Unmasking Bias in Code:

Navigating the Ethical Complexities of Technological Redlining

Sam Yard

CDS DS 380: Data, Society & Ethics

December 10th, 2023

In today's tech dominated world, establishing a protocol for standardizing AI should be at the top of our checklist. Features like speech recognition and personalized online advertisements have become staples in a lot of our day-to-day lives. However, the impacts of AI aren't always positive, and as it evolves, its negative effects become more prominent. One specific ethical concern is technological redlining, which occurs when technology resources are excluded from certain communities based on characteristics such as socioeconomic status, gender identity, or disabilities.[1] IBM, an industry giant, defines artificial intelligence bias as "AI systems that produce biased results that reflect and perpetuate human biases within a society, including historical and current social inequality. *Bias can be found in* the initial training data, the algorithm, or the predictions the algorithm produces."[2] As a result, there have been many cases of unintentional discrimination. Instances of detected bias have been present in various organizations including healthcare, image generation, online advertising, and job applicant systems. AI bias has great consequences that limit opportunities for marginalized groups and maintain structural injustices. This white paper emphasizes the need for two policies and advocates for an equitable technological environment.

Technological redlining underscores the pressing need for a thorough examination of its real-world implications. This issue is exemplified by notable cases, one being the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) developed by Northpointe, a system extensively used in pretrial hearings across the United States.[3] Members of the Accountability Working Group at the Data & Society Research Institute found that "despite its intended purpose of fairly assessing the risk of recidivism, COMPAS has been shown to

---

[1] DigitalC, "DigitalC — Bridging the Educational Gap: The Impact of Digital Redlining on Urban Schools," August 28, 2023.
[2] Team and Team, "Shedding Light on AI Bias with Real World Examples."
[3] Caplan, Robyn, and Lauren Hanson. Algorithmic accountability: A Primer - Data & Society.

exhibit racial bias, disproportionately affecting certain demographics during legal proceedings."[4] Additional instances of AI redlining are evident in credit algorithms, as seen in the case of Apple's creditworthiness algorithm, which faced allegations in 2019 of gender-based discrimination despite being programmed to be "blind" to gender.[5] Users reported instances where the algorithm provided lower credit limits to women compared to men with similar financial profiles. This case especially blurs ethical lines due to how difficult it is to prove such a claim on account of creditworthiness algorithms being notoriously muddled.[6] This event only emphasizes the need for strict guidelines for algorithm auditing and the transparency of algorithmic processes. In the realm of employment and hiring, Amazon's now scrapped, 2014 AI-driven hiring platform came under scrutiny for perpetuating gender biases by exhibiting preferences for male candidates. Strategic marketing advisor for AI startups, Roberto Iriondo, analyzes this by saying, "due to the biases that the machine learning model had, most ideal candidates were generated as men, which is a reflection of the male dominance across the tech industry—therefore the data fed to the model was not unbiased towards gender equality but au contraire."[7] As a step in the right direction, New York City has passed a law mandating employers to audit software that evaluates job candidates and employees; The law took effect on January 1, 2023.[8] This is a small stepping stone towards national AI ethics laws. These case studies highlight the consequences of AI redlining and necessitate the effort to ensure the development and deployment of unbiased AI systems.

[4] Caplan, Robyn, and Lauren Hanson. Algorithmic accountability: A Primer - Data & Society.
[5] Knight, Will. "The Apple Card Didn't 'see' Gender-and That's the Problem." Wired.
[6] Nedlund, Evelina. "Apple Card Is Accused of Gender Bias. Here's How That Can Happen | CNN Business." CNN.
[7] University, Carnegie Mellon. "Amazon Scraps Secret AI Recruiting Engine That Showed Biases against Women - Machine Learning - CMU - Carnegie Mellon University."
[8] Baeza-Yates, Ricardo. "Council Post: Wielding the Double-Edged Sword of Algorithmic Auditing."

