Invisible Injustice: AI, Human Rights, and the Accountability Crisis

Abhishek Tiwari[1]

ABSTRACT

We are increasingly relying on artificial intelligence (AI) systems to do important things in our lives, whether it be in the field of law, healthcare, the job market, or education. The environment of creativity has been changed by AI, which has become a revolutionary tool in many fields. AI works by basic guidelines that computers are given and then the systems can generate results from data thanks to machine learning. AI now analyzes data, automates processes, and resolves challenging issues. We are beginning to encounter unanticipated issues resulting from these technologies. This paper examines how AI systems can violate new human rights. Such violations range from biased algorithms, and decision-making processes that align with society's stereotypes, to amorphous accountability when things take a turn for the worse.

AI systems are likely to behave erratically, with decisions that heavily shape people's lives made with no transparency or accountability. When they are trained on biased or unrepresentative data, such systems can learn structural discrimination from that data, leading to discriminatory hiring, differential access to health care, and more punitive sentences in criminal courts. Algorithmic bias, therefore, flouts the right to equality and non-discrimination recognized under international human rights law.

Moreover, the paper details the effects of AI-powered surveillance and data profiling, arguing that uncontrolled collection of personal information and processing thereof compromise people's right to privacy and threaten civil liberties. Such practices disproportionately impact vulnerable and marginalized communities and serve to reproduce social inequalities rather than combat them. One of the major problems this research is tackling is liability: who is responsible when AI harms? The uncertainty of legal liability again, whether for deployers or developers of AI or AI itself, is an all-encompassing justice gap for victims of AI-caused harm.

[1] 3rd Year Law Student, School of Law, CHRIST (Deemed to be University), Bangalore.

Existing legal structures lack the tools to deal with the decentralised, autonomous nature of these new AI systems.

Keywords: Artificial Intelligence (AI), Algorithmic Bias, Human Rights, Privacy, Legal Liability

1. INTRODUCTION

Artificial Intelligence (AI) has transformed itself from just a mere speculative concept of algorithms within the field of computer science to a system that has a persuasive and dynamic force driving the technological transformation in the normal day-to-day world. From flying planes, to running autonomous cars, from autonomous trading through prediction to personalized health care, it has become an integral part of human day-to-day life. Governments and other major entities have been deploying AI to their work due to the increased efficiency, reduction in costs, and enhanced decision-making capability provided by the AI. However, these rapid developments and integration of AI into the normal human world may raise significant legal and ethical concerns, particularly in relation to the protection of fundamental human rights.

This growing use of AI in our day to day lives have started raising concerns regarding the impact on human rights i.e., the problems arising out of the harms which is caused by the AI systems and how these systems can be held accountable for the same as these systems are more widespread it is crucial to confront various ethical, legal, and social consequences they bring in. One of the major concerns that can be seen as the rise of AI systems is the possibility of these systems spreading or perhaps exaggerating the current societal prejudices and problems related to injustices present in society. One of the most concerning issues about AI and human rights is the potential for these systems to propagate and worsen existing biases and discrimination present in the current society.

Along with this, the issue of legal liability remains totally unsettled. Who should be held liable for the work of AI, whether it will be the AI system's parent company or the AI system itself, or if it is just not liable to anyone, is an answer that needs to be taken care of. There is often no clear legal subject to hold accountable. This blurs the line between tool and agent, challenging conventional legal notions of responsibility, negligence, and due process.

By analyzing these interconnected areas, this paper will be seeking to illuminate the latent risks in AI algorithms and machine learning and the risk of AI technologies in general and will be advocating for a more human- centered approach by setting legal frameworks that ensures the accountability while deploying AI systems and AI development which should be aligned with the constitutional values, the human rights values such as right to privacy, equality, international humanitarian rights standards, etc. the ultimate goal is not to halt the technological progress itself but give a more clear guiding parts towards a more sustainable way in a manner where all the human dignity, accountability, and justice is ensured by the dominated by intelligent machines itself.

