NEW YORK UNIVERSITY

**Julia Stoyanovich**
*Assistant Professor of Computer Science and Engineering, Tandon School of Engineering*

*Assistant Professor of Data Science, Center for Data Science*

**Department Computer Science and Engineering**
**Tandon School of Engineering**
2 MetroTech Center
Brooklyn, NY 11201
**P:** 212 998 3251
**M:** 917 470 8199
stoyanovich@nyu.edu

**Testimony of Julia Stoyanovich before New York City Council Committee on Technology regarding Int 1894-2020: "Sale of automated employment decision tools"**

**November 12, 2020**

Dear Chair Holden and members of the Committee:

My name is Julia Stoyanovich. I hold a Ph.D. in Computer Science from Columbia University. I am an Assistant Professor of Computer Science and Engineering at New York University's Tandon School of Engineering, and an Assistant Professor of Data Science at the Center for Data Science. I am the founding Director of the Center for Responsible AI at NYU, together with my colleague Steven Kuyan. In my research and public engagement activities, I focus on incorporating legal requirements and ethical norms, including fairness, accountability, transparency, and data protection, into data-driven algorithmic decision making.[1] I teach responsible data science courses to graduate and undergraduate students at NYU.[2] Some of the students who took my course are here today. Most importantly, I am a devoted and proud New Yorker.

I would like to applaud the Committee on Technology for their sustained efforts to regulate the use of automated decision systems (ADS) in New York City. The bill we are discussing today represents a potentially transformative opportunity to make the use of ADS in a crucial domain – hiring and employment – responsive to the needs of *all* New Yorkers. I am speaking here in strong support of the bill.

This bill cannot be more timely: The Covid-19 pandemic is hitting members of minority and historically disadvantaged groups particularly hard, with many losing their jobs and being unable to re-enter the workforce. If this bill passes, it will *benefit job seekers*, by ensuring that the unaccountable use of algorithmic decision making in hiring does not further exacerbate these inequities. The bill will also *benefit vendors* of hiring ADS, by helping create an economically and ethically sustainable ecosystem of technological innovation. Finally, the bill will *benefit employers* who use these tools, by helping them evaluate the claims made by vendors during procurement (through auditing), and build trust of job seekers and employees (through public disclosure).

---

[1] See https://dataresponsibly.github.io/ for information about this work, funded by the National Science Foundation through NSF Awards #1926250, 1934464, and 1922658.
[2] All course materials are publicly available at https://dataresponsibly.github.io/courses/

Despite my strong support for the bill, I am of the opinion that much additional work is needed to ensure that, when passed into law, this bill gives rise to auditing and public disclosure procedures that are more than a rubber stamp.

Hiring ADS are a quintessential example of systems that are as impactful – for individuals, population groups, and society at large – as they are contraversial. Vendors of these tools frequently and confidently make claims that, because humans are known to have biases, algorithmic tools are our only viable option. It is dangerous to take such claims on faith: *algorithmic systems themselves, and audits of these systems, are only as good as the standards and objectives to which we hold them*. These objectives are often in competition, and it is not up to data or algorithms to guide us on how to resolve the trade-offs that arise.

Therefore, the meaningfullness of auditing and public disclosure mechanisms for hiring ADS will hang on our ability to support a robust and open dialog about where to use and *not* to use these systems, what kinds of decisions we should be leaving up to them and what kinds are for humans to make, how to state the objectives agains which we check these systems, and how to negotiate the trade-offs that priritize one stakeholder group's intests over another.

I must counter the argument that the City is too resource-constrained, due to the COVID-19 pandemic and otherwise, to engage in the oversight of hiring ADS. We would not be comfortable selling food or medicine if we were unsure about both their actual benefits and their safety. We would not be building a bridge or allowing a car out on the street if we were unsure that these artifacts both work to their specification and are safe. *Algorithmic systems are engineering artifacts.* They have no common sense, no empathy, and no sense of humor. They cannot exercise agency or be held accountable for their actions. To think otherwise is to indulge in magical thinking. Algorithmic systems are what we make of them, and if we decide to use them, then we – all of us collectively – are responsible for how they work. *If we trust algorithmic hiring systems sufficiently to deploy them, then we must ensure both that they deliver on their clearly stated purpose and that they are safe.*

Meaningful oversight of hiring ADS is, of course, is a tall order. But the New York City I know and love does not give up in the face of a challenge. And the City government does not have to do this work alone. The academic community, including the Center for Responsible AI at NYU, are the City's disposal to help make the auditing and public disclosure requirements of the proposed bill actionable.