AI redlining is a serious threat to our modern society. When algorithms are trained on historical data, there is a risk that the data is biased, which could lead to unfair outcomes for certain communities. One issue that greatly affects technology users is the lack of transparency in algorithmic decision-making. This leaves affected individuals unaware of how and why the AI produced its output. The intricacy of AI systems makes it difficult to determine who or what is responsible for biased outcomes. In the words of the Dean of Law at Swinburne Law School, "Algorithms are not autonomous machine applications or processes. Instead, they are developed and programmed by people and their efficacy is determined by the quality of the design process."[9] This statement emphasizes that there are human creators behind AI systems, and insinuates that they are entirely responsible for their AI's decision process. Even so, we do not have laws in place that have procedures for these kinds of issues. These errors, whether international or not, interfere with established ethics codes such as the ACM Code of Ethics[10] and IEEE's Ethically Aligned Design[11]. Both of these codes emphasize the importance of transparency and accountability in AI systems. A particularly fitting code outlined by ACM's Code of Ethics is section 1.4, *be fair and take action not to discriminate.* This section not only promotes fairness, but urges contributors to take time to eliminate biases in their technology. Specifically, "The use of information and technology may cause new, or enhance existing, inequities. Technologies and practices should be as inclusive and accessible as possible and computing professionals should take action to avoid creating systems or technologies that disenfranchise or oppress people. Failure to design for inclusiveness and accessibility may

[9] "The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence."
[10] "The Code Affirms an Obligation of Computing Professionals to Use Their Skills for the Benefit of Society." Code of Ethics.
[11] "Ethically Aligned Design - IEEE Standards Association." ieee.org.

constitute unfair discrimination."[12] ACM's Code of Ethics clearly outlines that the model contributor is responsible for eliminating its technologies discriminatory practices, or to not deploy it at all. Addressing technological redlining requires a commitment to these ethical principles to build responsible and equitable AI technologies. The threats and consequences are clear, but the nuances of AI have made punishments difficult to assign.

To tackle artificial intelligence redlining, two policies should be considered as potential solutions. Firstly, a policy requiring continuous education and training for AI developers is crucial. This policy would mandate education about ethical considerations, bias detection, and the responsible deployment of AI systems for developers. This policy would be dedicated to fixing the problem at the source(s), the sources being the developers and companies that deploy potentially biased algorithms. Mai-Ann Nguyen, director of CrossCountry Consulting's Risk and Compliance team, says that "Oftentimes questions only get asked about an AI-enabled system *after* it has been designed, developed, and deployed — but waiting introduces unnecessary risk. A best practice is to ask questions as a team moves through each of the various stages of the AI life cycle — design, development, deployment, and monitoring.".[13] Here, Nguyen underlines the importance of staying aware of these issues during all phases of the creation process. By implementing this educational component, governments, AI specific associations, and private corporations can ensure that developers are equipped to identify and mitigate biases in their algorithms before they are deployed. Secondly, a policy that mandates tech companies to be audited for potential negative consequences of their algorithms. Even with active attempts to mitigate biases, it is possible for it to slip through the cracks. In those cases, AI auditing is a great solution. In concurrence, the IBM Data and AI Team believes that "identifying and

[12] "The Code Affirms an Obligation of Computing Professionals to Use Their Skills for the Benefit of Society." Code of Ethics.
[13] "5 AI Auditing Frameworks to Encourage Accountability | AuditBoard."

addressing bias in AI begins with AI governance, or the ability to direct, manage and monitor the AI activities of an organization."[14] Eliminating bias is an incredible challenge, but it is a challenge that needs to be taken on. Some leading tech companies have invested in algorithm monitoring in an attempt to provide their customers with transparency, but as Director of Research at Northeastern University's Institute for experiential AI stated, "without an official body to set, monitor and enforce guidelines, there can be no foolproof stopgap to ensure algorithmic integrity."[15] These regulations can be enforced by national governments, ensuring that businesses adhere to ethical standards and are held accountable for the impact of their AI technologies. By implementing these policies, we can foster a culture of responsibility, transparency, and fairness in AI development.

It is also important to consider the perspectives of potential stakeholders. In order to implement these policies, we have to appeal to the specifics of stakeholders. To do this, we can think about these issues from their points of view. For instance, developers and tech companies may have an interest in creating innovative AI solutions, but may also be concerned about regulatory burdens. Meanwhile, communities impacted by AI redlining have a shared interest in unbiased technologies that do not perpetuate institutional inequalities. To gain support for the proposed policies, we must consider the motivations and interests of relevant tech companies. Collaboration with these companies would involve highlighting how ethical AI practices enhance their reputation and gain public trust. In turn, this will allow companies to deploy more accurate and trusted models. Another effective solution to gain support from stakeholders is to create a team of people from different industries with relevant perspectives and skills. Including user advocacy groups and civil rights organizations is important for representing the interests of

