

Updates

February 17, 2023

How AI and Automated Systems Use Can Lead to Discrimination in Hiring



Artificial Intelligence (AI) and automated systems can increase efficiency and help reduce human error. However, the National Institute of Standards and Technology (NIST), the White House, and the Equal Employment Opportunity Commission (EEOC) are warning companies that uncritical reliance on AI can have legal consequences, including potentially building in bias that can lead to claims of employment discrimination. Employers' reliance on these technologies to target job advertisements, recruit applicants, train employees, and make or assist in hiring decisions can lead to adverse employment actions. But, NIST explains, "[w]ith proper controls, AI systems can mitigate and manage inequitable outcomes." The NIST study does not focus on specific legal risks arising from use of this technology, but it is useful for evaluating whether the systems meet accepted scientific standards.

NIST's Framework for Reducing the Risk of Harmful Bias in AI

NIST published the [AI Risk Management Framework](#) on January 26, 2023, with the stated intent of helping companies manage the risks associated with the use of AI systems and increase trustworthiness in AI. The framework is divided into two parts. Part 1 discusses the positive impact and negative risks of AI use. It recognizes the powerful potential of AI systems but also identifies various risks inherent in these systems, including biases that may arise in all AI systems. Part 2 outlines strategies to manage AI risks through a broad framework.

To reduce the risk of harm caused by AI, the NIST guidance explains that companies must consider and manage these biases throughout the development and implementation processes—including by recognizing and addressing systemic bias, computational and statistical bias, and human-cognitive bias. "Systemic bias can be

present in AI datasets, the organizational norms, practices, and processes across the AI lifecycle, and the broader society that uses AI systems. For example, systemic bias can be present in many 'off-the-shelf' systems that crunch data based on broader characteristics and biases present in the broader community. Computational and statistical biases can be present in AI datasets and algorithmic processes and often stem from systematic errors due to non-representative samples. AI systems that fail to adhere to basic statistical principles such as not accounting for small sample sizes are at risk for computations and statistical biases. Human-cognitive biases relate to how an individual or group perceives AI system information to make a decision or fill in missing information, or how humans think about purposes and functions of an AI system." An algorithm that promotes applicants with advanced degrees for positions that do not require advanced degrees, for example, could present risks of human-cognitive bias.

Part 2 of the framework provides a protocol for evaluating AI systems and managing risks in four broad steps: *govern, map, measure, and manage*. The core framework explains that during the development and implementation processes, companies can reduce bias by ensuring decision-making throughout each process involves a diverse team across different demographics, disciplines, experience, expertise, and backgrounds. Diverse teams could enhance an organization's ability to identify risks associated with bias before the use of AI becomes harmful.

Key Recent Governmental Guidance on Use of AI in Hiring

NIST's guidance comes on the heels of increased scrutiny of AI by the White House, the Equal Employment Opportunity Commission (EEOC), and the New York City Department of Consumer and Worker Protection (DCWP).

White House AI Bill of Rights & Discriminatory Impact

The White House Office of Science and Technology Policy (OSTP) issued the [Blueprint for an AI Bill of Rights](#) on October 4, 2022, which outlines principles to guide the design, use, and deployment of automated systems. The White House developed the AI blueprint out of concern for the harm caused by AI, including algorithms that may be "plagued by bias and discrimination." Pointing to guidance from the EEOC and the U.S. Department of Justice (DOJ), the White House explains that "employers' use of software that relies on algorithmic decision-making may violate existing requirements under Title I of the Americans with Disabilities Act (ADA)."

EEOC Provides Guidance on the Use of AI to Assess Job Applicants

According to guidance by the EEOC, published as [The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees](#), computerized applicant job assessments may make testing more difficult or reduce test accuracy for some disabled applicants. Similarly, employers' reliance on training programs delivered electronically can have a discriminatory impact on hearing and visually impaired employees. Self-paced electronic training programs may impact visually impaired employees who are unable to follow along with computer prompts without assistance. Similarly, virtually instructed training programs may affect employees with hearing impairments who are unable to follow the instructor without assisted technology.

Other discrimination claims may be triggered by AI use. Employers that rely on automated software may face claims of intentional discrimination, known as disparate treatment, under Title VII, and claims of unintentional

discrimination due to policies that seem neutral but have a discriminatory impact on protected groups. This unintentional discrimination, known as disparate impact, will likely be a focus of investigation and prosecution by enforcement agencies when evaluating cases involving AI and automated systems. For example, discriminatory impact claims could result from the use of automated software where applicants or employees are required to input data that may correlate to or be a proxy for a protected characteristic. For instance, applicant tracking software that weeds out candidates based on geographical location may inadvertently weed out candidates from racial groups that are less likely to live in the desired location.

EEOC Includes AI Focus in Strategic Plan

In line with the White House's focus on AI, the EEOC's new strategic plan includes a focus on technology-related employment discrimination. The enforcement agency "will focus on employment decisions, practices, or policies in which covered entities' use of technology contributes to discrimination based on a protected characteristic."

The strategic plan specifies the types of technologies used by employers that could be targeted, including the following:

- Software that incorporates algorithmic decision-making or machine learning (ML), including AI.
- Automated recruitment, selection, or production and performance management tools.
- Other existing or emerging technological tools used in employment decisions.

New York City Automated Employment Decision Law

New York City has passed landmark restrictions on the use of AI in hiring. Enforcement will start in April 2023. The new law will make it illegal for an employer to use an "automated employment decision tool" to screen a candidate or employee for an employment decision, unless the tool has undergone a bias audit and the results of the audit have been made public. The new law will also require employers that use automated employment decisions to notify employees and/or candidates of their use of the systems. The New York City Department of Consumer and Worker Protection (DCWP) is working on regulations and standards that will hopefully provide needed clarity on how employers are to comply.

Looking Forward: What Employers Can Do Now

Employers can reduce their risk of violating Title VII, the ADA, the Age Discrimination and Employment Act (ADEA), and other anti-discrimination statutes by identifying risks for possible intentional and unintentional discriminatory impact before implementing new technologies. Prior to implementation, employers should develop and test algorithms, automated systems, and AI to ensure that they do not improperly result in a disparate impact on protected demographics, such as a person's age, gender, sexual orientation, disability, or race.

After implementation, employers should regularly monitor the effects of the technology on protected groups. For instance, if an employer notices fewer applications from a certain protected group, it should consider whether the filters on any applicant tracking software in use are inadvertently weeding out groups with protected characteristics. Further, ensuring that these systems are accessible for persons with disabilities and do not screen them out is paramount to reduce legal risks.

Employment technologies that are most likely to be targeted include the following:

- Recruiting software.
- Online employment applications.
- Applicant screening and rating software.
- Automated learning and training programs.
- Virtual training programs.
- Onboarding software.
- Employee engagement and retention software.

Testing the software for potential discriminatory treatment and discriminatory impact, before and after deployment, is crucial. Enforcement agencies have made it clear that employers who take a "set it and forget it" attitude toward AI and automated systems when making key employment decisions are at risk if they are not evaluating whether those systems cause bias.

© 2023 Perkins Coie LLP

Authors

Explore more in

[Government Contracts](#) [Labor & Employment](#) [Artificial Intelligence & Machine Learning](#)

Related insights

Update

[**Two Tools for Trump To Dismantle Biden-Era Rules: the Regulatory Freeze and the Congressional Review Act**](#)

Update

[**The FY 2025 National Defense Authorization Act: What's New for Defense Contractors**](#)