

Testimony of Daniel Schwarz
On Behalf of the New York Civil Liberties Union
Before the New York City Council Committee on Technology
In Relation to Automated Decision Systems Used by Agencies

January 22, 2020



1 Whitehall Street, 3rd Fl.
New York NY 10004
nyclu.org

Donna Lieberman
Executive Director

Robin Willner
President

The New York Civil Liberties Union (“NYCLU”) respectfully submits the following testimony in relation to automated decision systems used by city agencies. The NYCLU, the New York affiliate of the American Civil Liberties Union, is a not-for-profit, non-partisan organization with eight offices throughout the state and more than 180,000 members and supporters. The NYCLU’s mission is to defend and promote the fundamental principles, rights, and values embodied in the Bill of Rights, the U.S. Constitution, and the Constitution of the State of New York. The NYCLU works to expand the right to privacy, increase the control individuals have over their personal information, and ensure civil liberties are enhanced rather than compromised by technological innovation.

Automated decision systems (“ADS”) comprise any software, system, or process that aims to automate, aid, or replace human decision-making. ADS are widely used in administering government services; allocating resources; and making inferences about individuals, groups, or places. Their ubiquity across government agencies means that ADS have the potential to impact a person’s eligibility for welfare benefits, education opportunities, and even their very liberty. In most instances these tools are deployed opaquely without regulation, transparency, impact assessments, or independent audits. If left unchecked, they risk severely undermining the civil, human, and privacy rights of New Yorkers.

Two years ago, the Council recognized the need to regulate the government use of automated decision systems, and enacted Local Law 49 of 2018 to create an ADS Task Force. Unfortunately, the final ADS Task Force report¹ offered New Yorkers little information on the actual use of ADS; as a result, we have joined other advocates and researchers in the publication of a Shadow Report that addresses some of the shortcomings and offers detailed recommendations to different stakeholders and institutions.²

¹ *New York City Automated Decision Systems Task Force Report*, November 2019, <https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf>.

² See: Rashida Richardson, ed., *Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force*, AI NOW INSTITUTE, December 4, 2019, <https://ainowinstitute.org/ads-shadowreport-2019.html>.



The public has a right to know about the automated systems used by our government, how widespread they are, and what type of decisions they make. A broad disclosure requirement for all agencies is long overdue. The NYCLU therefore offers qualified support for Intro. 1806-2019 as a first step towards accountability, equity, and due process.

The Need for Regulation of Automated Decision Systems

Government agencies justify their use of ADS by claiming that such systems allow them to provide new services or improve current services, to increase speed and efficiency, to cut costs, and for the algorithms' assumed accuracy and neutrality. While the use of computational tools undoubtedly boosts speed and scale, their accuracy and neutrality are consistently questioned by researchers and experts, despite the fact that these systems operate with little to no transparency. Many studies have challenged their opaque or "black box" operation,³ and provided evidence of harmful,⁴ discriminatory,⁵ sexist,⁶ and racist⁷ outcomes.

³ See e.g.: CATHY O'NEIL, *WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY* (2016); FRANK PASQUALE, *THE BLACK BOX SOCIETY* (2015).

⁴ See e.g.: VIRGINIA EUBANKS, *AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR* (2018); Ed Pilkington, *Digital dystopia: how algorithms punish the poor*, *THE GUARDIAN*, October 14, 2019, <https://www.theguardian.com/technology/2019/oct/14/automating-poverty-algorithms-punish-poor> (last visited Jan 15, 2020); Colin Lecher, *A healthcare algorithm started cutting care, and no one knew why*, *THE VERGE* (2018), <https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy> (last visited Jan 15, 2020).

⁵ SOLON BAROCAS & ANDREW D. SELBST, *Big Data's Disparate Impact* (2016), <https://doi.org/10.2139/ssrn.2477899> (last visited Jan 16, 2020).

⁶ See e.g.: Jeffrey Dastin, *Amazon scraps secret AI recruiting tool that showed bias against women*, *REUTERS*, October 10, 2018, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (last visited Jan 16, 2020); Galen Sherwin, *How Facebook Is Giving Sex Discrimination in Employment Ads a New Life*, *AMERICAN CIVIL LIBERTIES UNION*, <https://www.aclu.org/blog/womens-rights/womens-rights-workplace/how-facebook-giving-sex-discrimination-employment-ads-new> (last visited Jan 16, 2020).