2. IMPACTS ON MARGINALIZED COMMUNITIES (INDIA FOCUS)

AI as stated earlier depend upon the data it is fed , which generally is via the internet supervised or otherwise. AI in India often relies on data steeped in existing social hierarchies, reproducing caste, religious, gender, and class biases. For instance, Indian police records have historically labeled many Dalits (Scheduled Castes) and Adivasis (Scheduled Tribes) as "habitual offenders" a colonial legacy now digitized and embedded into predictive policing technologies. Most argue that such digitization hides caste discrimination and fosters discriminatory criminalization and surveillance of marginalized groups.[2] New Indian criminal justice statistics indicate that Muslims, Dalits, and Adivasis are disproportionately represented in arrests and jailings. When such AI models trained on data are applied to facial recognition or risk assessment, they lead to increased policing and surveillance in marginalized communities. For instance, during the 2020 Delhi riots, facial recognition software identified and arrested mostly Muslim young people suspecting AI-led religious profiling.[3]

Due to user prejudice, "safety apps" have often picked out and labeled places with significant numbers of Muslims or poor people as unsafe. Such biases make the principles in Articles 14 and 15 of the Indian Constitution equality before the law and preventing discrimination on religious, racial, caste, gender or birthplace grounds unattainable. Article 21 was interpreted by the Supreme Court to include rights such as dignity, living and privacy. AI with biases puts at

[2] Vidushi Marda & Shivangi Narayan, Data in New Delhi's Predictive Policing System, in FAT '20: PROCEEDINGS OF THE 2020 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 1 (ACM 2020).

[3] Ananya Bhattacharya, Delhi Police's Use of Facial Recognition Could Hurt Muslims More, QUARTZ INDIA (Aug. 24, 2021).

risk the same constitutional protections related to get access to welfare, being watched or receiving financial support.

Internationally, Article 7 of the Universal Declaration of Human Rights (UDHR) and Article 26 of the International Covenant on Civil and Political Rights (ICCPR) also mandate equality before the law and protection against discrimination.[4]AI bias in India is also evident in public welfare. In Telangana, the state's Samagra Vedika system linked databases to remove ineligible citizens from welfare rolls. However, the algorithm erroneously flagged many legitimate claimants including elderly widow Bismillah Bee, who was denied food rations because the AI believed her deceased husband owned a car.[5] More than 1.8 million food-security cards were canceled in five years, often without notice or explanation.

These exclusions violate not only the constitutional right to life under Article 21 but also India's obligations under the UDHR, which recognizes the right to food and an adequate standard of living. In employment and fintech, algorithmic bias disadvantages women and marginalized communities. AI-powered hiring tools, if trained on skewed historical datasets, risk excluding SC/ST candidates or women. Similarly, alternative credit scoring based on mobile phone usage favors digitally literate male users from urban areas deepening financial exclusion of women and rural populations.[6]In India, where only 33% of women use mobile internet, such biases can result in systemic denial of credit, loans, or insurance. Public health and education sectors also reveal such exclusions[7]. AI systems designed to optimize urban medical interventions rarely consider rural health indicators like tuberculosis prevalence. AI tutors trained on English-language data may poorly support vernacular or low-income students exacerbating educational gaps.

These gaps are not just technical; they are fundamentally human rights concerns. Marginalized communities are being structurally excluded by automated tools with little accountability. Unless India adopts mandatory algorithmic audits, anti-discrimination safeguards, and redress mechanisms, AI will continue to replicate and amplify social injustice. Legal scholars and

[4] G.A. Res. 217A (III), Universal Declaration of Human Rights, art. 7 (Dec. 10, 1948); International Covenant on Civil and Political Rights art. 26, Dec. 16, 1966, 999 U.N.T.S. 171.
[5] Tapasya, Kumar Sambhav & Divij Joshi, How an Algorithm Denied Food to Thousands of Poor in India's Telangana, AL JAZEERA (Jan. 24, 2024).
[6] Berkeley Center for Long-Term Cybersecurity, A NEW ERA FOR CREDIT SCORING 22–25 (2020).
[7] Berkeley Center for Long-Term Cybersecurity, supra note 5.

rights advocates increasingly call for applying constitutional scrutiny and human rights standards to algorithmic decision-making, as they would to state action or public policy.

