

# WHY YOU NEED TO CARE ABOUT AI BIAS IN 2026 AND HOW A BIAS AUDIT CAN HELP YOU AVOID DANGER

Insights  
Jan 7, 2026

As employers increasingly use AI to help screen job applicants, evaluate performance, forecast risk, and support everyday decision-making, it's critical you ensure your organization doesn't accidentally introduce "AI bias" into the mix. A slate of recent lawsuits alleging that the use of AI tools may disproportionately affect certain groups should serve as a warning, even if you believe your systems are facially neutral and used in good faith. As we roll into 2026, you should make sure you understand what AI bias is, how it arises, and why you should consider a bias audit if you use AI-driven decision-making tools.

## From Data to Decisions: Inputs, Features, and Weights

It's helpful to have a quick crash course in AI in order to understand how it works and how it can inadvertently lead to biased recommendations or decisions.

- AI systems are developed and trained with *data inputs*, also called training data. This could include structured data such as resumes, performance metrics, or credit history. It could also include unstructured data such as written text, video interviews, or audio recordings.
- From these inputs, AI developers select *features*, meaning the specific variables or characteristics the system will evaluate when making a decision or recommendation. For example, features of an AI system evaluating resumes may designate education level or prior job titles as features to consider when evaluating resumes.

## Related People



**Usama Kahf, CIPP/US**

Partner

949.798.2118



**Chelsea Viola**

Associate

213.403.9626

- The system then assigns *weights* to those features, which determine how much importance each feature has in producing the final output. For example, a system may give greater weight to education pedigree than years of experience, or prioritize uninterrupted work history over skills-based assessments.

## What Do We Mean by “Bias”?

Bias refers to a systematic tendency to favor certain outcomes, characteristics, or groups over others. Bias may be intentional or unintentional, explicit or implicit. Under the legal theory of *disparate impact*, liability can arise even in the absence of discriminatory intent when a facially neutral practice produces outcomes that disproportionately affect individuals based on protected characteristics such as race, sex, age, disability, or other protected statuses.

AI systems are shaped by the data on which they are trained. When that data reflects historical bias, structural inequities, design flaws, or incomplete information, the system may replicate or amplify those patterns. As a result, AI tools can unintentionally replicate existing inequities at scale.

Even when protected characteristics are excluded from the data inputs given to an AI system, certain features may function as *proxies* for protected traits, such as zip code correlating with race, or employment gaps correlating with disability or caregiving responsibilities. Outcomes are shaped by both the selection of features and the weight assigned to them, meaning these design choices can materially influence results and potentially create disparities at scale.

## Common Types of AI Tools Used by Employers and Businesses

AI appears in many familiar forms, often without being labeled as such.

- **Predictive Tools:** Predictive analytics uses historical data to identify patterns and forecast future outcomes, such as sales performance or creditworthiness. Since these tools rely on past decisions, they may reinforce existing disparities if historical practices were biased.



**David J. Walton, AIGP,  
CIPP/US**

Partner

610.230.6105

---

## Service Focus

AI, Data, and Analytics

Employment Discrimination  
and Harassment

---

## Resource Hubs

AI Governance Hub

- **Example:** An insurer uses predictive analytics to estimate the likelihood that a policyholder will file a future claim and to inform premium pricing based on prior claims data.
- **Machine Learning Systems:** Machine learning models learn patterns from large datasets without fixed decision rules. During training, the model continuously adjusts the weights assigned to different features based on prior outcomes. While this adaptability can improve accuracy, it can also make bias harder to identify, particularly as models evolve over time.
  - **Example:** A bank uses a machine learning model to assess loan applications by learning from historical lending data and adjusting the weight assigned to factors such as credit history and repayment behavior.
- **Scoring, Ranking, and Recommendation Tools:** Many AI systems generate scores, rankings, or recommendations that inform or influence human decisions, such as applicant rankings or performance scores. Even when humans remain involved, there is a risk of over-reliance on automated outputs, which can reduce meaningful oversight.
  - **Example:** A resume-screening tool ranks applicants by learning which candidates advance in the hiring process and adjusting its evaluation criteria based on prior hiring outcomes.
- **Language-Based and Generative Tools:** Some AI systems analyze or generate language, including resume-screening tools, chatbots, and performance-summary tools. Since these systems are trained on large volumes of text, they can replicate patterns and assumptions present in the training data.
  - **Example:** An AI system generates automated email or chat responses to customer inquiries based on patterns learned from prior communications.