---

[14] Team, Ibm Data and Ai, and Ibm Data and Ai Team. "Shedding Light on AI Bias with Real World Examples."
[15] Baeza-Yates, Ricardo. "Council Post: Wielding the Double-Edged Sword of Algorithmic Auditing." Forbes.

the communities that are affected by AI redlining. Marissa Gerchick and Olga Akselrod of the ACLU Racial Justice Program emphasize this point by saying that "technologists, civil rights advocates, policymakers, and interdisciplinary researchers should work together to ensure that algorithm audits live up to their potential."[16] Utilizing experts in relevant industries provides stakeholders with solutions from cross-industry perspectives. Additionally, partnering with educational institutions to develop and provide training programs for developers can ensure that the implementation of these programs are seamless. To increase interest in these policies, we can tailor our engagement strategies to stakeholders specific concerns and goals.

Anticipating potential critiques is crucial when advocating for policies to address AI redlining. One concern could be about the feasibility of implementing continuous education for developers and requiring AI auditing. Critics may question the practicality of enforcing educational mandates and argue that such measures might stifle innovation. In response to this, research fellows at the Brooking Institute offered that a beginning initiative could involve a bias impact statement which can help "probe and avert any potential biases that are baked into or are resultant from the algorithmic decision. As a best practice, operators of algorithms should brainstorm a set of initial assumptions about the algorithm's purpose prior to its development and execution."[17] At times, datasets can lack representation for certain groups. Ongoing education could help developers catch this in the early stages and provide additional data to the training set. Although this does not provide a solidified solution, it is a step towards equity in AI model creations. Another auditing policy could involve the process of deciding whether an algorithm is ethically sound. Auditing AI models is a highly complicated process. With that, on top of the

[16] Marissa Gerchick, Olga Akselrod. "Why Meaningful Algorithm Auditing Is Key to Protecting Civil Rights in the Digital Age: ACLU." American Civil Liberties Union.
[17] "Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms | Brookings."

pressing need for regulation, companies could be turned off to implementing AI audits.[18] With this in mind, it is important to understand a stakeholders mindset. Krishna Gade, founder and CEO of Fiddler (an AI monitoring company), says that “most CEOs don’t really know the immediate ROI for implementing ethical AI and responsible AI.”[19] He then said that he starts conversations by attempting to “convince prospective clients that monitoring their AI could create positive ROI.”[20] Gade emphasizes the importance of understanding how executives think, and framing AI auditing as a profitable investment. The cherry on top of a positive ROI is an increasingly positive reputation to gain their customers' trust. Regardless of how these policies are implemented, the push to standardize algorithm outcomes will be a cross-industry effort.

AI redlining, a popular ethical issue in the tech world, is causing real problems with biased algorithms. It affects everything, from the justice system, to credit decisions, and hiring practices. To tackle this, there are guidelines that need to be introduced. More specifically, policies that mandate developers to stay up-to-date and regularly check and monitor their AI models. This will provide more transparency for consumers, and help to gain public trust. These rules could aid in the mitigation of the negative impact of AI redlining. As tech gets more advanced, it's important to hold artificial intelligence users to high ethical standards. By enforcing transparency and ongoing learning, we can shape a future where digital tech benefits everyone fairly. Authors Michael Chui, James Manyika, and Mehdi Miremadi of the article “What AI can and can't do (yet) for your business”, state that “debiasing is proving to be among the most daunting obstacles, and certainly the most socially fraught, to date.”[21] The difficulty of

[18] “How to Get Rid of Talent Biases - Spiceworks.”
[19] Kaye, “These Startups Want to Help Prove Responsible AI Can Be Profitable.”
[20] Kaye, “These Startups Want to Help Prove Responsible AI Can Be Profitable.”
[21] Chui, Manyika, and Miremadi, “What AI Can and Can’t Do (yet) for Your Business.”

this task proves its importance. With the collaboration of cross-industry experts, we can push artificial intelligence to uphold the highest of ethical standards.

Bibliography

American Criminal Law Review | Georgetown Law. “The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence,” n.d. https://www.law.georgetown.edu/american-criminal-law-review/in-print/volume-59-number-1-winter-2022/the-solution-to-the-pervasive-bias-and-discrimination-in-the-criminal-justice-system-transparent-and-fair-artificial-intelligence/.