3. AI, DATA COLLECTION, AND THE RIGHT TO PRIVACY (INTERNATIONALLY FOCUSED)

Human Rights and Artificial Intelligence

The rapid growth of artificial intelligence (AI) and machine learning (ML) has given rise to now profound implications for human rights. AI technologies intersect with core rights such as privacy, equality, free expression, and due process and also both enabling and challenging ways. Below we examine how rights such as privacy, non-discrimination, freedom of speech, and fair legal process can be affected by AI systems.

3.1 Privacy and Surveillance

AI greatly amplifies the capacity for surveillance and data processing. Facial recognition systems, big-data analytics, and networked sensors can track individuals' movements and behaviors in unprecedented detail. A striking example is China's use of AI-driven surveillance: cities install public screens showing captured images of "jaywalkers" with personal data, turning privacy invasions into public shaming. As Human Rights Watch warns, "facial recognition surveillance… undermines our privacy rights" and poses a threat to freedom of expression and assembly. The embedded image below illustrates this point:

Facial recognition surveillance in China tags a jaywalker on a public screen, exemplifying AI's invasion of privacy and risk to individual dignity.[8][9]

Algorithms also process our phones, social media, and other digital footprints, often without meaningful consent. Companies and governments can infer sensitive details (health, religion, sexual orientation) from data and potentially expose or misuse them. In many countries, AI-driven profiling has raised alarm. For instance, litigation in the United States challenged Clearview AI's practice of scraping billions of photos from the internet to build a face-recognition database. In a recent Chicago court ruling, users alleged Clearview violated Illinois

[8] Nicol Turner Lee & Caitlin Chin-Rothmann, Police Surveillance and Facial Recognition: Why Data Privacy Is Imperative for Communities of Color, BROOKINGS (Apr. 12, 2022), https://www.brookings.edu/articles/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color/.

[9] Paul Mozur, Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras, N.Y. TIMES (July 8, 2018), https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html.

privacy law by collecting faces without consent. The judge approved a landmark settlement requiring the company to compensate victims for this privacy breach.[10]

Internationally, privacy rights enshrined in law are being tested by AI. The European Union's GDPR, for example, guarantees a "right not to be subject to a decision based solely on automated processing" that significantly affects a person, unless safeguards like human intervention are provided. Yet in practice, many AI systems (e.g., credit scoring or insurance decisions) operate opaquely. Civil society groups warn that without transparency and consent, AI erodes the privacy and autonomy that human rights protect.[11]

3.1.1 Inequality and the Discrimination

AI systems can entrench or even amplify existing social biases. Because algorithms learn from historical data, they often replicate prejudices present in that data. Studies have repeatedly found such biases. For instance, when there was a large-scale analysis conducted by National Institute of Standards and Technologies(NIST) found that many facial recognition algorithms were misidentify the African American and Asian faces at 10 to 100 times the error rate for white faces.[12] Similarly, a study in 2016 by ProPublica has documented that there was a criminal-risk algorithm (COMPAS) used in U.S. courts was almost twice as likely to be falsify the label of black defendants as high-risk as compared to white defendants with similar type of records and charges against them. In hiring and finance also, the automated screening was automatically screening out the women or racial minorities: one report noted that the AI hiring tools were trained on their past company's data where they used to often downgrade the qualified candidates just simply because of their gender or zip code.[13]

These patterns mean that AI can violate the right to equality before the law and the prohibition on discrimination. Human Rights Watch points out that facial recognition "exacerbates existing structural inequalities and hits marginalized and vulnerable folks hardest". Predictive policing tools, claiming neutrality, have also led to over-policing of racialized communities. As one

[10] *Id.*

[11] Kashish Maggo, Artificial Intelligence: Impact on Right to Privacy, JURIS CENTRE (Sept. 12, 2023), https://juriscentre.com/2023/09/12/artificial-intelligence-impact-on-right-to-privacy/.