## In my statement today I would like to make three recommendations:

1. **Auditing**: The scope of auditing for bias should be expanded beyond disparate impact to include other dimensions of discrimination, and to also convey information about a tool's effectiveness. Audits should be based on a set of uniform publicly available criteria.
2. **Disclosure**: Information about job qualifications or characteristics for which the tool was used to screen should be disclosed to a job seeker in a manner that is comprehensible and actionable.

3. **An informed public**: To be truly effective, this law requires an informed public. I recommend that New York City invests resources into informing members of the public about data, algorithms, and automated decision making, using hiring ADS as a concrete and important example.

In what follows, I will give some background on automated hiring systems, and will then expand on each of my recommendations.

## Automated hiring systems

Since the 1990s, and increasingly so in the last decade, commercial tools are being used by companies large and small to hire more efficiently: source and screen candidates faster and with less paperwork, and successfully select candidates who will perform well on the job. These tools are also meant to improve efficiency for the job applicants, matching them with relevant positions, allowing them to apply with a click of a button, and facilitating the interview process.

In their 2018 report, Bogen and Rieke[3] describe the hiring process from the point of view of an employer as a series of decisions that form a funnel: "Employers start by *sourcing* candidates, attracting potential candidates to apply for open positions through advertisements, job postings, and individual outreach. Next, during the *screening* stage, employers assess candidates—both before and after those candidates apply—by analyzing their experience, skills, and characteristics. Through *interviewing* applicants, employers continue their assessment in a more direct, individualized fashion. During the *selection* step, employers make final hiring and compensation determinations." Importantly, while a comprehensive survey of the space lacks, we have reason to believe that automated hiring tools are in broad use in all stages of the hiring process.

The entire hiring funnel, as well as each component of the funnel, are examples of automated decision systems (ADS). These systems:

1. Process data about people, some of which may be sensitive or proprietary;
2. Help make decisions that are consequential to people's lives and livelihoods;
3. Involve a combination of human and automated decision making; and
4. Are designed with the stated goals of improving efficiency and promoting, or at least not hindering, equitable access to opportunity.

ADS may or may not use Artificial Intelligence (AI), and they may or may not have autonomy, but they all rely heavily on data.

Despite their potential to improve efficiency for both employers and job applicants, hiring ADS are also raising concerns. I will recount some well-known examples here.

---

[3] Bogen and Rieke, "*Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias"*, Upturn, (2018) https://www.upturn.org/static/reports/2018/hiring-algorithms/files/Upturn%20--%20Help%20Wanted%20-%20An%20Exploration%20of%20Hiring%20Algorithms,%20Equity%20and%20Bias.pdf

*Sourcing*: One of the earliest indications that there is cause for concern came in 2015, with the results of the AdFisher study out of Carnegie Mellon University[4] that was broadly circulated by the press[5]. Researchers ran an experiment, in which they created two sets of synthetic profiles of Web users who were the same in every respect — in terms of their demographics, stated interests, and browsing and search patterns — with a single exception: their stated gender, male or female. In one experiment, the AdFisher tool stimulated an interest in jobs in both groups. Researchers showed that Google displayed ads for a career coaching service for high-paying executive jobs far more frequently to the male group (1,852 times) than to the female group (318 times). This brings back memories of the time when it was legal to advertise jobs by gender in newspapers. This practice was outlawed in the US in 1964, but it persists in the online ad environment today.