⁷ See e.g.: Kate Crawford, *Opinion | Artificial Intelligence's White Guy Problem*, *THE NEW YORK TIMES*, June 25, 2016, <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html> (last visited Jan 15, 2020); Alistair Barr, *Google Mistakenly Tags Black People as 'Gorillas,' Showing Limits of Algorithms*, *WSJ* (2015), <https://blogs.wsj.com/digits/2015/07/01/google-mistakenly-tags-black-people-as-gorillas-showing-limits-of-algorithms/> (last visited Jan 15, 2020).



Software systems are often wrongly perceived as more neutral or offering a scientific and objective truth.⁸ Their proponents are able to make these assertions because the vast majority of ADS are opaque systems, secretly developed and silently deployed, that are shielded from independent review and scrutiny due to their proprietary nature. This secrecy obscures the potential errors, flaws, subjective decisions, personal choices, and views that find their way into these systems.

While actually obtaining access to the underlying source code for ADS is difficult and resource intensive, the public's ability to view and evaluate the code is critical to understanding the extent to which such errors occur. For example, it was revealed that a Medicaid ADS in Arkansas had failed to correctly assess care needs of patients with cerebral palsy or diabetes: a fact only discovered through lengthy litigation and subsequent disclosure of the code.⁹ And here in New York City, an independent review of the source code of a DNA analysis tool used by the office of the chief medical examiner raised serious questions about its validity, including whether the code may have been intentionally skewed to create more matches.¹⁰

Many automated systems purport to predict the future by observing the past. Among them are “risk assessment tools,” designed to use past policing and court data to “predict” the future behavior of an individual criminal defendant. Specifically, risk assessment tools attempt to determine which attributes are shared by people who previously failed to show up to court. Certain weights are placed on each of the attributes to produce a formula and “score” a person's future risk of flight. For instance, in 2019, the city recently revamped its old pretrial risk assessment tool to develop a new one based on a dataset of cases from 2010 to 2014. Risk assessment tools reflect a troubling philosophy toward criminal justice policy: Using past cases to determine what might happen in future cases disregards time-specific influences that may have affected prior case outcomes, and freezes a government judgment in the realities of the past. Critically, it also strips the person who is awaiting trial of independent agency and the ability to make the case that they will appear in court.

But even those who philosophically agree with using past statistics to predict future individual human behavior acknowledge that the value of such a predictive system lies in the value of the data input into it. When an ADS deploys machine learning that relies on large historic datasets to train the

⁸ danah boyd & Kate Crawford, *Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon*, 15 INFORMATION, COMMUNICATION & SOCIETY 662–679 (2012).

⁹ Litigating Algorithms 2018, AI NOW INSTITUTE, <https://ainowinstitute.org/litigatingalgorithms.pdf>.

¹⁰ Lauren Kirchner, *Thousands of Criminal Cases in New York Relied on Disputed DNA Testing Techniques*, PROPUBLICA (2017), <https://www.propublica.org/article/thousands-of-criminal-cases-in-new-york-relied-on-disputed-dna-testing-techniques> (last visited Jan 16, 2020).



underlying models, the quality of that underlying data is of paramount importance. If that data includes false or biased data, every output will repeat this pattern and in turn result in false and biased decision-making. In the context of policing, utilizing data from unconstitutional and racially biased stop-and-frisk practices by the NYPD will create outputs reflecting these practices.¹¹ This behavior is commonly known by the computer-science idiom “garbage in, garbage out,” or in this scenario, as Sandra Mayson coined, “bias in, bias out”.¹² In another recent example, researchers discovered that a widely used health care algorithm used to identify patients’ health risks failed to identify many Black patients, making them less likely to be enrolled for medical treatment.¹³ And where these systems operate in the dark, people may not even realize that they are suffering at the hands of a flawed machine-learning system: one ADS in Indiana blocked hundreds of thousands of people from receiving vital support services and left them struggling to challenge these decisions.¹⁴

Given these enormous human impacts that automated systems make on our community – and the very real possibility of simply automating existing human error and bias – meaningful regulation is the bare minimum our democracy demands. The growing power imbalance between people affected by ADS and those who deploy them is at its height when affected people aren’t even aware that their lives have been changed by an ADS. Access to information on what systems are in use, whether their accuracy has been studied and their impact assessed, and the mechanisms to obtain redress for harm is essential for the public to be able to engage in a fully-informed discussion regarding what role—if any—these systems should have in government decision making.