[12] National Institute of Standards &Technology, Face Recognition Technology Evaluation (FRTE) 1:1 Verification, last updated Apr. 25, 2025, https://pages.nist.gov/frvt/html/frvt11.html.

[13] Jeff Larson, Surya Mattu, Lauren Kirchner & Julia Angwin, How We Analyzed the COMPAS Recidivism Algorithm, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

analysis notes, AI-driven surveillance and police tools often *automate prejudice*: areas with historically heavy policing (usually poorer, minority neighborhoods) are predicted as "crime hot spots," leading to even more policing there. In this way, AI can perpetuate cycles of inequality – undermining the ideal that human rights belong equally to all.

3.1.2 Freedom of Expression and Information

AI profoundly affects how information is created, filtered, and shared, with direct consequences for freedom of expression. On one hand, algorithms power social media platforms and search engines that millions use to communicate. They can help connect voices across languages (through translation AI) and curate personalized news feeds. On the other hand, the same technologies can suppress or manipulate expression. Automated content filters (often driven by AI) may wrongly censor lawful speech. Recommendation engines can create the "filter bubbles," which usually is isolating users in a homogenous information environment. There is also a growing concern about the AI-generated disinformation and such type of example can be seen as deep fake videos and synthetic text that can erode trust between masses and spread propaganda which the company might gain favor from, posing challenges to informed free discourse.

These issues have caught international attention. Human rights advocates note that algorithmic surveillance and curation can chill speech and assembly. In oppressive regimes, AI-powered social credit systems and surveillance cameras deter protests and target dissidents. In democratic societies, debates rage over how platforms should moderate hate speech and false content without overreaching. The UN's Advisory Body on AI, for example, has emphasized that AI governance must "prioritize human rights, free expression, and international law" and involve affected communities. Freedom of expression in the AI era thus means guarding against both direct censorship by smart systems and the more subtle distortion of information channels.

3.2 Due Process and Fair Trial

The use of AI in legal and administrative sphere is more or less related to decisions which raises questions about the due process of law and the right to receive a fair trial as the trail can be leveraged towards the general biasness of the AI system possess as mentioned above. Key aspects of due process include the ability to know and how the evidence against you can be challenged or charged. However, AI systems often operate as "black boxes," providing little to no explanation for their outputs and how they have reached to such a conclusion leaves room

for so much error. For instance, some courts use AI risk-assessment tools to inform the bail or actual deliver a sentence, and it is seen that the defendants frequently unable to reach or understand how to reach the same conclusion as the AI has reached through the algorithm. This creates a conflict between the legal principles like the right to confront one's accusers and to have an impartial tribunal.

Recognizing these challenges, some laws now limit automated decision-making. The EU's GDPR enshrines a form of due process for AI: it grants individuals a right to obtain human intervention, to express their viewpoint, and to contest automated decisions that significantly affect them. In practice, legal experts have debated whether fully automatic tools can violate constitutional due process. A landmark analysis by ProPublica highlighted the moral hazard of using COMPAS as de facto evidence, noting that "sentencing should not be easy" and that judges themselves hesitated to rely solely on the software. Similarly, an Amnesty/NIST study warned that biased face recognition not only violates privacy but also endangers the right to fair trial, since an innocent person could be misidentified and prosecuted.

3.3 AI and Human Rights in Practice: Violations and Promises

In the current world data is power and whoever has data can actually change the course of world that is why there are so many countries regulating data and other privacy concerns as a major threat to country and even though most of ai users might be unaware by during giving of consent in the permissions tab other than necessary data there are multiple data set that can be accessed by these companies and AI model as well. The data that is being getting collected, processed and is sold to other AI driven platforms especially to the social media and mobile applications, often operated on the vague consent mechanisms buried in lengthy term of services to trick the user for accessing there database. In current world scenarios, AI both threatens and can advance human rights.[14] On the one hand, there are stark examples of violations. Massive surveillance programs use AI to track citizens (often without accountability). In China, the expansion of facial recognition and social credit has intruded into basic liberties, drawing international criticism. Data-driven policing and immigration checks have ensnared marginalized groups around the globe. Even in free societies, data breaches and opaque profiling infringe privacy (as in the Clearview case), while social media algorithms