The findings of the 2015 AdFisher study started a line of inquiry into the reasons for gender-based and other types of discrimination in online ad delivery, particularly as they pertain to access to employment and housing opportunities. The current understanding is that there is a multitude of reasons for this, including both biased training data and the advertisement targeting mechanism itself. In their comprehesive 2019 analysis of Facebook's ad delivery mechanisms, Ali et al.[6] explain:

"The enormous financial success of online advertising platforms is partially due to the precise targeting features they offer. Although researchers and journalists have found many ways that advertisers can target—or exclude—particular groups of users seeing their ads, comparatively little attention has been paid to the implications of the platform's ad delivery process, comprised of the platform's choices about which users see which ads.

It has been hypothesized that this process can 'skew' ad delivery in ways that the advertisers do not intend, making some users less likely than others to see particular ads based on their demographic characteristics. In this paper, we demonstrate that such skewed delivery occurs on Facebook, due to market and financial optimization effects as well as the platform's own predictions about the 'relevance' of ads to different groups of users. We find that both the advertiser's budget and the content of the ad each significantly contribute to the skew of Facebook's ad delivery. Critically, we observe significant skew in delivery along gender and racial lines for 'real' ads for employment and housing opportunities despite neutral targeting parameters. Our results demonstrate previously unknown mechanisms that can lead to potentially discriminatory ad delivery, even when advertisers set their targeting parameters to be highly inclusive. This underscores the need for policymakers and platforms to carefully consider the role of the ad delivery optimization run by ad platforms themselves—and not just the targeting choices of advertisers—in preventing discrimination in digital advertising."

---

[4] Datta, Tschantz, Datta, "*Automated experiments on ad privacy settings",* Proceedings of Privacy Enhancing Technology (2015) https://content.sciendo.com/view/journals/popets/2015/1/article-p92.xml

[5] Gibbs, "*Women less likely to be shown ads for high-paid jobs on Google, study shows",* The Guardian (2015) https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study

[6] Ali, Sapiezynski, Bogen, Korolova, Mislove, Rieke, "Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes" (2019) https://arxiv.org/pdf/1904.02095.pdf

As a result of this and other lines of research, the US Department of Housing and Urban Development (HUD) is currently investigating Facebook[7], and it is also reported that Google and Twitter are being probed for housing discrimination[8].

*Screening*: In late 2018 it was reported that Amazon's AI resume screening tool, developed with the stated goal of increasing workforce diversity, in fact did the opposite thing: the system taught itself that male candidates were preferable to female candidates.[9] It penalized resumes that included the word "women's," as in "women's chess club captain," and downgraded graduates of two all-women's colleges. These disparities are most likely due to gender bias in hiring exhibiting itself in the data on which the screening tool was trained: they aligned with, and reinforced, a stark gender imbalance in the workforce at Amazon and other platforms, particularly when it comes to technical roles. Interestingly, despite essentially unlimited data, computational, and human resources, Amazon was unable to fix the problem of bias in hiring by means of a purely technological intervention.

*Interviewing:* In 2014, it was reported that online personality tests, often used as part of the interviewing process, disproportionately reject candidates suffering from mental illness such as depression and bipolar disorder *even if* they have the right skills for the job.[10] There is much to be said about discrimination based on disability status, which arises due to a multitude of factors, including the under-representation of individuals with disabilities in the training and validation data, and the choice of data representations and effectiveness metrics. Importantly, this type of discrimination is notoriously difficult to detect because individuals with disabilities often do not report their disability status.

In summary, numerous cases of discrimination based on gender, race, and disability status during sourcing, screening, interviewing, and selection[11] stages have been documented in recent reports. These examples show that, if left unchecked, automated hiring tools will replicate, amplify, and normalize results of historical discrimination in hiring and employment.