## **Where AI Bias Creates Legal Risk: Disparate Impact and Detection Challenges**

One of the most significant challenges with AI bias is that it can be difficult to detect. Many AI systems operate as “black boxes,” making it hard to understand how inputs translate

into outcomes. Without deliberate testing and documentation, biased results may go unnoticed until regulatory scrutiny or litigation arises.

## The 80/20 Rule

A potential measure for disparate impact when AI systems are involved is often evaluated using the 80/20 Rule. Under this framework, a selection rate for a protected group that is less than 80% of the rate for the most favored group may indicate potential adverse impact. The 80/20 Rule is not a definitive test of discrimination. Instead, it functions as a screening mechanism or warning indicator that may warrant closer review of a particular practice or decision-making process.

For example, if 60% of male applicants pass a screening assessment and only 45% of female applicants do, the resulting ratio of 75% falls below the 80% threshold. This outcome does not establish discrimination on its own, but it may signal potential bias and trigger further analysis.

## Why AI Bias Matters for Employers

AI bias matters because legal liability can arise even in the absence of discriminatory intent. Employers and businesses may be held responsible for practices that disproportionately impact protected groups when those outcomes are not job-related and consistent with business necessity.

As AI tools play an increasing role in employment decisions, courts are scrutinizing how these systems operate and whether they contribute to discriminatory results.

- One example is [\*Mobley v. Workday\*](#), a class action pending in California federal court, in which a job applicant alleges that Workday's AI-based screening tools systematically rejected him across more than 100 applications.
- Similarly, in [\*Harper v. Sirius XM\*](#), pending in Michigan federal court, an applicant alleges that the employer relied on an AI-powered applicant tracking system that embedded historical bias by using data points functioning as proxies for race, resulting in his candidacy being downgraded and eliminated before advancing in the hiring process.

## What Employers Can Do Now: Consider an AI Bias Audit

As AI-driven tools become more embedded in employment decision-making, employers should take proactive steps to assess and mitigate bias using a defensible, structured approach. Through its [AI Fairness and Bias Audit Solutions](#), Fisher Phillips helps employers evaluate risk and implement practical safeguards by:

- **Identifying where AI tools are used across the employment lifecycle**, including recruiting, hiring, onboarding, performance management, staffing and assignment decisions, employee relations, retention, and termination, to pinpoint where automated decision-making may create risk.
- **Conducting bias audits and compliance reviews of third-party AI vendors**, including reviewing vendor-provided bias audits or documentation, evaluating whether a tool qualifies as an automated decision tool under current and emerging laws, and advising on practical risk-mitigation strategies.
- **Auditing internally developed or custom-built AI tools**, including statistical testing for disparate impact across protected categories where data are available, explainable AI-based root-cause analysis, and recommendations to reduce or correct identified bias.
- **Providing privileged legal assessments and regulator-ready documentation**, including multi-jurisdictional compliance analyses, guidance on disclosure obligations, and summaries suitable for internal governance or external review.
- **Establishing AI monitoring and governance frameworks**, including regular bias audits, AI-assisted monitoring, regulatory updates, compliance workshops, and policy refresh guidance to address evolving legal requirements over time.

[Fisher Phillips delivers these services](#) in collaboration with AI employment analytics firm BLDS and AI fairness software provider SolasAI, offering employers an integrated and legally defensible approach to AI bias auditing, compliance, and governance.

## Conclusion

We will continue to monitor AI litigation and related developments and provide the most up-to-date information directly to your inbox, so make sure you are subscribed to [Fisher Phillips' Insight System](#). If you have questions, contact your Fisher Phillips attorney, the authors of this Insight, or any attorney in our [AI, Data, and Analytics Practice Group](#).