[14] Kashish Maggo, supra note 10.

amplify extremist content and misinformation, undermining pluralistic discourse. These harms reflect human rights concerns across privacy, equality, expression, and justice.[15]

On the other hand, AI also offers tools that *promote* human rights, often by empowering human rights defenders and governments to protect rights more effectively. For instance, AI for human rights monitoring is a burgeoning field. Non-profits and international agencies use AI-driven satellite imagery analysis to document war crimes and abuses in remote areas. As one recent review explains, organizations like Amnesty International have deployed AI to identify damage to civilian infrastructure in conflict zones (e.g., destroyed hospitals in Syria). Data analytics and machine learning sift through social media and media reports to detect patterns of hate speech or imminent violence (aiding early warnings in Myanmar and elsewhere). Even smartphones equipped with AI can translate emergency information across languages, assisting journalists and activists. In healthcare, AI diagnosis tools can help realize the right to health by extending medical expertise to underserved regions. These "AI for good" applications illustrate how technology can be harnessed to secure rights – if deployed ethically.

Civil society and governments are responding to AI's dual nature. International bodies have begun to set standards: UNESCO's 2021 Recommendation on the Ethics of AI explicitly ties AI development to human rights norms. The UN's High-Level Advisory Body on AI recommended transparent, rights-respecting governance. Some countries are enacting laws with AI such as the EU's proposed AI Act, for example, bans the most harmful uses of AI (such as facial surveillance in public spaces) and classifies other high-risk AI used for strict such oversight. In March 2025, UN member states adopted their first resolution on AI, addressing that "AI applications should be respecting individual rights, such as privacy, freedom of expression and the right to non-discrimination." These efforts signal a growing consensus that AI must be aligned with human rights principles.[16]

4. THE ACCOUNTABILITY GAP

One of the grave and urgent issues concerning artificial intelligence (AI) involves the legal and ethical responsibility spheres. With AI technologies increasingly being incorporated into

[15] *Id.*
[16] Kashish Maggo, supra note 10.

central areas of transportation, healthcare, law enforcement, and finance, this raises a fundamental question: who is at fault in the case of harm?

At the core of this crisis is what researchers call the "**accountability gap**": a profound disconnect between the decision-making of artificial intelligence systems and prevailing paradigms of human accountability. Unlike previous technologies, AI systems, especially those based on machine learning methods, are autonomous, modifying their outputs to data patterns rather than explicit programming. This autonomy creates a big challenge of attribution when things go wrong: do we hold the creator, the organization that deployed the system, or the provider of the data, or the AI system itself, responsible?

This is exacerbated by so-called "black box" technology. Much artificial intelligence software, particularly that employing deep learning methods, is opaque, and their own authors do not know how they make decisions. Therefore, the victims of biased AI find it extremely difficult to determine where an error or the bias originated, much less issues relevant to proving negligence or fault in court. One notable example of this impunity is the 2018 fatal crash by a self-driving Uber vehicle, which killed an Arizona pedestrian.[17] The vehicle was operating autonomously when its system failed to properly recognize the victim. Investigations revealed a breakdown in both the algorithm and monitoring by human operators, but Uber was not charged criminally, with only the human safety operator being charged with negligence. The incident generated much debate over the assignment of responsibility in hybrid human-machine systems and the difficulties presented by existing tort law in assigning liability in such situations.[18]

Likewise, the application of the COMPAS algorithm within the criminal justice system in the United States is an example of how opaque AI technology can come to perpetuate systematic discrimination with impunity. Investigative reporting uncovered that the algorithm systematically marked Black defendants as high-risk relative to white defendants, even though the actual rates of recidivism did not justify such discrepancies. The proprietary software company, however, exercised proprietary privilege regarding its algorithm and refused accountability for judicial wrongdoing.[19]

[17] Daisuke Wakabayashi, *Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam*, N.Y. Times (Mar. 19, 2018)
[18] Daisuke Wakabayashi, supra note 16.
[19] Julia Angwin et al., Machine Bias, PROPUBLICA (May 23, 2016).