¹¹ Rashida Richardson et al., *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. REV. ONLINE 192 (2019), <https://ssrn.com/abstract=3333423>.

¹² Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE LAW JOURNAL (2019), <https://www.yalelawjournal.org/article/bias-in-bias-out> (last visited Oct 28, 2019). Archived at: <http://archive.is/nzP1D>.

¹³ See: Beth Haroules & Simon McCormack, *How an Algorithm Puts Black People’s Health in Danger*, NEW YORK CIVIL LIBERTIES UNION (2019), <https://www.nyclu.org/en/news/how-algorithm-puts-black-peoples-health-danger> (last visited Jan 15, 2020); Ziad Obermeyer et al., *Dissecting racial bias in an algorithm used to manage the health of populations*, 366 SCIENCE 447–453 (2019).

¹⁴ Alyssa Edes & Emma Bowman, “Automating Inequality”: *Algorithms In Public Services Often Fail The Most Vulnerable*, NPR.ORG (2018), <https://www.npr.org/sections/alltechconsidered/2018/02/19/586387119/automating-inequality-algorithms-in-public-services-often-fail-the-most-vulnerab> (last visited Jan 16, 2020); Virginia Eubanks, *We created poverty. Algorithms won’t make that go away*, THE GUARDIAN, May 13, 2018, <https://www.theguardian.com/commentisfree/2018/may/13/we-created-poverty-algorithms-wont-make-that-go-away> (last visited Jan 16, 2020).

Use of Automated Decision Systems in New York City Agencies

In 2016 ProPublica published *Machine Bias*, a seminal report on the disparate impact of the risk assessment tool COMPAS.¹⁵ It inspired then-Councilmember James Vacca to introduce legislation with the intent to mandate transparency and accountability requirements for the use of ADS in city government.¹⁶ Initially proposing broad transparency requirements, the legislation was later amended to create an ADS Task Force and was enacted as Local Law 49 of 2018.¹⁷ Its mandate included creating criteria for identifying which ADS should be subject to regulation and oversight; procedures by which a person affected by ADS receive an explanation of that decision; procedures to determine whether an ADS disproportionately impacts people based upon protected status and how to address such bias; a process for making information publicly available to enable public oversight of government use of ADS; and procedures for archiving ADS and related data. Although an important first step in addressing government use of ADS, the Task Force's mandate was very modest, and included no ability to actually audit ADS and a wide carve-out for law enforcement use.



Despite the Task Force's limited mandate, it represented the first real attempt at establishing a formal oversight mechanism over ADS in the United States. Unfortunately, the Task Force's final report, published in November 2019, fell short of these expectations. It did not examine any individual ADS and limited its recommendations to broad guidelines. Non-governmental Task Force members requested access to, but were blocked from reviewing, specific ADS. And disappointingly, the Task Force completely missed the opportunity for broad public education and community engagement envisioned by its enacting legislation. The NYCLU and our partners repeatedly sought to offer input and recommendations to the Task Force, including through open letters

¹⁵ Julia Angwin et al., *Machine Bias*, ProPublica (2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

¹⁶ Lauren Kirchner, *New York City Moves to Create Accountability for Algorithms*, ProPublica (2017), <https://www.propublica.org/article/new-york-city-moves-to-create-accountability-for-algorithms>.

¹⁷ The New York City Council - File #: Int 1696-2017, <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0>.



in January 2018,¹⁸ August 2018,¹⁹ March 2019,²⁰ and finally in a comprehensive Shadow Report in December 2019.²¹

A major recommendation from the Task Force report was the establishment of a “centralized ADS Organizational Structure within City government”. The Mayor created this structure, titled the Algorithms Management and Policy Officer (AMPO),²² through an Executive Order at the time of the report publication.²³ Beyond having the practical effect of further postponing any actual release of information to the public on the use of ADS in city government and punting on setting up a system for democratic oversight, this order also severely limited the AMPO’s effectiveness by including a broad “public safety” carve-out. We can expect the NYPD to interpret this exemption in the broadest sense, interfering with the important oversight function. But it is precisely the Police Department’s use of ADS that warrants the greatest scrutiny due to the potential for police ADS systems to rely on data resulting from unconstitutional and racially biased policing. Especially when an encounter with law enforcement can result in the deprivation of one’s liberty or, in the most tragic of cases, the loss of one’s life, the public deserves to know whether police are making enforcement decisions based on flawed or biased data.

