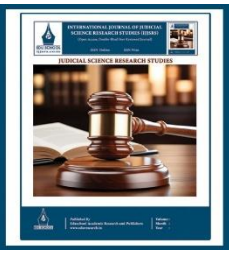
INTERNATIONAL JOURNAL OF JUDICIAL SCIENCE RESEARCH STUDIES (IJJSRS)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-0618 ISSN Print:



Algorithmic Bias and the Law: Ensuring Fairness in Automated Decision-Making

Binu Mole K

Assistant professor, School of Legal Studies, CUSAT, Kochi, Kerala, India.

Article information
Received: 13th May 2025
Received in revised form: 30th May 2025
Accepted: 24th July 2025
Available online: 21st August 2025

Volume: 2
Issue: 3
DOI: https://doi.org/10.5281/zenodo.16946253

Abstract

Algorithmic decision-making systems have become pervasive across critical domains including employment, housing, healthcare, and criminal justice. While these systems promise enhanced efficiency and objectivity, they increasingly demonstrate patterns of discrimination that perpetuate and amplify existing societal biases. This paper examines the evolving legal landscape governing algorithmic bias, analyzing recent regulatory developments, landmark litigation, and emerging compliance frameworks. Through comparative analysis of the fragmented U.S. approach and the European Union's comprehensive regulatory strategy, this study identifies persistent enforcement gaps and structural limitations in current legal frameworks. The research reveals that existing civil rights protections, while foundational, prove insufficient for addressing the novel challenges posed by automated decision-making systems. Key findings indicate that recent legal developments, including the Colorado AI Act and landmark cases such as *Mobley v. Workday*, represent significant progress toward establishing algorithmic accountability. However, substantial gaps remain in transparency requirements, technical standards for bias detection, and effective remediation mechanisms. This paper proposes an integrated legal framework combining rights-based protections, technical standards, and institutional oversight to ensure algorithmic fairness while fostering innovation.

Keywords:- Algorithmic bias, artificial intelligence law, automated decision-making, civil rights, discrimination, regulatory frameworks

I. INTRODUCTION

The proliferation of artificial intelligence and automated decision-making systems across critical sectors has fundamentally transformed how institutions allocate opportunities and resources. Algorithmic decision-making systems are widely used in various sectors, including criminal justice, employment, and education, promising enhanced efficiency, consistency, and objectivity in human decision-making processes. However, mounting evidence demonstrates that these systems frequently exhibit discriminatory patterns that systematically disadvantage protected groups, raising profound questions about fairness, accountability, and the protection of fundamental rights.

Algorithmic bias concerns primarily privacy and AI regulation in general, representing a complex intersection of technology, law, and social justice. The challenge extends beyond technical considerations to encompass broader questions of procedural fairness, democratic accountability, and the distribution of power in algorithmic societies. It is not an exaggeration to say that the regulatory framework for this phenomenon is characterized by a multiplicity of initiatives at national, regional, and global level, not to mention the rules of platforms and AI companies themselves.

The legal response to algorithmic bias has emerged as a critical frontier in civil rights enforcement, testing the adaptability of traditional anti-discrimination frameworks to address novel technological challenges. In the two years since generative AI was launched into popular use, legislators have attempted to regulate AI in all these ways and more, reflecting the urgency of developing effective legal responses to algorithmic discrimination.

This paper addresses three fundamental research questions: First, how effectively do existing legal frameworks address algorithmic bias and discrimination? Second, what regulatory approaches show the greatest promise for ensuring algorithmic

fairness while preserving innovation? Third, what legal and technical standards are necessary to create comprehensive algorithmic accountability systems?

The significance of this inquiry extends beyond academic discourse to affect millions of individuals subject to algorithmic decision-making daily. If predictive policing disproportionately targets certain communities based on flawed data, it risks undermining the very principles of fairness and justice it seeks to uphold. The stakes are particularly high in domains such as employment, housing, and criminal justice, where algorithmic decisions can determine life opportunities and fundamental liberties.