---

[7] US Department of Housing and Urban Development, Office of Administrative Law Judges (2018) https://www.hud.gov/sites/dfiles/Main/documents/HUD_v_Facebook.pdf

[8] Robertson, "*HUD reportedly also investigating Google and Twitter in housing discrimination probe*", The Verge (2019) https://www.theverge.com/2019/3/28/18285899/housing-urban-development-hud-facebook-lawsuit-google-twitter

[9] Dastin, "*Amazon scraps secret AI recruiting tool that showed bias against women*", Reuters (2018) https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[10] Emerging Technology from the arXiv, "*Racism is Poisoning Online Ad Delivery, Says Harvard Professor"*, MIT Technology Review (2013) https://www.technologyreview.com/s/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor/

[11] Emerging Technology from arXiv, *"Racism is Poisoning Online Ad Delivery, Says Harvard Professor*", MIT Technology Review (2013) https://www.technologyreview.com/2013/02/04/253879/racism-is-poisoning-online-ad-delivery-says-harvard-professor/

## Recommendation 1: Expanding the scope of auditing

Bias audits should take a broader view, going beyond disparate impact when considering fairness of outcomes. Others will speak to this point, and I will not dwell on it here. Instead, I will focus on another important dimension of due process that is closely linked to discrimination – substantiating the use of particular features in decision-making.

Regarding the use of predictive analytics to screen candidates, Jenny Yang states: "Algorithmic screens do not fit neatly within our existing laws because algorithmic models aim to identify statistical relationships among variables in the data whether or not they are understood or job related.[...] Although algorithms can uncover job-related characteristics with strong predictive power, they can also identify correlations arising from statistical noise or undetected bias in the training data. Many of these models do not attempt to establish cause-and-effect relationships, creating a risk that employers may hire based on arbitrary and potentially biased correlations."[12]

In other words, identifying what features are impacting a decision is important, but it is insufficient to alleviate due process and discrimination concerns. I recommend that an audit of an automated hiring tool should also include information about the *job relevance* of these features.

A subtle but important point is that even features that can legitimately be used for hiring may capture information differently for different population groups. For example, it has been documented that the mean score of the math section of the SAT (Scholastic Assessment Test) differs across racial groups, as does the shape of the score distribution.[13] These disparities are often attributed to racial and class inequalities encountered early in life, and are thought to present persistent obstacles to upward mobility and opportunity.

Some automated hiring tools used today claim to predict job performance by analyzing an interview video for body language and speech patterns. Arvind Narayanan refers to tools of this kind as "fundamentally dubious" and places them in the category of *AI snake oil*.[14] The premise of such tools, that (a) it is possible to predict social outcomes based on a person's appearance or demeanor and (b) it is ethically defensible to try, reeks of scientific racism and is at best an elaborate random number generator.

The AI snake oil example brings up a related point: that an audit should also evaluate the effectiveness of the tool. *Does the tool work?* Is it able to identify promising job candidates better than a random coin flip? What were the specific criteria for the evaluation, and what evaluation methodology was used? Was the tool's performance evaluated on a population

---

[12] Yang, "*Ensuring a Future that Advances Equity in Algorithmic Employment Decisions*", Urban Institute (2020) https://www.urban.org/research/publication/ensuring-future-advances-equity-algorithmic-employment-decisions

[13] Reeves and Halikias "*Race gaps in SAT scores highlight inequality and hinder upward mobility*", Brookings (2017) https://www.brookings.edu/research/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility

[14] Narayanan, "*How to recognize AI snakeoil*" (2019) https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf

with demographic and other characteristics that are similar to the New York City population on which it will be used? Without information about the statistical properties of the population on which the tool was trained (in the case of machine learning) and validated, we cannot know whether the tool will have similar performance when deployed.[15]

In summary, I recommend that the scope of auditing for bias should be expanded beyond disparate impact to include other dimensions of discrimination, and also contain information about a tool's effectiveness. To support compliance and enable a comparison between tools during procurement, these audits should be based on a set of uniform criteria. To enable public input and deliberation, these criteria should be publicly available.

## Recommendation 2: Explaining decisions to the job applicant

Information about job qualifications or characteristics that the tool used for screening should be provided in a manner that allows the job applicant to understand, and, if necessary, correct and contest the information. I argued in Recommendation 1 that it is important to disclose *why* these specific qualifications and characteristics are considered job relevant.