Even though ADS have faced greater scrutiny in the last few years and the field of study has significantly grown, it remains difficult to identify the full scope of ADS in use. Much of what we know about their use in New York City is pieced together from disparate sources such as public records requests, litigation, procurement data, employee information, and press statements. It is safe to assume that ADS are used by virtually all City agencies, including:

- the New York City Department of Education for teacher evaluations and student placements;
- the NYPD for predictive policing, the gang database, the case recommendation tool Patternizr, automated license plate readers, social media monitoring, and facial recognition;

¹⁸ Letter to Mayor de Blasio: Regarding NYC Automated Decision Systems Task Force, NEW YORK CIVIL LIBERTIES UNION (2018), <https://www.nyclu.org/en/publications/letter-mayor-de-blasio-regarding-nyc-automated-decision-systems-task-force> (last visited Jan 16, 2020).

¹⁹ Open Letter to Automated Decision Systems Task Force, NEW YORK CIVIL LIBERTIES UNION (2018), <https://www.nyclu.org/en/publications/open-letter-automated-decision-systems-task-force> (last visited Jan 16, 2020).

²⁰ Letter to the Automated Decision Systems Task Force - March 1, 2019, NEW YORK CIVIL LIBERTIES UNION (2019), <https://www.nyclu.org/en/publications/letter-automated-decision-systems-task-force-march-1-2019> (last visited Jan 16, 2020).

²¹ *See supra* 2.

²² *New York City Algorithms Management and Policy Officer*, <https://www1.nyc.gov/site/ampo/index.page>.

²³ The City of New York Office of the Mayor, Executive Order 50, November 19, 2019, <https://a860-gpp.nyc.gov/bitstream/gpp/24020/1/Executive%20Order%2050%20-%20Algorithms%20Management%20and%20Policy%20Officer.pdf>.



- the Mayor's Office of Criminal Justice for their failure-to-appear tool;
- the New York City Fire Department to anticipate where fires may spark;
- the New York City Department of Health to identify serious pregnancy complications;
- the Administration for Children's Services for child welfare, ecomap and genogram software;
- the Center for Innovation through Data Intelligence to predict families' risk of homelessness and to identify buildings that are likely to house at-risk families;
- the New York City Human Resources Administration to identify fraudulent benefit recipients;
- the New York City Housing Authority to predict malfunctions in apartments and buildings and report alerts in real time;

...and many other instances. Without giving the public tools to know that these systems even exist and to provide them with the information needed to assess their usefulness and impact, we are in grave danger of outsourcing government decision-making to ever-more opaque tools that could automate bias and strip us of our most fundamental rights.

In November 2018, New York City joined the Cities Coalition for Digital Rights and signed its Declaration. It clearly states that people have “sovereignty over their data, including the right to know what happens to their data, who uses it and for what purposes. [...] Everyone should have access to understandable and accurate information about the technological, algorithmic and artificial intelligence systems that impact their lives, and the ability to question and change unfair, biased or discriminatory systems.”²⁴ We urge the Council to uphold this promise by enacting legislation that will serve our democratic values and create the regulatory mechanisms necessary to protect against harmful and discriminatory algorithms.

Intro 1806-2019 - Reporting on Automated Decision Systems Used by City Agencies

The legislation would require city agencies to provide basic information about every automated decision system in use. The disclosure requirement would include what each automated decision system is intended to measure or reveal, a description of the decisions made, the name of the entity that developed the ADS, and how long the system has been in use. The Mayor’s Office of Operations would be required to compile the information received by city agencies and report it to the Mayor and the Speaker of the Council every year.

²⁴ Declaration of Cities Coalition for Digital Rights, https://citiesfordigitalrights.org/assets/Declaration_Cities_for_Digital_Rights.pdf.