II.THEORETICAL FRAMEWORK

2.1 Conceptualizing Algorithmic Bias

Algorithmic bias manifests as systematic and unfair discrimination embedded in automated decision-making systems. Bias can emerge from many factors, including but not limited to the design of the algorithm or the unintended or unanticipated use or decisions relating to the way data is coded, collected, selected or used to train the algorithm. Understanding algorithmic bias requires distinguishing between several interconnected phenomena that contribute to discriminatory outcomes.

We identify five primary types of algorithmic bias: bias by algorithmic agents, discrimination based on feature selection, proxy discrimination, disparate impact, and targeted advertising. Each category represents distinct pathways through which bias enters algorithmic systems, requiring tailored legal and technical responses.

Computational bias emerges from technical design choices, including data selection, feature engineering, and model architecture decisions. However, systemic biases result from institutions operating in ways that disadvantage certain social groups, such as discriminating against individuals based on their race. Human biases further complicate the picture, relating to how people use data to fill in missing information, such as a person's neighborhood of residence influencing how likely authorities would consider the person to be a crime suspect.

When human, systemic and computational biases combine, they can form a pernicious mixture — especially when explicit guidance is lacking for addressing the risks associated with using AI systems. This recognition has prompted regulators to adopt socio-technical approaches that consider both technical and social dimensions of algorithmic fairness.

2.2 Legal Theory and Algorithmic Decision-Making

Traditional civil rights law operates through two primary theories of discrimination: disparate treatment (intentional discrimination) and disparate impact (policies that disproportionately affect protected groups). (Wang et al., 2024) argue that the US legal framework for algorithmic bias is rooted in fundamental civil rights protections and the Fourteenth Amendment, with a strong emphasis on three core principles: equality, non-discrimination and transparency.

The application of disparate impact theory to algorithmic systems presents unique challenges. Unlike traditional employment practices, algorithmic systems process vast datasets using complex mathematical operations that may obscure the causal mechanisms producing discriminatory outcomes. The EEOC guidance emphasizes that the legal consequences of disparate impact are not eliminated by offloading the decision-making to an AI system or to a third-party vendor.

Procedural fairness considerations add another layer of complexity. Huq notes that the EU General Data Protection Regulation (GDPR) affirmatively creates such a right by enshrining "the right not to be subject to a decision based solely on automated processing", establishing important precedents for algorithmic due process rights.

2.3 Regulatory Theory and Technological Governance

The governance of algorithmic systems reflects broader tensions between innovation and regulation in technological governance. Two main regulatory trends can be identified in the case of algorithmic bias: the European, highly restrictive but user-centric approach and the American, more liberal, "mixed" approach.

These divergent approaches reflect different conceptions of the state's role in technological governance and varying priorities regarding innovation, fundamental rights, and market efficiency. The European approach emphasizes comprehensive ex-ante regulation with strict compliance requirements, while the American model relies more heavily on ex-post enforcement through litigation and sectoral regulation.

III. LEGAL ANALYSIS

3.1 United States Federal Framework

The federal approach to algorithmic bias relies primarily on existing civil rights statutes, including Title VII of the Civil Rights Act of 1964, the Age Discrimination in Employment Act (ADEA), the Americans with Disabilities Act (ADA), and the Fair Housing Act. The "mixed" regulatory regime mentioned above stems from the fact that algorithmic bias is governed by both the relevant laws, primarily related to job search, employment and job placement, such as the Fair Credit Reporting Act (FCRA) and the Equal Employment Opportunity Commission (EEOC).

3.1.1 EEOC Guidance and Enforcement

The Equal Employment Opportunity Commission has emerged as a leading federal agency addressing algorithmic bias in employment. The EEOC has added technical guidance specifically to address the growing concerns raised by algorithmic bias and AI systems. This guidance clarifies that employers remain liable for discriminatory outcomes produced by algorithmic tools, regardless of whether these tools are developed internally or by third-party vendors (Equal Employment Opportunity Commission, 2022).

The EEOC's approach emphasizes that traditional disparate impact analysis applies to algorithmic decision-making systems. Employers must demonstrate that algorithmic tools are job-related and consistent with business necessity when they produce discriminatory outcomes. This standard requires employers to validate their algorithmic systems and consider less discriminatory alternatives.