*I recommend to build explanations for job seekers around the popular nutritional label metaphor*, drawing an analogy to the food industry, where simple, standardized labels convey information about the ingredients and production processes.[16]

An applicant-facing nutritional label for an automated hiring system should be *comprehensible*: short, simple, and clear. It should be *consultative*, providing actionable information. Based on such information, a job applicant may, for example, take a certification exam to improve their chances of being hired for this or similar position in the future. Labels should also be *comparable*: allowing a job applicant to easily compare their standing across vendors and positions, and thus implying a standard.

Nutritional labels are a promising metaphor for other types of disclosure, and can be used to represent the process or the result of an automated hiring system for auditors, technologists, or employers.[17]

## Recommendation 3: Creating an informed public

To be truly effective, proposed law relies on an informed public. Individual job applicants should be able to understand and act on the information disclosed to them. In Recommendation 1, I spoke about the need to make auditing criteria for fairness and effectiveness publicly available. Empowering members of the public to weigh in on these

---

[15] Stoyanovich and Howe, "*Follow the data: Algorithmic transparency starts with data transparency*" (2019) https://ai.shorensteincenter.org/ideas/2018/11/26/follow-the-data-algorithmic-transparency-starts-with-data-transparency

[16] Stoyanovich and Howe, "*Nutritional labels for data and models*", IEEE Data Engineering BUlletin 42(3): 13-23 (2019) http://sites.computer.org/debull/A19sept/p13.pdf

[17] Stoyanovich, Howe, Jagadish, *"Responsible Data Management*", PVLDB 13(12): 3474-3489 (2020) https://dataresponsibly.github.io/documents/mirror.pdf

standards will strengthen the accountability structures and help build public trust in the use of ADS in hiring and beyond.

I recommend that New York City invests resources into informing members of the public about data, algorithms, and automated decision making, using hiring ADS as a concrete example. This aligns with a set of recommendations by the Automated Decision Task force[18], on which I served, but we have not yet seen the City act on these recommendations.

We heard from members of the administration that public engagement activities have slowed down due to the COVID-19 pandemic. This does not have to be the case: based on our own experience, described below, there is substantial interest from the public to participate, and an opportunity to effectively use online platforms to educate and engage them. *One of the activities we have been ramping up at the Center for Responsible AI at NYU in recent months has focused specifically on public education and engagement around the use of ADS in hiring.* In collaboration with the Queens Public Library, the Center conducted a series of sessions called "Uncovering Hidden Decisions: AI in Hiring", where we gave an introduction to data, algorithms, and AI, gave actionable advice to job seekers, told participants about Int 1894, and asked for their thoughts on the bias auditing and public disclosure components of the proposed law. The final session will take place on November 17, 2020. I will be happy to share what we learned during these sessions with the Committee on Technology once the series completes.

## Conclusion

In conclusion, I would like to quote from the recently released position statement by IEEE-USA, titled "Artificial Intelligence: Accelerating Inclusive Innovation by Building Trust".[19] IEEE is the largest professional organization of engineers in the world; I have the pleasure of serving on their AI/AS (Artificial Intelligence / Autonomous Systems) Policy Committee.

"We now stand at an important juncture that pertains less to what new levels of efficiency AI/AS can enable, and more to whether these technologies can become a force for good in ways that go beyond efficiency. We have a critical opportunity to use AI/AS to help make society more equitable, inclusive, and just; make government operations more transparent and accountable; and encourage public participation and increase the public's trust in government. When used according to these objectives, AI/AS can help reaffirm our democratic values.

If, instead, we miss the opportunity to use these technologies to further human values and ensure trustworthiness, and uphold the status quo, we risk reinforcing disparities in access to goods and services, discouraging public participation in civic life, and eroding the public's trust in government. Put another way: Responsible development and use of AI/AS to further human

---

[18] See Section 2 of the New York City Automated Decision Systems Task Force Report (2019) https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf

[19] IEEE-USA, "*Artificial Intelligence: Accelerating Inclusive Innovation by Building Trust*" (2020) https://ieeeusa.org/wp-content/uploads/2020/10/AITrust0720.pdf

values and ensure trustworthiness is the only kind that can lead to a sustainable ecosystem of innovation. It is the only kind that our society will tolerate.”