This legislation is a first step toward closing the overwhelming information gap around the use of ADS in New York City. New Yorkers currently lack even the most basic information about what these systems are and how public officials are using them. A disclosure requirement will help the public and policymakers alike understand the current terrain, craft better and more targeted oversight mechanisms, aid people in finding help when they feel they are unfairly impacted by a decision, and drive public education opportunities. The sponsor of the 2017 ADS oversight bill, former Councilmember James Vacca, has already publicly endorsed the legislation as further realizing the intent of his original bill.²⁵ And some agencies have already shown that this is a workable model and that it is feasible to release information about their use of ADS. The Criminal Justice Agency published details about their Risk Assessment tool on their website allowing the public to understand the scoring.²⁶ However, the Criminal Justice Agency should go further and make sure that the thresholds and specific data used to determine their risk calculations are made public.²⁷

It is worth noting that the proposed legislation defines automated decision systems very broadly. As drafted, it would cover a myriad of software tools, scripts, and processes. Though likely unintended, this could include search scripts, automated software updates, virus scanners, and other programs. This over-inclusivity could make the disclosure requirement unworkable for agencies to compile, and tedious for the public to review. We therefore recommend adding a narrow carve-out that would exclude certain tools:

“ADS do not include: 1. routine software tools for internal cybersecurity procedures such as update schedulers, anti-virus, and network security, or 2. routine software tools for data back-ups, retention, and deletion.”

In any definition of ADS adopted by the Council, it is imperative that the disclosures at a minimum include information on when and whether such systems are making decisions that impact the lives and rights of New Yorkers – and where they do, require detailed documentation and racial equity analyses.

Intro. 1806 has the potential to accomplish what the Task Force failed to do by affirming New Yorkers' right to know what types of automated decision systems are being used by city agencies and how these tools are

²⁵ Noel Hidalgo, *Council Member Peter Koo Introduces Bill to Require Reporting on Automated Decision Systems (ADS) used by City Agencies*, BETANYC (2019), <https://beta.nyc/2019/11/27/council-member-peter-koo-introduces-bill-to-require-reporting-on-automated-decision-systems-ads-used-by-city-agencies/> (last visited Jan 16, 2020).

²⁶ *Release Assessment*, New York City Criminal Justice Agency, <https://www.nycja.org/release-assessment/3704>.

²⁷ The Leadership Conference on Civil and Human Rights, “The Use of Pretrial Risk Assessment Instruments: A Shared Statement of Civil Rights Concerns” (2018), available at <https://civilrights.org/edfund/pretrial-risk-assessments/>.

impacting their lives and environments. Similarly, the Council will benefit from these disclosures and be better able to fulfill its oversight role over City agencies and be better capable of pursuing further legislation on ADS going forward.

The proposed legislation is a modest, but necessary first step to protecting civil liberties and civil rights as the City moves its operations into the Digital Age.

Intro 1447-2019 - Annual Inventory of Agency Data



The NYCLU supports the underlying purpose of Intro. 1447, but we believe the bill's numerous exceptions will prevent it from achieving that goal. This bill would require the Director of the Office of Data Analytics to submit an annual report to the Mayor and the Speaker of the Council describing data collected and maintained by mayoral agencies. For each data set the report would include descriptions of the content, the agency collecting or maintaining it, which agencies have access, whether it is available on the open data portal, collection method used, update frequency, and size.

The NYCLU agrees with the need for public reporting of agency data collections. Unfortunately, however, the legislation includes an exception broad enough to critically undermine its potential utility. The legislation includes a broad carve-out for information whose disclosure: "(i) is expressly prohibited under federal, state law or local law; (ii) would pose a risk to individual privacy, public safety, or the cyber-security of agency systems; or (iii) would violate a confidentiality agreement or trade secret protection." This carve-out would allow for vast exemptions in the disclosure of covered data sets. Most disappointingly, we could expect that reliance on these exemptions would be most pronounced in the very areas like policing that would benefit the most from public reporting. Thus, reporting required by the bill would not add significantly to the already existing Open Data Law (Local Law 11 of 2012). For this legislation to actually succeed in increasing transparency, this exception must first be removed. Otherwise, it is doubtful whether it would give the public a better understanding of what data agencies collect, retain, share, and act on.

Conclusion

We thank the Committee for the opportunity to provide testimony and for recognizing the need for oversight and regulation of automated decision systems. The NYCLU urges the Council to pass Intro. 1806 to enhance transparency around automated decision systems as a first step for accountability, fairness, due process, and the protection of New Yorkers' civil rights and liberties.