Before the
National Institute of Standards and Technology

In the Matter of Artificial Intelligence Risk Management Framework	Docket No. NIST-2021-0004

Comments of

COLOR OF CHANGE

Gabrielle M. Rejouis
Color Of Change
gabrielle.rejouis@colorofchange.org
colorofchange.org

Filed September 15, 2021

Table of Contents

Introduction

In its request for comments, the National Institute of Standards and Technology (NIST) asks for input on the development of the Artificial Intelligence Risk Management Framework (AI RMF).[1] Color Of Changes urges NIST to take into consideration how proactive algorithmic accountability will protect Black communities and increase trust in artificial intelligence (AI). NIST must disrupt trends to use algorithms to augment and hide discrimination. Humans are deploying artificial intelligence that magnifies the scale of discrimination we've witnessed in the past. Biased algorithms disadvantage Black people and other protected classes in the following areas:

- Housing
- Education
- Healthcare
- Employment

Firms must suspend the use of AI found to discriminate until the bias is removed. The AI RMF must do the following to guide firms in this antidiscrimination evaluation:

- Require proactive algorithmic accountability
- Direct firms to frequently evaluate their systems
- Create neutral, robust standards to evaluate for bias

Firms that design and develop AI (AI developers) and firms that use AI (AI users) must be held accountable for how AI harms Black communities. Algorithmic accountability requires more than an audit of these systems; it requires the thorough examination of the system for bias, assessment of where bias might enter or be amplified, and formation of guardrails to prevent discriminatory effect.

I. Algorithms are making important life decisions and cannot be trusted to do so if there is undetected and unaddressed bias in their decision-making.

AI deployers are using algorithms to determine critical points in the public's life. These important life decisions include ones made about an individual's health, housing, education,

[1] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework - Docket No. NIST-2021-0004, Federal Register, August 24, 2021, https://www.federalregister.gov/documents/2021/08/24/2021-18108/artificial-intelligence-risk-management-framework.

employment, credit, and insurance. These decisions impact an individual's livelihood and way of being. There are civil rights or privacy laws protecting almost every one of these life decisions.[2] By automating decisions in these sectors, AI has considerable power over the individuals who need these eligibility determinations. AI developers and users are not considering how this automation severely disadvantages Black people, gender minorities, and other protected classes. Stories of discriminatory algorithms and disparate impact underscore the need for algorithmic accountability.

A. Mortgage algorithms bar Black homebuyers from purchasing homes.

Mortgage lenders use algorithms that discriminate against Black homebuyers by denying them mortgages they qualify for or giving them higher interest rates than their White peers. A mortgage approval algorithm denied the McDanielses, a Black couple, a mortgage despite the fact that they both made six figures and had excellent credit.[3] The McDanielses should have qualified for the loan. But lending officials did not approve the couple's application after an algorithm denied their request around fifteen times. The Markup's investigation found that mortgage lenders were 80 percent more likely to reject Black applicants.[4] Even if a Black homebuyer is approved for a loan, there is still opportunity for bias to make home buying more burdensome. In a similar study, lenders who used AI charged Black and Latino homebuyers between 5.6 and 8.6 percent higher interest rates on their mortgage.[5] Algorithms are making it harder for Black people to purchase homes and increase their wealth.

---

[2] "Constitutional Amendments and Major Civil Rights Acts of Congress Referenced in Black Americans in Congress," History, Art, and Archives, House of Representatives, https://history.house.gov/Exhibitions-and-Publications/BAIC/Historical-Data/Constitutional-Amendments-and-Legislation/; Thorin Klosowski, "The State of Consumer Data Privacy Laws in the US (And Why It Matters)," *The New York Times,* September 6, 2021, https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/; Equal Credit Opportunity Act, 15 U.S.C. §§ 1691-1691f.

[3] Emmanuel Martinez and Lauren Kirchner, "The Secret Bias Hidden in Mortgage-Approval Algorithms," The Markup, August 25, 2021, https://themarkup.org/denied/2021/08/25/the-secret-bias-hidden-in-mortgage-approval-algorithms.

[4] *Id.*

[5] Laura Counts, "Minority homebuyers face widespread statistical lending discrimination, study finds," Haas School of Business, University of California, Berkeley, November 13, 2018, https://newsroom.haas.berkeley.edu/minority-homebuyers-face-widespread-statistical-lending-discrimination-study-finds/.

