INTERNATIONAL JOURNAL OF ADVANCED LEGAL RESEARCH

ALGORITHMIC BIAS AND HUMAN RIGHTS: SAFEGUARDING EQUALITY IN AGE OF AI

- Janaan Naseeb Mulamparambil[1] & Adiba Khan[2]

ABSTRACT

The digital landscape is at the juncture of intersection with human rights violation. The world has come to the point where humans have entered to the arena of algorithmic surveillance. Slowly and gradually the algorithmic bias has seeped into fourth generational human right concern. The ideologies and perspectives a person possess is scrutinized under political will and undermined by shadow banning. The right to expression and right to privacy are curtailed by algorithmic scrutinization. From employment sectors to criminal justice enforcement mechanisms, the rampant use of biased data could be traced. This has reached to the extent that Kerala High Court in India had to emphasis the judicial officers not to use digitalized mechanisms for judicial purposes. Question of equality is to be raised when traditionally biased data and information are used as sole basis for inspection of criminal behaviors among the marginalized sections of society. The gender-based discrimination is also a gross violation of human rights under digital surveillance model algorithm. The values and rights protected under various conventions are at stake when algorithm is politicized and biased. The paper aims to highlight the major violation of human rights and equality concerns that have to be addressed in light of algorithmic bias. Need of the hour is to attain the balanced approach that adopt algorithm-based technology and human centric perspective ensuring human rights and equality.

INTRODUCTION

---

[1] Student at Aligarh Muslim University, Faculty of Law, Aligarh
[2] Student at Aligarh Muslim University, Faculty of Law, Aligarh

In the world of technology, one may face infringement of rights at various levels. The arbitrariness and inequality are rampant. While AI promises efficiency and innovation it also raises significant concerns about fairness particularly in the form of algorithmic bias. Algorithmic bias occurs when automated systems produce prejudiced outcomes that disproportionately disadvantage certain groups based on race, gender, ethnicity, or other protected characteristics. There are instances where normal citizen handled Instagram accounts are getting shadow banned for posting about certain topics. When the war happens and the world countries wants the people to believe in a political propaganda then the main tool that they use is algorithm. Protecting human rights in the age of AI demands a critical examination of how algorithms are designed, deployed, and regulated. Without adequate safeguards algorithmic bias risks increase and amplifying existing social inequalities. This undermines the principles of justice and equal treatment enshrined in international human rights law and domestic constitutions. These problems matter because they take away people's dignity and equality. They can make discrimination worse based on things like race, gender, caste, or how much money someone has often without anyone noticing. When important decisions about housing, healthcare, jobs, or freedom are made by machines it becomes harder to spot or fight unfairness.

Today AI affects many parts of our lives so dealing with algorithmic bias isn't just a tech problem it is about protecting human rights. We need to ask how can we make sure these systems are fair to everyone? How do we protect those who might be hurt by automatic discrimination?

ALGORITHMIC BIAS –

When an AI system makes consistent, repeating mistakes that result in unfair outcomes like favoring one group over another based on factors like ethnicity, gender, age, socioeconomic status, or other characteristics this is known as algorithmic bias.[3]

Biases frequently result from the datasets used to train AI when they are incomplete, unrepresentative, or reflect historical inequalities. Design bias when developer's assumptions and

---

[3] Black's Law Dictionary, 11th edn. (Thomson Reuters, 2019), p. 88

choices embed subjective or cultural preferences. Feedback bias when the AI system learns from outcomes that reinforce existing stereotypes or discriminatory practices.

"Algorithmic bias refers to the tendency of AI systems to make decisions that reflect and perpetuate human biases, leading to discrimination and unequal treatment that may violate international human rights standards"[4] Algorithm bias typically results from improper design, biassed data, or social injustices represented in training datasets. Facial recognition software, for instance, has been shown to more often misidentify women and people with darker skin tones, illustrating how technology can reinforce historical inequality. In addition to biassing judgement, these prejudices undermine fundamental human rights like equality, privacy, and due process. In the 2024 case of Mobley v. Workday, Inc.[5] a U.S. federal court granted a lawsuit claiming that an AI-based hiring tool discriminated against candidates on the basis of age, race, and handicap. The case illustrates how employers' algorithmic systems may inadvertently replicate bias found in training data, in violation of anti-discrimination legislation. Therefore, human rights particularly the rights to equality before the law and nondiscrimination guaranteed by Article 7 of the Universal Declaration of Human Rights are directly threatened by algorithmic bias. Governments and businesses must respect the idea of human rights by design and guarantee openness, responsibility, and equity in the development and application of AI in order to protect equality in the era of AI.