3.1.2 Proposed Federal Legislation

Recent congressional activity reflects growing awareness of algorithmic bias issues. The Eliminating Bias in Algorithmic Systems Act of 2024 represents one of the most comprehensive federal proposals, requiring agencies that use, fund, or oversee algorithms to establish dedicated civil rights offices focused on bias and discrimination.

However, federal legislative progress remains limited, with most comprehensive proposals failing to advance through the legislative process. This has left enforcement primarily to existing civil rights statutes and agency guidance, creating uncertainty about legal standards and compliance requirements.

3.2 State-Level Innovations

3.2.1 The Colorado AI Act

On May 17, 2024, Colorado took the lead, becoming the first state to enact legislation to address algorithmic bias. Commonly known as the Colorado AI Act, the legislation is designed to provide "Consumer Protections in Interactions with Artificial Intelligence Systems". The Act establishes comprehensive obligations for both developers and deployers of high-risk AI systems. (Colorado AI Act, 2024)

The Colorado Act defines high-risk artificial intelligence systems as those making consequential decisions in domains including employment, healthcare, education, and financial services. The more prominent aspect of the Colorado law focuses on eliminating algorithmic bias in consequential decisions when automated AI systems are used to make selections involving employment, health care, legal services, or other high-risk categories.

Key provisions include requirements for impact assessments, bias testing, and ongoing monitoring. As of February 1, 2026, deployers must adopt the NIST's Artificial Intelligence Risk Management Framework or its equivalent, comply with any risk management framework designated by the Colorado Attorney General, complete an assessment or impact statement within the first ninety days of initial deployment, and update the impact statement annually.

3.2.2 Other State Initiatives

Utah's Artificial Intelligence Policy Act, S. 149, 2024 Gen. Sess. § 7 (Utah 2024) has a scope that is narrower than the Colorado AI Act, primarily focusing disclosure of the use of generative AI while clarifying that deployers of generative AI systems are responsible for the output of those systems when the output violates various consumer protection laws.

Illinois and California have also enacted targeted legislation addressing algorithmic bias in specific sectors. Illinois enacted legislation that would have also barred using zip code "as a proxy" for protected characteristics, though the version of the bill that ultimately passed replaced this input-based restriction with an output-based one.

3.3 European Union Regulatory Framework

3.3.1 The AI Act

In June 2024, the EU adopted the world's first rules on AI. The AI Act (Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence) is the first-ever comprehensive legal framework on AI worldwide. The Act establishes a risk-based approach to AI regulation, with specific requirements for high-risk AI systems that include algorithmic bias mitigation measures.

Data sets should contain accurate information, and potential bias must be identified and mitigated as well as possible. These requirements apply to training data, validation data, and testing data. The Act mandates comprehensive data governance requirements for high-risk AI systems, including bias detection and mitigation measures.

3.3.2 Bias Detection and Mitigation Requirements

Article 10 of the AI Act establishes detailed data governance obligations. High-risk AI systems which make use of techniques involving the training of AI models with data shall be developed on the basis of training, validation and test data sets that meet the quality criteria referred to in paragraphs 2 to 5.

The Act requires examination in view of possible biases that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations. This creates an affirmative obligation to identify and address potential sources of bias throughout the AI lifecycle.

Notably, Article 10(5) AI Act allows providers to collect sensitive data to fulfil the obligation. Article 10(5) AI Act aims to prevent discrimination, creating a limited exception to data protection rules for bias detection purposes.

3.4 Comparative Analysis

The regulatory approaches reflect fundamentally different philosophical orientations toward technological governance. The EU's comprehensive ex-ante regulatory approach contrasts sharply with the United States' reliance on existing civil rights frameworks and post-hoc enforcement.

The findings underscore persistent enforcement gaps, particularly concerning opaque black-box algorithmic design, which hampers bias detection and remediation. Both approaches face challenges in addressing the technical complexity of modern AI systems and ensuring effective enforcement across diverse deployment contexts.

The European approach offers greater legal certainty and comprehensive coverage but may impose significant compliance costs and potentially stifle innovation. The American approach preserves flexibility and encourages innovation but creates uncertainty about legal standards and may provide insufficient protection against discriminatory outcomes.