B. Algorithms in education admissions and screening can bar Black students from opportunities.

New York City high schools that use screening algorithms admitted fewer Black and Latino students.[6] Black students are not admitted to top-performing schools at the same rate at which they apply. These screening algorithms can use factors that appear to be neutral but actually reflect racist systems. For example, when a higher rate of Black students have chronic health issues, their attendance may suffer.[7] Looking at school discipline records to predict success ignores that Black students experience disproportionately higher and harsher punishments than their peers.[8] Screening algorithms are ill-equipped to reduce racial inequalities in New York City schools.

C. The use of biased AI can deny Black patients treatment.

Medical professionals turn to algorithms to sort which patients should receive treatment, like vaccines and organ transplants.[9] However, when these decisions rely on racist assumptions and flawed science, they severely disadvantage Black patients. Jordan Crowley has a kidney disease and will need a transplant as his condition grows more severe.[10] The need for a kidney transplant increases based on a patient's estimation of glomerular filtration rate (eGFR).[11] In the United States, patients become eligible for a transplant when their eGFR falls below 20. Race plays a major factor in determining eGFR. Jordan is biracial; he has three White grandparents and a Black grandparent. If Jordan's doctors told the algorithm their patient was White, Jordan's eGFR would be 17; but because the doctors calculated eGFR for a Black patient, Jordan's eGFR

---

[6] Colin Lecher and Maddy Varner, "NYC's School Algorithms Cement Segregation. This Data Shows How," The City, May 26, 2021, https://www.thecity.nyc/2021/5/26/22453952/nyc-high-school-algorithms-segregation.
[7] James H. Price, Jagdish Khubchandani, Molly McKinney, and Robert Braun, "Racial/Ethnic Disparities in Chronic Diseases of Youths and Access to Health Care in the United States," *BioMed Research International*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3794652/pdf/BMRI2013-787616.pdf.
[8] Lauren Camera, "School Suspension Data Shows Glaring Disparities in Discipline by Race," *U.S News and World Report,* October 13, 2020, https://www.usnews.com/news/education-news/articles/2020-10-13/school-suspension-data-shows-glaring-disparities-in-discipline-by-race.
[9] Drew Harwell, "Algorithms are deciding who gets the first vaccines. Should we trust them?," *The Washington Post*, December 23, 2020, https://www.washingtonpost.com/technology/2020/12/23/covid-vaccine-algorithm-failure/; Tom Simonite, "How an Algorithm Blocked Kidney Transplants to Black Patients," *Wired*, October 26, 2020, https://www.wired.com/story/how-algorithm-blocked-kidney-transplants-black-patients/.
[10] Jennifer Tsai, "Jordan Crowley Would Be in Line for a Kidney—if He Were Deemed White Enough," Slate, June 27, 2021, https://slate.com/technology/2021/06/kidney-transplant-dialysis-race-adjustment.html.
[11] *Id.*

was 21. His doctors refused to use their discretion to give an advantage to their biracial patient. This is one example of an algorithm that does not serve Black patients well, and instead puts them at risk for prolonged suffering.

D. Algorithms in the hiring process can bury discrimination.

Employers have increasingly turned to algorithms to streamline the recruitment process.[12] But it has come at the cost of rejecting candidates who might otherwise be qualified. Amazon pulled an algorithm used to select job applicants after finding it only recommended male applicants.[13] Advocates warned about how AI can further existing bias in hiring.[14] While human evaluators can use AI thinking it eliminates bias, it may instead repackage and hide the prejudice of the recruiter.[15] Both AI developers and AI users must take care to reduce the introduction of bias in the hiring process.