Against this backdrop, the legal frameworks across the EU and the world are beginning to address the issue of accountability in AI. Last year, the European Commission added the AI Liability Directive to go along with the proposed AI Act. The main goal of these frameworks is to develop one simple system for understanding AI liability. Artificial intelligence developers and users are required to tell users about the risks, clearly classify how risky their products are, allow people to monitor their decisions and comply with the law, under the AI Act.

In addition, the United States still relies on statutes and legal guidelines made for each industry such as common law rules. like general principles of common law. The regulatory bodies, e.g., the Federal Trade Commission (FTC),[20] have begun to define discriminatory or biased artificial intelligence systems as crimes under consumer protection laws, and they ask corporations to refrain from unfair or discriminatory algorithmic practices. However, because of the absence of general federal legislation on AI, the victims of such practices typically face extreme difficulties in the legal system.

In India, the regulatory infrastructure is mostly absent. The Digital Personal Data Protection Bill, 2022, addresses some of the aspects of automated decision-making but is short on dealing with more comprehensive issues regarding accountability for harm inflicted by artificial intelligence. The judiciary and legislators have not yet seriously explored the tangled issues regarding the autonomy of AI and its legal consequences.[21]

Legal experts and policy analysts have advocated a range of options for bridging the accountability gap. These involve: Demanding algorithmic impact assessments, which analyze the risk of AI systems prior to deployment; Developing audit trails to improve traceability and transparency of algorithmic choices; Inferring strict liability on organizations that employ high-risk AI, akin to the ultra-hazardous activities tort law doctrine; Creating autonomous AI regulatory bodies or courts that are charged with examining harms and offering redress.

Global standards facilitate such changes. The OECD AI Principles and the UNESCO Recommendation on the Ethics of Artificial Intelligence both emphasize that accountability

[20] Fed. Trade Comm'n, *Aiming for Truth, Fairness, and Equity in Your Company's Use of AI* (Apr. 2021), https://eaccny.com/news/chapternews/ftc-aiming-for-truth-fairness-and-equity-in-your-companys-use-of-ai/
[21] Digital Personal Data Protection Bill, 2022, Bill No. 233-C of 2022 (India).

must be an inherent value in AI governance. As mandated in these guidelines, all the stakeholders within the AI life cycle, i.e., developers, deployers, and regulators, must have clear responsibility for the systems they manage. Ultimately, fixing the accountability deficit will necessitate a change in legal doctrine as well as institutional design. As AI systems increasingly become the source of everyday decisions, the law itself will have to adapt to acknowledge new types of agency and corresponding risk. Legal accountability cannot be permitted to devolve into technical vagueness. Instead, the rule should remain the same: where there is the capability to impact human lives by algorithmic means, there should be accountability.

4.2 Comparative legal Responses

There are multiple organization and states which want to combat the same the unethical use of AI they are the European Union 's AI act that introduces the risk based on the regulations and accountability towards AI including the documentation, human oversight, and penalties for non- compliance by the AI companies. In contract to this the U.S. policy still remains fragmented, with sector- specific approaches and in general limited federal oversight. And similar to EU's and US laws India also jumped into action by providing new draft for the Digital Personal Data Protection Bill signals the growing concerns that the world is facing due to the emergence of AI all the parts that has been added are mentioned below

4.2.1 Indian Constitutional Law and AI

The Indian Constitution, though it was drafted in the mid-20th century, but still provides a robust framework which can be helped to address modern technological challenges through its Fundamental Rights, particularly under Part III.

a. Right to Privacy (Article 21)