IV. CASE STUDIES AND LANDMARK LITIGATION

4.1 Mobley v. Workday: Defining Vendor Liability

The Mobley v. Workday litigation represents one of the most significant legal challenges to algorithmic bias in employment screening. The May 16 decision is a major development in Mobley v. Workday, one of the country's most closely watched legal challenges to the use of artificial intelligence in employment decisions.

4.1.1 Case Background and Legal Theory

Plaintiff Derek Mobley filed a putative class action lawsuit in February 2023, alleging that his applications for 80-100 jobs with employers who use Workday's screening tools were rejected because the tools allegedly allow the employers to make discriminatory judgments when evaluating applicants.

The case presents novel questions about vendor liability for discriminatory algorithmic systems. Mobley further alleged that Workday's administration and dissemination of the tools amounted to intentional and disparate impact discrimination in violation of Title VII of the Civil Rights Act, the Age Discrimination in Employment Act and the Americans with Disabilities Act.

4.1.2 Legal Developments and Significance

On July 12, 2024, the Court denied Workday's Motion to Dismiss with respect to an agency theory of liability, holding that Mobley plausibly alleged that Workday's employer-customers delegated to Workday and its AI screening tools their "traditional function of rejecting candidates or advancing them to the interview stage".

On May 16, 2025, Judge Rita Lin of the U.S. District Court for the Northern District of California granted preliminary certification under ADEA, allowing the lawsuit to move forward as a nationwide collective action. This development significantly expands the potential scope and impact of the litigation.

The court's analysis focused on the delegation of traditional employment functions to algorithmic systems. The Court viewed this delegation of duties traditionally exercised by the employer as the type of conduct Congress intended to prohibit by including the term "agent" in the definition of "employer" under the relevant anti-discrimination statutes.

4.1.3 Implications for Algorithmic Accountability

Mobley v. Workday Inc. is significant because it demonstrates that a vendor in the hiring process may be directly liable to candidates for discriminatory outcomes produced by its AI-enabled applicant screening tools. This establishes important precedent for holding AI system developers accountable for discriminatory outcomes.

The EEOC argued that Workday's software might enable discriminatory practices by allowing employers to exclude applicants from protected categories, violating Title VII of the Civil Rights Act of 1964. Federal agency support for the plaintiffs signals broader regulatory alignment on algorithmic accountability issues.

4.2 Meta Advertising Discrimination Settlement

On June 21, the United States Department of Justice announced that it had secured a "groundbreaking" settlement resolving claims brought against a large social media platform for allegedly engaging in discriminatory advertising in violation of the Fair Housing Act.

4.2.1 Algorithmic Targeting and Discrimination

The case involved three distinct algorithmic discrimination theories. First, it alleges that the company's advertising platform employed a machine-learning algorithm that allowed advertisers to target ads to consumers who "look like" a particular kind of "source audience," defined, in part, by reference to protected characteristics, including sex, or proxies for protected characteristics.

Second, it alleges that the company's "personalization algorithms" resulted in certain housing ads being targeted to potential customers — or not provided to potential customers — based on protected characteristics, or proxies for those characteristics.

4.2.2 Settlement Terms and Compliance Measures

The settlement required comprehensive algorithmic remediation measures. Develop a new system that addresses disparities between advertisers' target audiences and the consumers who actually receive ads based on the company's personalization algorithms. The new system will be reviewed by an independent third party to ensure that it meets agreed upon metrics for reducing the alleged disparities.

In reality, Meta quickly agreed to deploy a new "Variance Reduction System," or VRS — a second algorithmic overlay designed to reduce certain biases of the ad tool by rebalancing its results to render it less discriminatory. This approach demonstrates how technical solutions can be integrated with legal compliance requirements.

4.3 HireVue Accessibility Case

The ACLU's complaint also cites research showing that automated speech recognition systems and AI scoring tools often perform worse for individuals who speak English with non-white accents, including Black, Hispanic, and Indigenous applicants. This case highlights intersectional discrimination issues in algorithmic hiring systems.