The opacity of hiring algorithms makes it difficult to know why the AI rejects applicants, but the AI often magnifies existing bias.[16] Employers and colleagues create cultures that penalize women for having and caring for children.[17] When women return to the workforce after leaving to care for their families, employers are reluctant to hire them, often citing the gap on their résumés.[18] Hiring algorithms appear to replicate this bias. Sonam Oberai and Verina LeGrand both took time away from work to care for their children.[19] When they tried to return to the

---

[12] Kathryn Dill, "Companies need more workers. Why do they reject millions of résumés?" *The Wall Street Journal,* September 4, 2021, https://www.wsj.com/articles/companies-need-more-workers-why-do-they-reject-millions-of-resumes-11630728008.
[13] 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.
[14] Miranda Bogen and Aaron Rieke, Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias, Upturn, December 2018, https://www.upturn.org/reports/2018/hiring-algorithms/.
[15] *Id.*
[16] Dill, "Companies need more workers. Why do they reject millions of résumés?" *The Wall Street Journal*, https://www.wsj.com/articles/companies-need-more-workers-why-do-they-reject-millions-of-resumes-11630728008.
[17] Gayle Tzemach Lemmon, "The Pregnancy Penalty: How Working Women Pay for Having Kids," *The Atlantic,* December 13, 2012, https://www.theatlantic.com/sexes/archive/2012/12/the-pregnancy-penalty-how-working-women-pay-for-having-kids/266239/.
[18] Patricia Cohen, "A Child Care Gap in the Résumé: Whether to Explain or Not," *The New York Times,* May 20, 2016, https://www.nytimes.com/2016/05/20/business/economy/a-child-care-gap-in-the-resume-whether-to-explain-or-not.html.
[19] Dill, "Companies need more workers. Why do they reject millions of résumés?" https://www.wsj.com/articles/companies-need-more-workers-why-do-they-reject-millions-of-resumes-11630728008.

workforce, hiring AI did not advance their applications despite their having the qualifications and experience for the positions they sought.

II. Black communities need proactive detection of bias in algorithms and AI.

The public's trust in AI will be eroded by widespread and unmitigated discrimination and discriminatory impact. Communities are looking for innovation that reverses racial inequality. To cultivate the public's trust in the design, development, use, and evaluation of AI technologies, AI developers and AI users must detect and remove bias from AI. Proactive algorithmic accountability will ensure that communities are receiving fair and accurate results.

A. When discriminatory impact is a possibility, it is negligent to deploy AI without testing and correcting that bias.

AI developers and AI users should test and eradicate bias in their systems. They should take steps to prevent Black communities from experiencing bias. Innovation should not come at the cost of discrimination. Putting profit and speed over Black communities' ability to purchase homes, secure jobs, or get medical treatment can no longer be the route the AI community takes. AI developers and AI users must gauge the risk of bias and discriminatory impact in AI making important life decisions. Guidance from the AI RMF must condemn failure and refusal to conduct independent, robust evaluations of AI for bias.

B. The impacts of discrimination can be catastrophic.

The impacts of discrimination can be hard to reverse or irreversible. In healthcare, the decision to give a patient treatment can mean the difference between treating an illness or leaving a patient to suffer. Waiting can cause otherwise preventable complications or pain. When biased algorithms are widely adopted, this increases the possibility that Black communities will encounter discriminatory outputs. AI increases the scale of harm of discriminaton. The use of AI should neither be taken lightly nor implemented without precaution.

The ability to override is insufficient protection. When humans have the ability to override decisions from AI, they often do not. Jordan's doctors chose to identify their biracial patient as Black to his detriment.[20] Although Amazon managers can override automatic

[20] Tsai, "Jordan Crowley Would Be in Line for a Kidney—if He Were Deemed White Enough," https://slate.com/technology/2021/06/kidney-transplant-dialysis-race-adjustment.html.

terminations, they rarely do so to uphold the rate of work in fulfillment centers.[21] This suggests that AI will have the final say. When biased AI is deployed in important life decisions, that spells disaster for Black communities. Preemptive accountability is needed to prevent these threats.

C. AI deployers are in the best position to prevent bias and must be proactive in eradicating such threats.

Proactive algorithmic accountability is necessary to mitigate harm to Black communities because the average person is not equipped to raise concerns that AI is biased. It is unlikely that the average person will either know AI was used or that the AI returned a biased result. Firms will know if an algorithm was deployed, the purpose for the algorithm, as well its level of sophistication. By the time the average person discovers a biased algorithm was used, it is often too late. Evaluating for bias must occur before deployment. Firms must consider how seemingly neutral factors, like what cars are parked in a neighborhood, can return decisions along racial lines and discriminate.[22] Preventative algorithmic accountability reduces the risk that Black communities will experience harm.