EFFECT ON NATIONAL Legal and Constitutional Frameworks -

Article 14 – "The State shall not deny to any person equality before the law or the equal protection of the laws within the territory of India"

The principle of equality before the law and equal protection of laws serves as a fundamental safeguard against discrimination. However, this guarantee can be undermined when state authorities employ AI-driven tools that produce biased outcomes. With the rapid integration of

---

[4]United Nations Human Rights Office of the High Commissioner, The Right to Privacy in the Digital Age (United Nations Human Rights Council, 2021) https://www.ohchr.org/en/documents/thematic-reports/right-privacy-digital-age
(Last visited on 20 October 2025.)
[5]. N.D. Cal. July 12, 2024

Artificial Intelligence in diverse domains such as hiring, financial services, policing, and judicial decision making, concerns about fairness and accountability have become more pressing. As AI systems increasingly effecting and influencing key aspects of governance and human interaction algorithmic bias poses serious challenges to the principle of equality embodied in Article 14 of the Indian Constitution and reinforced by numerous international human rights conventions. The risk lies in automated systems reinforcing existing social inequalities thereby threatening the constitutional and moral commitment to impartiality and justice.[6]

ARTICLE 15

Prohibition of discrimination on grounds of religion, race, caste, sex or place of birth- The State shall not discriminate against any citizen on grounds only of religion, race, caste, sex, place of birth or any of them.

Article 16 Equality of opportunity in matters of public employment

(1) There shall be equality of opportunity for all citizens in matters relating to employment or appointment to any office under the State

(2) No citizen shall, on grounds only of religion, race, caste, sex, descent, place of birth, residence or any of them, be ineligible for, or discriminated against in respect of, any employment or office under the State.

The Constitution explicitly forbids discrimination based on religion, race, caste, sex, or place of birth. When artificial intelligence systems are employed in areas such as public recruitment or distribution of state resources and their outcomes disproportionately disadvantage Scheduled Castes, Scheduled Tribes, Other Backward Classes, or women, such practices may warrant judicial scrutiny. Algorithms that help in welfare distribution, credit evaluation, or employment selection must therefore be designed with awareness of social and economic hierarchies. Articles 15 and 16 of the Indian Constitution provide the bedrock for ensuring non-discrimination and equal opportunity particularly within the sphere of public employment and governance. As India

---

[6] Prashant Reddy, "Algorithms and Accountability: Understanding the Legal Framework for Artificial Intelligence in India" (2020) 62(2) JILI 201, 205

progressively includes algorithmic technologies into administrative and employment processes, it becomes essential to reinterpret and expand these constitutional guarantees.[7] This evolution is necessary to confront algorithmic bias a contemporary manifestation of structural discrimination embedded within automated decision-making systems.

21. Protection of life and personal liberty- No person shall be deprived of his life or personal liberty except according to procedure established by law.

Over the years judicial interpretation has expanded this provision beyond mere physical existence to include the right to live with dignity, the right to privacy, and the right to fair procedure.

EFFECT ON INTERNATIONAL LEGAL FRAMEWORK

ARTICLE 1 UNDHR - "All human beings are born free and equal in dignity and rights"

Article 7- "All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination."[8]

ICESCR Articles 6 and 9 protect the right to work and social security both threatened by biased recruitment or welfare algorithms.[9]

International Covenant on Civil and Political Rights (ICCPR, 1966) - Article 17 guarantees the right to privacy.[10]

UNESCO Recommendation on the Ethics of Artificial Intelligence (2021) - UNESCO's Recommendation is the first global normative instrument on AI ethics adopted by 193 member states. It emphasizes transparency, accountability, fairness, and non-discrimination as ethical imperatives. It explicitly warns against algorithmic systems that reproduce stereotypes,

---

[7] P. D. Sujith & V. R. Dinkar, Algorithmic Bias and the Right to Equality: Re-imagining Constitutional Guarantees in the Age of Artificial Intelligence (2024) 11(2) IJRDO 872

[8] Universal Declaration of Human Rights, 1948, arts. 1 and 7.

[9] International Covenant on Economic, Social and Cultural Rights, 1966, arts. 6 and 7.

