, such as systems that automatically scan the resumes of applicants, to gamified assessments designed to evaluate personality traits for managerial roles. Even in the absence of clear standards of oversight and evaluation, these systems are already being used to make important decisions throughout the hiring process, from who gets targeted with a job ad, to who may be called in to an interview, to what salary might be offered to a candidate.¹ Job candidates are often unaware when these systems are in use, what qualifications were taken into account in making decisions about hiring, what bias a given system might encode, or why they didn’t get the job, and are thus unable to identify or marshal evidence to identify when discrimination takes place - let alone aggregate the data across multiple candidates to necessary to challenge it.

Even more, we lack basic evidence to understand whether and how these tools work, to what extent they are effective in their evaluation, and the risk of errors influencing their outcomes. What we do know, according to research by a team of scholars at Cornell, is that while these companies express a sensitivity to issues of bias and discrimination there is a worrying lack of well-defined best practices as to appropriate methods for debiasing or audits, and with no transparency about the validity of their assessments, assuming a system is assessed at all.² Indeed, there is currently no standard for algorithmic

¹ Bogen, M. and Rieke, A. (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. *Upturn*. Retrieved from <https://www.upturn.org/reports/2018/hiring-algorithms/>.

² Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2019). Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3408010.

bias audits more generally, and the field of research examining these issues is still very young. Other studies indicate significant doubt as to whether these systems work as advertised, and even more concerningly, that they may in fact introduce other, new forms of employment discrimination.³

In July, the AI Now Institute joined 23 other civil rights, employment and privacy organizations by signing a set of Civil Rights Principles for Hiring Assessment Technologies. These principles state that hiring assessments:

- Should not discriminate
- Should measure traits and skills that are important to job performance
- Should be transparent to job applicants
- Should be thoroughly and regularly audited
- Should be subject to meaningful oversight by state and federal regulators

For these reasons we are particularly supportive of the City Council's attention to this important issue.

Building on these principles, this testimony makes two primary points: one, that this is a space in urgent need of increased accountability and oversight. Two, bias as it surfaces in these tools cannot be separated out from historic and present-day patterns of employment discrimination - and in fact, research suggests that these tools could introduce new forms of bias. Thus, while in its intent the Fair Shot Act addresses an area sorely in need of close scrutiny, we are concerned that at present this bill could provide a rubber stamp for the perpetuation of discriminatory hiring practices, in a way that makes hiring discrimination much harder to see and could, in the end, could even compound its effects.⁴

Automated Employment Decision Tools Deserve Greater Scrutiny.

First, automated employment decision tools deserve greater scrutiny than they are currently subject to. We know next to nothing about how the models used in automated employment decision tools work and their outcomes, beyond the marketing claims provided by the companies. We have reasons to doubt these claims: one such company, Hirevue, is currently facing charges in front of the Federal Trade Commission of engaging in unfair and deceptive business practices.⁵ Opening up these tools to expert review and scrutiny is fundamental to any meaningful program of accountability and oversight.

Given their significant impact, candidates and employers alike deserve independent, third party evaluations of these hiring systems. But as written, the bill doesn't provide workers the information that they need in order to challenge these tools' use. Many job applicants are evaluated by automated employment decision tools without ever knowing that they were used, or if a human even looked at their

³ Kim, P. (2017). Data-Driven Discrimination At Work. *William & Mary Law Review*, 48, 857-936. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2801251

⁴ Ibid.

⁵ Electronic Privacy Information Center. (2019). In Re HireVue. Retrieved from https://epic.org/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf

application. This diminishes workers' rights, leaving them unable to request reasonable accommodations or to make corrections when errors are introduced in their data.

As we highlighted in the 2019 AI Now report, it is important to note that it is employers, not workers, who are the "customers" that these companies seek to court with promises of efficiency and fewer worries about accountability and liability. In fact, several prominent vendors of these tools actively offer to cover any of their customers' legal fees or liabilities that might arise from the use of their products or services.⁶ If a goal of this legislation is to combat longstanding patterns of employment discrimination, it is critical that it be designed to serve the needs of those who are discriminated against, rather than the firms responsible for perpetuating discriminatory hiring practices.

Bias Mitigation Efforts Cannot Be Separated from Other Forms of Employment Discrimination

Transparency and disclosure is an important step toward ensuring accountability in the use of automated employment decision tools, but it is only the first of many.⁷

Some of the companies selling these systems say that they conduct audits that demonstrate their validity. But in many cases we lack sufficient information about how these audits work in practice to make an independent assessment of whether and to what extent they are effective in mitigating discriminatory outcomes.⁸ The gold standard for auditing would be conducted by an independent third party who can provide an impartial and accountable view into the system.⁹ But even still, it's important to note that the field of algorithmic auditing is very young, and there is currently no accepted standard to which such audits adhere, with different approaches often using very different mathematical definitions of "fairness" and "bias," some that may capture the kinds of employment discrimination these tools may perpetuate, and some that may not.

Moreover, in order to be effective, audits need to take into account the actual data used by each employer, and each model, in order to surface potential forms of discrimination. A black box audit of a system without the actual data used both to train a given model, and as input on the employer side, will be able to tell us very little about its real-world activity - this is roughly akin to fumbling in the dark without a flashlight.¹⁰ By design, algorithmic systems mirror patterns identified in large sets of data. In practice, this means that the introduction of automated employment decision tools into a company's hiring practices

⁶ Crawford, K. et al. (2019). AI Now 2019 Report. Retrieved from https://ainowinstitute.org/AI_Now_2019_Report.pdf.

⁷ Civil Rights Principles for Hiring Assessment Technologies. (2020). Retrieved from <https://civilrights.org/resource/civil-rights-principles-for-hiring-assessment-technologies/>.