D.K.—a Deaf Indigenous woman—applied for a promotion to Seasonal Manager at Intuit after working for the company in various seasonal roles since 2019. As part of the 2024 application process, she was required to complete an asynchronous video interview through HireVue, a platform that relies on automated speech recognition and AI-driven scoring.

The case demonstrates how algorithmic systems can create compounded barriers for individuals with multiple protected characteristics, requiring legal frameworks that address intersectional discrimination.

V. TECHNICAL STANDARDS AND COMPLIANCE FRAMEWORKS

5.1 NIST AI Risk Management Framework

In 2020, Congress directed the National Institute of Standards and Technology (NIST) to develop an AI Risk Management Framework with the public and private sectors. The framework provides voluntary guidelines for managing AI risks, including algorithmic bias.

5.1.1 Framework Structure and Core Functions

The NIST AI RMF introduces the four core functions and categories that the framework organizes its actionable guidelines under: Govern, Map, Measure, and Manage. Each function addresses different aspects of AI risk management throughout the system lifecycle.

Organizations assess potential risks to stakeholders, including AI actors and end-users. This function helps in understanding: Data Quality Issues: Biases or inaccuracies in training data. Algorithmic Risks: Potential for unintended behavior or outcomes. Operational Risks: Failures in deployment environments.

5.1.2 Bias-Specific Guidance

The Framework considers approaches to develop characteristics of trustworthiness, including accuracy, explainability and interpretability, reliability, privacy, robustness, safety, security, and mitigation of unintended and/or harmful uses.

NIST's Reva Schwartz, the main distinction between the draft and final versions of the publication is the new emphasis on how bias manifests itself not only in AI algorithms and the data used to train them, but also in the societal context in which AI systems are used.

5.2 IEEE Standards Development

Currently, a new IEEE standard is being drafted that aims to specify methodologies which help creators of algorithms eliminate issues of bias and articulate transparency. The IEEE 2857-2024 standard provides practical guidance for bias management in algorithmic systems.

In 2022, the IEEE released a standard aimed at specifying methodologies to help creators of algorithms address issues of bias and promote transparency regarding the function and potential effects of their algorithms. This standard complements regulatory requirements with technical implementation guidance.

5.3 Organizational Compliance Strategies

5.3.1 Bias Auditing and Testing

Effective compliance requires systematic bias auditing throughout the AI lifecycle. The legal team should work closely with HR and the IT team to conduct bias audits on a regular basis. If an audit reveals the tool has disparate impacts at any point, the company should consider working with the vendor to implement bias-mitigating techniques.

Conduct regular internal AI audits to detect compliance violations or biases. Organizations must establish ongoing monitoring systems to detect emergent bias and ensure continued compliance with legal requirements.

5.3.2 Governance and Documentation

Clear documentation helps to systematically implement policies and procedures, and standardizes how an organization's bias management is implemented. Comprehensive documentation serves both compliance and accountability purposes.

Prepare detailed internal documents clearly explaining the AI tool's operation and selection criteria based on the review mentioned in item a to protect the company in case of government investigations or lawsuits.

VI. CRITICAL EVALUATION AND LIMITATIONS

6.1 Enforcement Gaps and Implementation Challenges

The paper highlights how current regulatory efforts disproportionately affect marginalized communities and fail to provide effective protection across jurisdictions. It also identifies structural imbalances in legal instruments, particularly in relation to risk classification, transparency, and fairness standards.

Current legal frameworks face several critical limitations. First, the complexity of modern AI systems creates significant challenges for bias detection and attribution. Persistent enforcement gaps, particularly concerning opaque black-box algorithmic design, which hampers bias detection and remediation remain a fundamental challenge.

Second, the technical-legal gap creates difficulties in translating legal standards into actionable technical requirements. Policy formulations on this topic are rather vague, favouring flexibility on one hand, but leaving the contours of law-abiding bias management undefined for practitioners, contributing to legal risk and uncertainty.

6.2 Transparency and Explainability Limitations

To prevent bias and discrimination, transparency is essential. Transparency is necessary in respect of both the technical accuracy of the algorithm and its deployment, particularly towards the people affected. The Act does not require sufficient transparency.