The Supreme Court in its landmark judgement of *Justice K.S. Puttaswamy (Retd.) v. Union of India (2017)*,[22]has already recognized the right to privacy as a fundamental right which is given under Article 21 of the Indian constitution. The judgment can be seen in a fashion where it has emphasized on what an informational privacy is and the autonomy in the digital age given the principles that are increasingly challenged by AI-driven surveillance, facial recognition systems, and data-mining practices. There are certain Legal Risk which can be seen as the AI

[22] Justice K.S. Puttaswamy (Retd.) v. Union of India, AIR 2017 SC 4161.

surveillance tools, such as those used in "Aadhaar-linked authentication"[23] or law enforcement facial recognition, may or may not violate the proportionality standard laid out in Puttaswamy unless there is clear legislation ensuring the necessity, proportionality, and procedural safeguards. And it can be said although the Puttaswami case has laid the foundation for protection of rights privacy, but that system can still be questioned as that data of Aadhar card screening is totally give in the hands of the government and to what extent that data can be used without being known by other judicial body.

b. Right to Equality and Non-discrimination (Articles 14 and 15)

AI systems which are getting widely used in recruitment of various jobs and stuff so there is a dire need of policy making in this aspect so the welfare distribution which is risking the violation of Article 14 (equality before the law) and Article 15 (non-discrimination) if they are replicate or exacerbate the structural inequalities will going to increase.[24][25]For example, the predictive policing algorithms may be disproportionately targeting the marginalized communities, which will be violating the equal protection guarantee given under the constitution of India. The use of opaque decision-making systems with no avenue for appeal might also undermine that procedural fairness a very crucial component of Article 14.

c. Intellectual Property Rights (IPR)

While India has a Copyright Act, 1957 which recognizes human authorship,[26] but it does not have shown or given any room to accommodate AI-generated works. This creates a major legal vacuum in determining as to who to give the ownership rights over the work that has been created by AI, raising significant questions on whether under Article 19(1)(g) (freedom to practice any profession or carry on any occupation) the creators work can be protected or not.Therefore, it can be seen that there is a Policy Gap which is there since it can be seen in the modern times that who owns a painting, or a poem generated by AI? The absence of a legal framework addressing AI authorship could hamper the innovation or could also lead to unjust enrichment.

[23] "Unique Identification Authority of India, Which Devices Can Be Used for Face Authentication?", UNIQUE IDENTIFICATION AUTHORITY OF INDIA, last visited May 25, 2025, https://uidai.gov.in/en/304-english-uk/faqs/authentication/for-aadhaar-number-holders/16557-which-devices-can-be-used-for-face-authentication.html.
[24] INDIA CONST. art. 14.
[25] INDIA CONST. art. 15.
[26] The Copyright Act, No. 14 of 1957, INDIA CODE (1957).

4.2.2 European Union: GDPR and Fundamental Rights

The EU have been constantly adopted a rights-centric approach to AI regulation, with the General Data Protection Regulation (GDPR) and the EU Charter of Fundamental Rights serving as the bedrock for digital rights protection. And different articles supporting the same are listed below

d. Data Protection and Privacy (Articles 7 & 8 of EU Charter, GDPR Articles 5–22)

The GDPR actually mandates that the data minimization which means very less data from the users shall be collected and through this the purpose of limitation is achieved, they also mandate under this that there should be explicit consent shall be mentioned and given by the users, and transparency shall be present in AI systems that process personal data. Yes there are some Legal Basis which is mentioned under Article 22 of GDPR that guarantees the right not to be subject to automated decision-making with legal or significant effects.[27] AI systems making credit, hiring, or sentencing decisions must provide meaningful human oversight.[28] A Case Law regarding the same in Schrems II (2020), where the CJEU invalidated the Privacy Shield agreement due to inadequate U.S. surveillance protections thereby highlighting how AI-powered mass data transfers must comply with EU privacy norms.

e. Equality and Bias

The AI Act (proposed in 2021) introduces the concept of high-risk AI systems, includes those that affect the fundamental rights of citizens through access to education, employment, or justice. Discriminatory outcomes in AI are subjected to scrutiny under Articles 20 and 21 of the Charter, which majorly protects against unequal treatment and discrimination.

f. IPR and AI-generated Content

EU copyright law (Directive 2019/790) presumes only the human authorship but leaves AI-generated content in a grey area. The EU Parliament has long debated as to whether AI systems can be or should be granted legal personhood or whether creators of training datasets should be given the rights.

