



**FEBRUARY 2022**

Desiree' Orr, *Research Program Analyst*

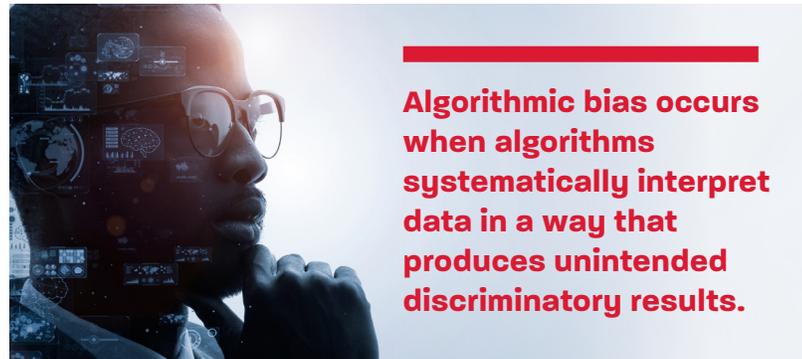
---

# The Unintended Consequences of Algorithmic Bias

---

## Introduction

The word 'algorithm' is often associated with customized news feeds on social media applications and targeted advertisements that surface in the digital space. This same machine learning technology is used to perform a myriad of tasks and make everyday decisions previously performed by human hands. From autonomous vehicles and voice-activated home systems to the deployment of automated decision systems (ADS) in education, healthcare, housing, criminal justice, and employment, machine learning is deeply intertwined with current technology and causes real-time consequences for consumers. While Artificial Intelligence (AI) and ADS technology increase productivity and spur economic growth, these innovations outpace relevant policy and national standards. Gaps in standardization reveal the lack of adequate safeguards against algorithmic bias.



## What is Algorithmic Bias?

Algorithmic bias occurs when algorithms systematically interpret data in a way that produces unintended discriminatory results. As private companies, federal agencies, and law enforcement become increasingly dependent on AI, systemic racism entrenched at the core of social institutions can be inadvertently embedded into the algorithms these entities deploy. Lack of diversity in the programming field, the unconscious bias of data scientists, and the lack of representative datasets used to train systems contribute to the creation and deployment of machine learning software that perpetuate systemic inequities. Without adequate oversight, the use of AI can cause a litany of unintended consequences, including civil rights violations, and stands to exacerbate racial discrimination by ingraining it into the digital realm.

## Who Trains these Systems?

Lack of diversity in the programming field and the subconscious preferences of those shaping these technologies play an integral role in perpetuating algorithmic bias. According to the 2014 U.S. Equal Employment Opportunity Commission Diversity in High Tech [report](#), compared to the overall private industry, the high tech sector employed a larger share of white Americans and a smaller share of African Americans. With a more diverse workforce, companies can be more inclusive in the design and implementation of algorithms, avoid perpetuating historical bias in machine learning and create products that serve the public more equally. Historical bias replicated in machine learning is perhaps the most common thread across sectors in the deployment of ADS. For example, courts across the nation use risk assessments to guide data-driven



**Lack of diversity in the programming field and the subconscious preferences of those shaping these technologies play an integral role in perpetuating algorithmic bias.**

decision-making about sentencing times, bond amounts, and other decisions regarding a defendant's freedom. These risk assessment tools predict the likelihood of a re-offense based on aggregated statistical data from samples of people in the justice system by modeling commonalities and risk factors. These assessments produce sentencing disparities that reflect historical bias resulting in harsher penalties for Black defendants. According to a [2016 report](#) by ProPublica, Black defendants were 77 percent more likely to be rated as at a higher risk of committing a future violent crime and 45 percent more likely to be predicted to commit a future crime of any kind in risk assessments. Similar historical biases are reflected in the health sector. Machine learning is used at nearly all points of the healthcare process, from diagnosis to quality of care received. A 2019 [study](#) found that a widely-used healthcare risk-prediction algorithm demonstrated racial bias as it used patients' history of healthcare spending as a proxy for future medical needs. This algorithm resulted in African Americans receiving low-quality care due to historical medical expenditures. Increased diversity in the field could allow for a more inclusive approach to algorithmic design and perhaps help to avoid the risk of perpetuating historical inequities in machine learning technology.

## What Data is Used to Train these Systems?

In addition to the need for a more diverse workforce, the lack of representative data used to train AI systems also leads to biased decision-making. Unrepresentative data can train algorithms to draw the wrong conclusions. For example, in recent years, [Amazon had to discontinue the use of one of its recruiting tools](#) because the data used to train the system was overly representative of white men resulting in disproportionate recruitment of white men. The lack of diverse and holistically representative data used to train the AI system resulted in discriminatory recruitment processes. Another example of unrepresentative data sets leading to algorithmic bias is the [deployment of facial recognition software among law enforcement](#). According to a 2019 National Institute of Standards Technology (NIST) [report](#), facial recognition systems are the least reliable when identifying people of color. Standard datasets used to train algorithms are [primarily representative of white males](#). In other cases, cameras are not sufficiently able to capture unique



facial characteristics in people of color due to the depth in skin tones resulting in low-quality database images for these groups. In these cases, algorithms are taught with insufficient data resulting in unreliable unfounded automated decisions.

## What Legislation is Protecting Consumers?

Over the past several years, legislators, civil society groups, and industry leaders have called on increased oversight and mitigation strategies to protect marginalized communities from the unintended consequences of algorithmic bias. In June 2020, the Association for Computing Machinery (ACM) U.S. Technology Policy committee released an official statement calling for the cessation of the private and federal use of AI in facial recognition software until meaningful regulation has been enacted. The Federal Trade Commission (FTC) has long held safeguards to protect consumers from algorithmic bias, such as [Section 5 of the FTC Act](#), the [Fair Credit Reporting Act](#), and the [Equal Credit Opportunity Act](#) and has recently published [guidelines](#) for businesses using machine learning technology and ADS. According to the National Conference of State Legislatures' September 2021 [report](#), general artificial intelligence bills or resolutions were introduced in 17 states in 2021 and enacted in Alabama, Colorado, Illinois, and Mississippi. While state legislatures are slowly working to regulate this technology, federal legislation is at a standstill. In May of 2021, the [Algorithmic Justice and Online Platform Transparency Act](#) was introduced to the Senate to control the discriminatory use of personal information in algorithmic processes but has not made any headway. In partnership with the FTC, legislators must uniformly work to address the ways algorithmic bias can infringe on civil rights and create enhanced oversight and regulations to protect consumers from the rampant misuse of data in public and private systems. This issue expands beyond American borders and should be assessed as a civil and human rights priority for lawmakers as AI tools are primarily designed by developers in the West and can [exacerbate inequalities](#) across the globe.

## What Next?

Technology aimed at advancing society cannot do so equitably without deliberately addressing the flaws that will perpetuate systemic racism and implicit bias. These flaws holistically threaten healthcare access, criminal justice, economic equity, and more. While many strides have been made in advancing civil rights, the digital space is a new frontier that can either expand opportunity or perpetuate historical inequalities. To rectify this deepening issue in emerging technologies, legislators must assess and address the various civil rights issues resurfacing in the digital space by providing increased consumer protections and holding private companies accountable when unlawful discrimination occurs. Technology companies must build enhanced transparency and accountability safeguards by: intentionally diversifying their workforce, implementing data representation standards with a racial justice lens, and uprooting opportunities for racial bias to be embedded in the technology they design.