III. AI developers and AI users must share the responsibility of auditing for bias.

Because certain types of AI continue to evolve after being introduced to the world, the responsibility of algorithmic accountability should be shared amongst AI developers and AI users. Focusing on who will be held at fault rather than how to prevent harm misses the urgent need for guidance and protection. Algorithmic accountability requires frequent and thorough assessments throughout the design, development, and use phases. NIST and the AI RMF must ensure that all participants in the process hold full responsibility to prevent algorithmic discrimination.

Managing AI models that continue learning requires continuous testing for bias. Studies found that although the designers do not direct them to nor intend them to, AI can return

[21] Colin Lecher, "How Amazon automatically tracks and fires warehouse workers for 'productivity'," The Verge, April 25, 2019, https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations.

[22] Catherine Shu, "Researchers can now make neighborhood voting predictions from Google Street View images," TechCrunch, January 1, 2018, https://techcrunch.com/2018/01/01/researchers-can-now-make-neighborhood-voting-predictions-from-google-street-view-images/.

decisions that have a discriminatory impact.[23] Our society uses racist structures that AI will replicate unless there are guardrails to prevent it. Testing should be as frequent as possible when AI routinely changes.

A. Industry claims that we must decide who should hold liability are a distraction.

Industry arguments that either AI developers or AI users must hold full liability for discrimination are a distraction. A zero-sum approach ignores how more accountability and shared collaboration between the AI developer and AI user will create more protection for Black communities. They should innovate ways to detect and eradicate bias.

Both AI developers and AI users should manage, respond to, and communicate risk with appropriate parties. For example, a firm that creates a hiring AI system has a duty to design it in a way that proactively rejects racial proxies, train it on data to reinforce racial equities, and rigorously test it for bias. The firm then has a duty to communicate all risks to the vendor who sells the tech. The vendor must independently test the algorithm against a neutral, robust testing standard for all hiring algorithms. They must also communicate any risks to the procuring AI user. The firm that purchases the tech should routinely evaluate the output of the system by comparing the applicant pool demographic to the demographic of applicants selected for the interview. The acquiring firm has a duty to report to the developing firm and vendor any issues to troubleshoot and offer appropriate remedies to applicants discriminated against. Although the actions differ, all parties must share responsibility. NIST must signal the need for a new paradigm and ensure that relevant experts and stakeholders can shape new processes.

B. The responsibility of evaluating AI systems must be shared by the developer and user of such technology.

Evaluation should occur during the AI system's development and deployment, as bias can enter at multiple points along this process.[24] Because of this, liability must be shared. AI

[23] Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, and Aaron Rieke, "Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes," Proceedings of the ACM on Human-Computer Interaction 2019, p. 11-13, September 12, 2019, https://arxiv.org/abs/1904.02095.
[24] Karen Hao, "This is how AI bias really happens—and why it's so hard to fix," *MIT Technology Review,* February 4, 2019, https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.

developers and AI users have different levels of control and information about the AI and should be responsible for preventing bias in their jurisdictions. AI developers and AI users should have incentives to be vigilant against bias. AI developers should avoid training AI in a way that will damage Black communities, such as using training data that includes racist stereotypes or insufficient data from Black people.[25] And AI users must refrain from using AI in a way that reinforces racist systems.[26]

C. Sharing liability will require all participants in the process to do their due diligence.

When there is shared liability, AI developers and AI users will search for and address bias. The shared liability framework should discourage willful ignorance. Eradicating bias is a team effort requiring all participants to do their part. This collaboration will create AI systems that the public can trust.

IV. The federal government must develop independent and robust algorithm accountability standards.

The role of developing algorithmic accountability measures must rest with the federal government because the tech industry has demonstrated that it will not proactively advance racial justice. During this era of self-regulation, tech companies have actively upheld racist structures, ignoring the impact of its technology on Black communities.[27] When Facebook was sued for violating the Fair Housing Act, it falsely claimed that Section 230 immunity made it exempt from liability.[28] Google attempted to bury a research paper that cautioned against the use of natural learning models because of the models' contributions to climate change, which would