AuditBoard. “5 AI Auditing Frameworks to Encourage Accountability | AuditBoard,” n.d. https://www.auditboard.com/blog/ai-auditing-frameworks/#authors.

Baeza-Yates, Ricardo. “Council Post: Wielding the Double-Edged Sword of Algorithmic Auditing.” Forbes, October 5, 2023. https://www.forbes.com/sites/forbestechcouncil/2022/07/25/wielding-the-double-edged-sword-of-algorithmic-auditing/?sh=395f956a30c9.

Brookings. “Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms | Brookings,” June 27, 2023. https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.

Caplan, Robyn, and Lauren Hanson. Algorithmic accountability: A Primer - Data & Society, April 18, 2018. https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL.pdf.

Chui, Michael, James Manyika, and Mehdi Miremadi. “What AI Can and Can’t Do (yet) for Your Business.” McKinsey & Company, January 11, 2018. https://www.mckinsey.com/capabilities/quantumblack/our-insights/what-ai-can-and-cant-do-yet-for-your-business.

“The Code Affirms an Obligation of Computing Professionals to Use Their Skills for the Benefit of Society.” Code of Ethics, n.d. https://www.acm.org/code-of-ethics.

DigitalC. “DigitalC — Bridging the Educational Gap: The Impact of Digital Redlining on Urban Schools.” DigitalC, August 28, 2023. https://www.digitalc.org/line-of-sight-blog/2023/8/21/bridging-the-educational-gap-the-impact-of-digital-redlining-on-urban-schools#:~:text=Digital%20redlining%20refers%20to%20the,to%20housing%20and%20economic%20opportunities.

“Ethically Aligned Design - IEEE Standards Association.” ieee.org, n.d. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf.

ITPro Today: IT News, How-Tos, Trends, Case Studies, Career Tips, More. “5 Ways to Prevent AI Bias,” September 23, 2022. https://www.itprotoday.com/artificial-intelligence/5-ways-prevent-ai-bias#close-modal.

Kaye, Kate. “These Startups Want to Help Prove Responsible AI Can Be Profitable.” *Protocol*, January 4, 2022. https://www.protocol.com/enterprise/ai-audit-2022.

Knight, Will. “The Apple Card Didn’t ‘see’ Gender-and That’s the Problem.” Wired, November 19, 2019. https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem/.

Marissa Gerchick, Olga Akselrod. “Why Meaningful Algorithm Auditing Is Key to Protecting

Civil Rights in the Digital Age: ACLU." American Civil Liberties Union, November 30, 2023. https://www.aclu.org/news/racial-justice/why-meaningful-algorithm-auditing-is-key-to-protecting-civil-rights-in-the-digital-age#:~:text=As%20automated%20systems%20become%20more,systems%20are%20used%20in%20practice.

Nedlund, Evelina. "Apple Card Is Accused of Gender Bias. Here's How That Can Happen | CNN Business." CNN, November 12, 2019. https://www.cnn.com/2019/11/12/business/apple-card-gender-bias/index.html.

Spiceworks. "How to Get Rid of Talent Biases - Spiceworks," March 6, 2023. https://www.spiceworks.com/tech/artificial-intelligence/guest-article/audit-your-ai/#:~:text=A%20bias%20audit%20can%20clarify,tools%20they%20employ%20during%20recruitment.

Susarla, Anjana. "Forget Dystopian Scenarios – AI Is Pervasive Today, and the Risks Are Often Hidden." The Conversation, n.d. https://theconversation.com/forget-dystopian-scenarios-ai-is-pervasive-today-and-the-risks-are-often-hidden-218222.

Team, Ibm Data and Ai, and Ibm Data and Ai Team. "Shedding Light on AI Bias with Real World Examples." IBM Blog, October 16, 2023. https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/.

University, Carnegie Mellon. "Amazon Scraps Secret AI Recruiting Engine That Showed Biases against Women - Machine Learning - CMU - Carnegie Mellon University." Machine Learning | Carnegie Mellon University, n.d.

https://www.ml.cmu.edu/news/news-archive/2016-2020/2018/october/amazon-scraps-secret-artificial-intelligence-recruiting-engine-that-showed-biases-against-women.html.