[10] International Covenant on Civil and Political Rights, 1966, art.17

marginalize vulnerable communities, or distort public discourse. This framework recognizes that biased algorithms can erode democratic participation and human dignity, urging states to ensure human oversight, algorithmic audits, and impact assessments before deploying AI in governance.

NATIONAL AND INTERNATIONAL INSTANCES OF ALGORTHMIC BIASNESS

The Delhi Police's use of facial recognition systems for crowd monitoring, protest surveillance, and crime identification has drawn criticism for algorithmic inaccuracies and biases against darker skin tones and women. Studies by the Internet Freedom Foundation (IFF) revealed that such systems had inaccurate results exceeding 80% in identifying suspects. Although not yet adjudicated by the courts, the use of such AI tools without transparency or judicial oversight raises serious concerns under Articles 14, 19, and 21, particularly regarding privacy, equality, and freedom of expression. It also contradicts the Supreme Court's privacy doctrine established in Puttaswamy which mandates proportionality and necessity in surveillance activities. In recent years digital lending platforms such as Paytm, CASHe, LazyPay, and KreditBee have revolutionized access to credit in India. However these platforms often rely on AI algorithms to evaluate creditworthiness, leading to growing concerns about algorithmic bias. A significant issue is the exclusion or penalization of applicants from Tier-2 and Tier-3 cities, women, gig workers, and individuals without formal credit histories or CIBIL scores. The Reserve Bank of India's recent circular states that credit decision algorithms must be designed to flag potential discrimination factors and must be auditable.[11]

At international level there are various instances wherein the algorithmic bias can be traced and in shutting down the prominent legal and political issues for various nations. On platforms operated by Meta Platforms Inc. such as Instagram and Facebook investigators found that pro-Palestinian posts, stories and accounts were subjected to reduced visibility by shadow-banning removal of engagement features such as likes and comments, account suspensions or complete de-lists even when the content did not appear to breach policy. Research shows over 1,000

---

[11] Shravya S. Shekar, "Algorithmic Bias in Employment: A Comparative Legal Analysis of Anti-Discrimination Frameworks in the EU and India" (2025) 5(3) IJALR 1-17

documented cases from more than 60 countries of content takedowns or suppression of Palestinian related posts on Facebook and Instagram between October–November 2023.

Vulnerable groups at greater risk

Algorithmic bias disproportionately affects vulnerable and marginalized groups because AI systems are typically trained on historical data that reflects existing social inequalities. When these biased datasets feed machine learning models. The resulting algorithms tend to replicate and sometimes amplify discrimination instead of eliminating it. [12]

Fair Housing Council of San Fernando Valley v. Roommates.com, LLC (2008)[13]

Roommates.com operated a website where users sought roommates by filling out questionnaires about gender, sexual orientation, familial status, and other protected characteristics. The site's matching algorithm allowed and even encouraged discrimination by allowing users to exclude potential roommates based on these protected categories.The Fair Housing Act prohibits discrimination in housing based on race, sex, religion, familial status, national origin, or disability. The question was whether the website's algorithm which facilitated discrimination by filtering users based on these characteristics, violated the FHA. The court held that Roommates.com was liable because its algorithm actively facilitated discriminatory housing choices violating the Fair Housing Act. The site's role in enabling bias through automated processes made it responsible under the law.

Is There an Invisible Censorship - Algorithms and Free Speech

In the digital age censorship no longer always appears in the form of overt content takedowns or government bans. Increasingly it operates through algorithmic systems subtle, opaque, and often untraceable. This form of control often termed **i**nvisible censorship is driven by algorithms on social media platforms that filter, suppress, or demote content without users ever being notified. While these algorithms are designed to manage vast amounts of data and ensure community standards they can unintentionally or sometimes systematically suppress marginalized voices,

---

[12]Pooja Agrawal, "Algorithmic Bias in Artificial Intelligence: A Socio-Legal Perspective" (2024) 4(1) GlIJ 2.

[13] Fair Housing Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir 2008).

dissenting opinions, and politically sensitive content. [14]This form of censorship undermines the fundamental human right to freedom of expression as protected under Article 19 of the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights.

Are Human Rights Laws Keeping Up with Algorithmic Advancements?