Even comprehensive regulatory frameworks like the EU AI Act face limitations in balancing transparency requirements with trade secret protection and system security. The AI Act exempts law enforcement from many of the relevant general requirements, despite the fact that law enforcement agencies are most likely to interact intrusively with the fundamental rights of persons.

6.3 Intersectional and Systemic Bias Challenges

Current legal frameworks struggle to address intersectional discrimination and systemic bias effectively. Different fields and society at certain times and places define fairness and discrimination in different ways, creating challenges for developing universal standards.

Even when data are collected representatively, biases can occur. Social bias can occur regardless of sample representativeness. For instance, in loan application settings, if certain groups (such as women) systematically receive lower wages, the data will contain inherent bias even with perfect representation.

VII. RECOMMENDATIONS AND FUTURE DIRECTIONS

7.1 Integrated Legal Framework

The analysis reveals the need for a comprehensive legal framework that combines rights-based protections, technical standards, and institutional oversight. This framework should include:

- Substantive Rights: Clear legal rights to algorithmic fairness, transparency, and contestability, building on existing civil rights protections while addressing the unique challenges of automated decision-making.
- Procedural Protections: Due process requirements for algorithmic decision-making, including rights to explanation, human review, and effective remedy mechanisms.
- Technical Standards: Mandatory compliance with technical standards for bias detection, testing, and mitigation, integrated with legal liability frameworks.

7.2 Institutional Reform and Capacity Building

Eliminating Bias in Algorithmic Systems Act of 2024 requires agencies that use, fund, or oversee algorithms to have an office of civil rights focused on bias, discrimination, and other harms of algorithms. This approach should be expanded to create specialized algorithmic oversight institutions with technical expertise and enforcement authority.

Educate developers, data scientists, and executives on responsible AI use. Implement mandatory AI governance training to ensure awareness at all levels. Capacity building must extend beyond technical training to include legal compliance and ethical considerations.

7.3 International Coordination and Harmonization

The global nature of AI development and deployment requires international coordination to ensure effective governance. For multi-jurisdictional enterprises, it is imperative to prioritize cross-border compliance strategies by aligning AI systems with the most EU standards, ensuring operational and legal consistency across regions.

International standards organizations should develop harmonized approaches to algorithmic bias assessment and mitigation, building on existing frameworks like the OECD AI Principles and ISO/IEC standards.

7.4 Research and Development Priorities

Future research should prioritize several critical areas:

- Technical Solutions: Development of more effective bias detection and mitigation techniques that can operate across diverse AI architectures and deployment contexts.
- Legal Innovation: Exploration of novel legal mechanisms for algorithmic accountability, including liability insurance, certification schemes, and alternative dispute resolution mechanisms.
- Empirical Research: Comprehensive studies of algorithmic bias impacts across different domains and populations to inform evidence-based policy development.

VIII. CONCLUSION

The challenge of algorithmic bias represents one of the most significant civil rights issues of the digital age. As automated decision-making systems become increasingly pervasive across critical domains, the need for effective legal frameworks to ensure algorithmic fairness has become urgent and undeniable.

This analysis reveals a legal landscape in transition, characterized by significant innovation alongside persistent gaps and limitations. Recent developments, including landmark litigation such as *Mobley v. Workday* and comprehensive regulatory

initiatives like the EU AI Act and Colorado AI Act, represent substantial progress toward establishing algorithmic accountability. However, the complexity of modern AI systems and the global nature of their deployment present ongoing challenges for effective governance.

The comparative analysis of regulatory approaches demonstrates that neither purely market-based nor comprehensive regulatory strategies alone can address the multifaceted challenges of algorithmic bias. Instead, the evidence points toward the need for integrated frameworks that combine rights-based protections, technical standards, institutional oversight, and international coordination.

The stakes of this endeavor extend far beyond regulatory compliance to encompass fundamental questions of justice, equality, and democratic governance in algorithmic societies. "If we are to develop trustworthy AI systems, we need to consider all the factors that can chip away at the public's trust in AI. Many of these factors go beyond the technology itself to the impacts of the technology".