[25] Todd Feathers, "Google's new dermatology app wasn't designed for people with darker skin," Vice, May 20, 2021, https://www.vice.com/en/article/m7evmy/googles-new-dermatology-app-wasnt-designed-for-people-with-darker-skin.
[26] Clare Garvie, "Garbage In, Garbage Out: Face Recognition on Flawed Data," Center on Privacy & Technology at Georgetown Law, https://www.flawedfacedata.com/.
[27] Amber M. Hamilton, "Silicon Valley pretends that algorithmic bias is accidental. It's not," Slate, July 7, 2021, https://slate.com/technology/2021/07/silicon-valley-algorithmic-bias-structural-racism.html.
[28] Statement of Interest of the United States of America, National Fair Housing Alliance, et. al vs. Facebook, 18 Civ. 2689 (JGK) pp. 14-17, August 17, 2018, https://www.justice.gov/crt/case-document/file/1089231/download.

disproportionately impact communities of color and communities living on low income.[29] NIST must be suspicious of any claims by the AI industry that it can adequately self-regulate. Algorithmic accountability is being treated like a luxury. Federal regulation is needed to end this trend. Allowing companies to contract auditors and report the results may lead to misrepresentation of results or even industry capture.[30] NIST should also beware of shallow attempts to address concerns.[31] Federal auditors will ensure true transparency for the public and their advocates. To develop these standards, the federal government and AI RMF must invest in both technological experts and civil rights experts.[32]

NIST must also consider how the AI RMF can advance racial equity. Black communities deserve more than AI that does not discriminate; this should be the bare minimum. Instead, algorithmic accountability should also consider how to ensure racial justice is incorporated in the development and use of AI. For instance, AI can assist with affirmative action decisions and ensure equitable distributions of employment opportunities. This will require intentional design, development, use, and evaluation of AI systems.

Conclusion

NIST must use the AI RMF to protect Black communities through robust algorithmic accountability. AI is being entrusted with more important decisions. Our communities need preventative measures to ensure equal opportunity.

---

[29] Karen Hao, "We read the paper that forced Timnit Gebru out of Google. Here's what it says," *MIT Technology Review,* December 4, 2020, https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru.
[30] Alex C. Engler, "Independent auditors are struggling to hold AI companies accountable," Fast Company, January 26, 2021, https://www.fastcompany.com/90597594/ai-algorithm-auditing-hirevue.
[31] AI Now Institute, AI Now Report 2018, p. 29, December 2018, https://ainowinstitute.org/AI_Now_2018_Report.pdf.
[32] Ashley Gold, "How team Biden could tackle tech's civil rights failings," Axios, December 18, 2020, https://www.axios.com/biden-civil-rights-facial-recognition-ai-3ecf108c-d7fa-4c7e-b4f7-e382102c340c.html.

COC Algorithmic Accountability Legislative Principles

As more and more crucial decisions are being made by algorithms—who gets approved for a home loan, what people pay for car insurance, who gets the opportunity to interview for a job—these systems could serve to expand the scale of current racial discrimination. In order to prevent replicating current inequities that harm Black people, algorithmic accountability measures should:

Require regular, public independent audits of algorithms.

- Audits must examine algorithms for discrimation and disparate impact before and during implementation.
- They must surface any discrimination that may arise due to protected class including: race, color, class, national origin, disability, gender, gender identity, sexual orientation, and criminal record.
- Like Securities and Exchange Commission financial disclosure documents, audit reports can provide information for non-government campaigns without disclosing proprietary information.

Empower the federal government to evaluate algorithms.

- Enforcement is key. Federal regulators must provide measures of redress, including suspending the use of a discriminatory algorithm and deleting the dataset used to train that algorithm.
- These enforcers should take the onus off of consumers to flag harms and ensure evaluations of algorithms are, in fact, independent.
- Staff positions at these enforcement agencies will be an alternative to private positions in AI Ethics.
- The federal government must offer strong whistleblower protections for company employees who report issues with the technology.
- Guidance must be given for smaller companies and start-ups to develop algorithms that are compliant with federal regulation. This will maintain a low barrier to entry for new and smaller firms while also keeping people safe from discriminatory algorithms.

Equip regulators with the resources and expertise they need to be effective.

- To be effective, regulators will need bigger budgets and staff.
- Enforcers will need both to hire staff with an understanding of the technology behind algorithms as well as an understanding of the potential harms to historically marginalized communities.