As artificial intelligence and algorithmic systems increasingly shape decisions about finance, employment, healthcare, policing, and online speech human rights laws are struggling to keep pace. While core human rights principles like equality, non-discrimination, privacy, freedom of expression, and the right to remedy are well established under international instruments such as the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights their application to automated decision making remains unclear and underdeveloped in most legal systems. Traditional human rights frameworks assume the presence of evidence, intention, and traceable process of human rights violation. In contrast algorithmic decisions are often made through black-box systems trained on massive datasets, with no obvious accountability when things go wrong. [15]

Moreover existing anti discrimination laws such as those found in U.S. civil rights legislation or European human rights law focus on human intent not systemic bias embedded in code or data. Courts are only beginning to interpret whether algorithmic bias constitutes indirect discrimination. .

Global Responses and Regulatory Efforts to Address Algorithmic Bias

One of the most comprehensive efforts is the European Union's AI Act which introduces a risk based regulatory framework. Under this model AI systems are categorized as posing minimal, limited, high, or unacceptable risk. High risk systems such as those used in credit scoring, predictive policing, and hiring are subject to strict requirements regarding transparency, data quality, human oversight, and accountability. The Act prohibits certain harmful uses outright

---

[14]Algorithmic Censorship by Social Platforms: Power and Accountability" by J. Cobbe (2021) accessible at https://www.springernature.com/gp/open-science/about/the-fundamentals-of-open-access-and-open-research (last visited on 18 Oct 2025)

[15] Jennifer Cobbe, "Algorithmic Censorship by Social Platforms: Power and Resistance" (2020) 33(2) Philosophy & Technology 315–335.

such as AI based social scoring and real time biometric surveillance. Its extraterritorial scope means even non-EU companies must comply if their AI systems impact EU residents. International law instruments have also gained attraction. The UNESCO Recommendation on the Ethics of AI2021 encourages the development of AI that is transparent, accountable, and inclusive. The OECD Principles on AI adopted by over 40 countries similarly emphasize respect for human rights and democratic values in the design and deployment of AI. In 2023 the Hiroshima AI Process launched by Japan during its G7 presidency brought together over 49 nations to establish shared principles for regulating generative AI technologies like ChatGPT. Countries like China have also enacted national laws such as the 2023 Interim Measures for the Management of Generative AI Services requiring content moderation, user consent, and ideological alignment with national values. Meanwhile regulatory bodies like the European Centre for Algorithmic Transparency are now operational auditing algorithmic systems for bias and transparency. Despite these advances significant challenges remain. Many countries especially in the Global South lack regulatory frameworks, technical capacity, or political will. Even in developed regions enforcement of AI laws is still evolving and the technical complexity of auditing AI systems for bias remains a barrier. Moreover global regulatory fragmentation poses challenges for multinational companies and developers trying to comply with differing standards. [16]

CASE LAWS

State v. Loomis

State v. Loomis is an important case about fairness in using algorithms in court. Eric L. Loomis was given a harsher sentence based partly on a risk assessment tool called COMPAS which predicts how likely someone is to reoffend. Loomis argued that this tool was unfair because it was a private black box algorithm with the meaning no one could see how it worked and it might use biased data like race or gender.[17]

---

[16] International Legal Frameworks and Initiatives Governing Algorithmic Bias (2021–2025)
https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (,last visited on 19 Oct 2025)
[17] State v. Loomis, 881 N.W.2d 749 (Wis. 2016)

The court decided that using COMPAS was not illegal but gave important warnings. Judges must know the tool's limits and must not rely on it alone and defendants must be allowed to question the data it uses. This case matters because it was one of the first to deal with algorithmic bias in the legal system. It showed the risks of using computer programs in sentencing without transparency. The court didn't ban the tool but made it clear that fairness, transparency, and the ability to challenge such tools are essential in any system using algorithms to make decisions.

R (Bridges) v Chief Constable of South Wales Police[18]

UK based case about the use of facial recognition technology by police and the risks of bias and lack of oversight in such systems. Edward Bridges a civil rights campaigner challenged the South Wales Police's use of a live Automated Facial Recognition system called AFR Locate. This system scanned public spaces and compared people's faces to a watchlist of suspects. Bridges argued that this was a serious invasion of privacy and possibly discriminatory. The Court of Appeal agreed and ruled that the police use of AFR was unlawful for several reasons. One of this was that the police failed to check whether the facial recognition software was biased especially in terms of race and gender. This breached their Public Sector Equality Duty under the Equality Act 2010 which requires public bodies to consider how their actions affect different groups.