⁸ Sanchez-Monedero, J., Dencik, L. and Edwards, L. (2020). What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from <https://arxiv.org/pdf/1910.06144.pdf>; Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2019). Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3408010.

⁹ Civil Rights Principles for Hiring Assessment Technologies. (2020). Retrieved from <https://civilrights.org/resource/civil-rights-principles-for-hiring-assessment-technologies/>.

¹⁰ Raji, I.D. and Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. Association for the Advancement of Artificial Intelligence. Retrieved from https://dam-prod.media.mit.edu/x/2019/01/24/AIES-19_paper_223.pdf

will be likely to exacerbate, rather than diminish, historical patterns of employment discrimination reflected in the data a given model is trained on, reflecting and amplifying existing conditions of workplace inequality.¹¹ Pre-sale audits don't enable this kind of assessment, let alone account for the kinds of ongoing tweaks and adjustments that are involved in calibrating many of automated employment decision tools once they are in use in a given firm or institution.

On their own, bias mitigation techniques such as auditing won't be enough to prevent employment discrimination.¹² Bias in these tools cannot be separated from historical and present day employment discrimination. The criteria used to tailor the design of these systems to individual positions often still involves assessments of candidate 'fit', criteria that have long reinforced discriminatory practices.¹³ Audits should thus be holistic in nature, considering the use of automated employment decision tools not in isolation, but in relation to employers' overall hiring practices and histories.

Even if we were to take at face value the claims that 'debiasing' processes could mitigate discrimination against protected groups, we know that these biases may be reintroduced through proxy variables.¹⁴ For example, zip code can serve as a proxy for race due to historical patterns of housing discrimination. As the Civil Rights Principles highlight, "Machine learning algorithms can discover subtle correlations and proxies for protected characteristics, even when they are purposefully omitted from the model-building process...Merely removing demographic data from the model-building process will not accomplish this goal."¹⁵

This is especially important because the research suggests that in fact, the use of automated employment decision tools could even introduce new forms of bias, including on the basis of protected categories like race, gender, age, and ability.¹⁶ For example, some assessments introduce a number of means to discriminate on the basis of accessibility, through interfaces that discriminate against people with

¹¹ Bogen, M. and Rieke, A. (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. *Upturn*. Retrieved from <https://www.upturn.org/reports/2018/hiring-algorithms/>.

¹² Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2019). Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3408010.; Corbett-Davies, S. and Goel, S. (2018) The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning. arXiv preprint. Retrieved from <https://arxiv.org/abs/1808.00023>.

¹³ Sanchez-Monedero, J., Dencik, L. and Edwards, L. (2020). What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from <https://arxiv.org/pdf/1910.06144.pdf>; Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2019). Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3408010.

¹⁴ Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2019). Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. ACM Conference on Fairness, Accountability, and Transparency (FAT*), 2020. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3408010.

¹⁵ Civil Rights Principles for Hiring Assessment Technologies. (2020). Retrieved from <https://civilrights.org/resource/civil-rights-principles-for-hiring-assessment-technologies/>. See also: Bogen, M. and Rieke, A. (2018). Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias. *Upturn*. Retrieved from <https://www.upturn.org/reports/2018/hiring-algorithms/>.

¹⁶ Kim, P. (2017). Data-Driven Discrimination At Work. *William & Mary Law Review*, 48, 857-936. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2801251; Ajunwa, I. (2018). Age Discrimination by Platforms. 40 *Berkeley J. Emp. & Lab. L.* 1 (2019), Retrieved from: <https://ssrn.com/abstract=3142979>; Whittaker, M. et al. (2019). Disability, Bias and AI. *AI Now Institute*, Retrieved from <https://ainowinstitute.org/disabilitybiasai-2019.pdf>.

disabilities such as blindness, deafness or speech disorders, or through the use of algorithmically-mediated personality tests that screen out people with disabilities related to mental health and autism.¹⁷ Facial and voice analysis technologies work less well for people of color, English speakers with non-native accents, and trans people.¹⁸ Automated employment decision tools may also exacerbate the digital divide, by necessitating that job candidates have the ability to connect to high-speed internet in order to complete an assessment, or have access to a smartphone or another expensive device.

In conclusion, greater transparency around the development, data, use, and logics of these automated hiring tools is urgently needed, not only by employers, but by workers and by regulators. Without such transparency, it will be impossible to ensure accountability and effective oversight of a space of growing importance to the hiring landscape. And while audits can be a useful tool, it is important that, at a minimum, they be conducted by independent third parties, incorporate the evaluation of the actual data used by developers and employers, and that they go beyond statistical evaluations to account for how automated employment decision tools are incorporated into hiring practices. Even still, auditing and mandating disclosure of the use of automated employment decision tools will not be enough to combat discrimination in the hiring process, and in some instances, the use of these tools risks making discrimination worse.

Given their prevalence, it is deeply concerning how little we know about whether automated employment decision systems work, let alone what kinds of harms they introduce. In a moment where workers are facing increasing precarity in the wake of the pandemic, the City Council's intervention on their behalf is badly needed. It's likewise critical that regulation of this space be designed to provide the support that workers will need in protecting themselves against employment discrimination.

¹⁷ Fruchterman, J. and Mellea, J. (2018). Expanding Employment Success for People with Disabilities. *Benetech*. Retrieved from <https://benetech.org/about/resources/expanding-employment-success-for-people-with-disabilities/>

¹⁸ Civil Rights Principles for Hiring Assessment Technologies. (2020). Retrieved from <https://civilrights.org/resource/civil-rights-principles-for-hiring-assessment-technologies/>.