Moving forward, the legal community must work collaboratively with technologists, policymakers, and civil society organizations to develop comprehensive solutions that protect fundamental rights while fostering beneficial innovation. The window for proactive intervention remains open, but effective action requires immediate and sustained commitment to developing robust legal frameworks for algorithmic accountability.

The future of algorithmic governance will be determined by choices made today about the values we embed in our technological systems and the legal frameworks we create to govern them. Ensuring algorithmic fairness is not merely a technical challenge but a fundamental requirement for maintaining democratic values and social justice in an increasingly automated world.

REFERENCES

Anti-Defamation League. (2025). *Antisemitic bias in large language models*. ADL Research Report.

Artificial Intelligence Policy Act, S. 149, 2024 Gen. Sess. § 7 (Utah 2024).

, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(3), 671–732.

Ceccon, M., Cornacchia, G., Dalle Pezze, D., Fabris, A., & Susto, G. A. (2025). Underrepresentation, label bias, and proxies: Towards data bias profiles for the EU AI Act and beyond. *arXiv preprint arXiv:2507.08866*.

Colorado AI Act, S. 24-205, 74th Gen. Assemb. (Colo. 2024).

Cornacchia, G., et al. (2023). Algorithmic fairness datasets: The story so far. *Data Mining and Knowledge Discovery, 37*(6), 2074–2152.

Deck, C., et al. (2024a). Compatibility of fairness metrics with legal anti-discrimination frameworks. *AI & Society, 39*(2), 487–503.

Deck, C., et al. (2024b). From algorithmic fairness to legal compliance: Bridging the gap. *Computer Law & Security Review, 52*, 105892.

Eliminating Bias in Algorithmic Systems Act of 2024, H.R. 10092, 118th Cong. (2024).

Equal Employment Opportunity Commission. (2022). *Technical assistance document on the consideration of algorithms in employment decisions under Title VII*. EEOC.

European Parliament. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence*. *Official Journal of the European Union*.

Fabris, A., et al. (2022). Algorithmic audit of Italian job portals. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 234–244).

Gebru, T. (2019). Race and gender in machine learning research. *Communications of the ACM, 62*(5), 42–44.

Gebru, T., et al. (2021). Datasheets for datasets. *Communications of the ACM, 64*(12), 86–92.

Hacker, P. (2018). Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Market Law Review, 55*(4), 1143–1185.

IEEE. (2024). *IEEE Standard 2857-2024: Privacy engineering for system life cycle processes*. IEEE Standards Association.

Lin, L. (2025, January 1). Algorithmic justice or bias: Legal implications of predictive policing algorithms in criminal justice. *Johns Hopkins Undergraduate Law Review*.

Lünich, M., & Kieslich, K. (2024). Exploring the roles of trust and social group preference on the legitimacy of algorithmic decision-making vs. human decision-making for allocating COVID-19 vaccinations. *AI & Society, 39*(1), 309–327.

MacCarthy, M. (2018). Standards of fairness for disparate impact assessment of big data algorithms. *Cumberland Law Review, 48*, 67–98.

Mehrabi, N., et al. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 1–35.

Mobley v. Workday, Inc., No. 3:23-cv-00770-RFL (N.D. Cal. filed Feb. 1, 2023).

NAACP. (2024, February 15). *Artificial intelligence in predictive policing issue brief*. NAACP.

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1).

Obermeyer, Z., et al. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453.

Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence. (2024). *Official Journal of the European Union, L 1689*.

Schwartz, R., et al. (2022). Towards a standard for identifying and managing bias in artificial intelligence (NIST Special Publication 1270). *National Institute of Standards and Technology*.

Sousa e Silva, F. (2024). Algorithmic discrimination and the rule of law. *Computer Law & Security Review, 52*, 105863.

United States v. Meta Platforms, Inc. (2022). Settlement Agreement, U.S. District Court for the Southern District of New York.

Wachter, S. (2024). The AI Act and algorithmic discrimination: Challenges and opportunities. *European Law Review, 49*(3), 287–312.

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review, 41*, 105567.

Wang, X., Wu, Y. C., Ji, X., & Fu, H. (2024). Algorithmic discrimination: Examining its types and regulatory measures with emphasis on US legal practices. *Frontiers in Artificial Intelligence, 7*, 1320277.