In the case of **State of West Bengal v. Anwar Ali Sarkar[19]** the Court held that equality demands not only equal laws but equal application. This can be extended to our understanding of algorithms reproducing structural inequalities would offend this settled principle.

In the case of **Arijit Singh v. Codible Ventures LLP**[20] the Bombay High Court ruled against AI voice cloning, affirming that personality attributes like voice and likeness were protectable rights. This precedent could be expanded to protect individuals from other algorithmic harms.

OBSERVATION & CONCLUSION

---

[18] [2020] EWCA Civ 1058

[19] AIR 1952 SC 75.

[20] 2024 SCC OnLine Bom 2445

It is important to think about how we can protect people from unfair treatment caused by biased algorithms. One way to do this is by making AI systems more open and easier to understand. People should be able to see how decisions are made and ask questions when something feels wrong. This will help build trust and give people a chance to challenge unfair outcomes. Another important step is to regularly check and test AI tools for bias before they cause harm. These checks should look at how AI affects different groups especially those who have been treated unfairly in the past. To make sure AI treats everyone fairly the data used to teach these systems must include a wide range of people. This means including voices and experiences from all backgrounds not just the majority. We also need clear rules that hold companies responsible when their AI causes harm. People should have ways to get help and be compensated if they are treated unfairly by a machine. Even when AI is involved humans must always have the final say especially in important areas like health jobs or legal matters. This human oversight can prevent mistakes and protect people's rights. It is also very important to involve the public especially marginalized communities in deciding how AI should be used. When people feel heard and included they are more likely to trust technology and feel safe. Using education to help everyone understand how AI works will empower people to speak up against unfairness. Lastly since AI is used worldwide countries should work together to create shared rules and standards. This cooperation can help prevent harmful uses of AI and protect human rights everywhere.

As artificial intelligence becomes deeply connected to our daily lives it is crucial to recognize that these technologies are not neutral. Algorithms reflect the values and data they are built on and when unchecked they can replicate and even worsen existing social biases. This reality raises serious concerns about the protection of fundamental human rights especially the right to equality and freedom from discrimination. Without proper safeguards AI risks turning into a tool that entrenches injustice rather than promotes fairness.

The challenge of algorithmic bias is complex because it often operates invisibly making unfair outcomes difficult to detect and challenge. People affected by biased algorithms may never know why they were denied a loan excluded from a job opportunity or subjected to harsher treatment by the criminal justice system. This opacity not only undermines trust in technology but also

threatens the dignity and equal standing of individuals and communities especially those already vulnerable due to race, gender, caste, or socioeconomic status.

What I have observed from this research is that the algorithmic bias is real and discrimination based on this technical error can affect the livelihood of people. It shows that unfair algorithms can hurt people's chances in jobs, healthcare, and justice without them even knowing. The paper points out that protecting human rights and fairness must be a priority when using AI. It also suggests ways to make AI fairer and safer. Overall the paper reminds us that technology should help everyone equally and respect human dignity.

Bibliography

1. Black's Law Dictionary, 11th edn. (Thomson Reuters, 2019), p. 88

2. United Nations Human Rights Office of the High Commissioner, The Right to Privacy in the Digital Age (United Nations Human Rights Council, 2021) https://www.ohchr.org/en/documents/thematic-reports/right-privacy-digital-age (Last visited on 20 October 2025

3. Prashant Reddy, "Algorithms and Accountability: Understanding the Legal Framework for Artificial Intelligence in India" (2020) 62(2) JILI 201, 205

4. P. D. Sujith & V. R. Dinkar, Algorithmic Bias and the Right to Equality: Re-imagining Constitutional Guarantees in the Age of Artificial Intelligence (2024) 11(2) IJRDO 872

5. Shravya S. Shekar, "Algorithmic Bias in Employment: A Comparative Legal Analysis of Anti-Discrimination Frameworks in the EU and India" (2025) 5(3) IJALR 1-17

6. Pooja Agrawal, "Algorithmic Bias in Artificial Intelligence: A Socio-Legal Perspective" (2024) 4(1) GIIJ 2.

7. International Legal Frameworks and Initiatives Governing Algorithmic Bias (2021–2025) https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (,last visited on 19 Oct 2025)

8. Jennifer Cobbe, "Algorithmic Censorship by Social Platforms: Power and Resistance" (2020) 33(2) Philosophy & Technology 315–335.

9. International Legal Frameworks and Initiatives Governing Algorithmic Bias (2021–2025)
    https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (,last visited on
    19 Oct 2025