4.2.3. United States: Bill of Rights and Common Law Protections

[27] Regulation (EU) 2016/679 (General Data Protection Regulation), art. 22, 2016 O.J. (L 119) 1.
[28] *Id.*

Just like India and EU, U.S. has also designed some legal framework, although decentralized, which protects the core liberties through the U.S. Constitution's Bill of Rights, supplemented by sectoral regulations.

g. Right to Privacy (Fourth and Fourteenth Amendments)

Although not explicitly stated in US constitution, the right to privacy has been inferred from multiple amendments. There has been a regulation which protects data from AI-driven surveillance and predictive analytics that have been raising constitutional concerns, especially when deployed by the state. There is a very relevant case law as well In Carpenter v. United States (2018), the Supreme Court of USA has ruled that the cell-site location the data collection from that without a warrant violates the Fourth Amendment of US, setting a precedent for all AI surveillance tools.[29]

h. Right to Equality (Equal Protection Clause of the Fourteenth Amendment)

The Discriminatory algorithmic practices which can trigger the equal protection claims, can be particular when used in criminal justice. Tools like COMPAS are being used for recidivism prediction but have been constantly criticized for racial biasness. Through this it can be seen that there is a Legal Vacuum whereby there is no federal law that actually mandates the fairness or transparency in AI decision-making, although states like Illinois (Biometric Information Privacy Act) and New York have started passing targeted AI regulations but there is no one concrete regulation to being with.

i. Intellectual Property Rights (U.S. Copyright Act)

The case of Thaler v. Perlmutter (2023), the U.S. Copyright Office refused to register a copyright for an AI-generated image, asserting that only human authorship is protectable.[30] So this reflects a conservative stance on AI. Also this can be said in the case that work of ai is still not recognized under one domain and can be said that there is a lack of proper bifurcation for the same as to whom can the rights be give to.

5. POLICY FRAMEWORK FOR ETHICAL AI IN INDIA: ELIMINATING ALGORITHMIC BIAS AND UPHOLDING RIGHTS

[29] Carpenter v. United States, 138 S. Ct. 2206 (2018).
[30] Thaler v. Perlmutter, 687 F. Supp. 3d 140 (D.D.C. 2023).

5.1. Data Governance Model for Fair AI

a) Representative, High-Quality Data: Mandate that all (especially high-risk) AI systems use "relevant, sufficiently representative, error-free" training/validation data. Datasets must undergo bias and quality certification before use. For example, the EU AI Act requires high-risk systems to be trained on data free from distortions. India can adopt similar rules, requiring agencies to document data sources, check for population skews, and continually monitor for new biases.

b) Algorithmic Impact Assessments: Require pre-deployment fairness and privacy impact assessments (AIAs) for public-sector AI (analogous to environmental impact statements). Such AIAs would force agencies to identify potential bias, justice, or privacy harms in advance. Evidence shows that internal self-assessments help agencies "better understand and explicate" an AI's disparate impacts before deployment.

c) Continuous Auditing & Transparency: Establish ongoing, independent audits of AI systems to detect bias in live use. Agencies should publish summaries of AI decision-making criteria and allow third-party review. Public registries (inspired by OECD and NY principles) could list deployed AI tools in welfare, policing, credit, etc., increasing transparency.

d) Human Oversight and Redress: Enshrine a right to human review and explanation for significant automated decisions. Like the UK's GDPR (DPA 2018), India should guarantee that individuals "obtain human intervention… express their point of view… and obtain an explanation of the decision and challenge it". All critical AI-driven decisions (e.g. welfare exclusion, policing) must allow appeals and correction of errors, preventing unjust outcomes.

5.2. Constitutional & Human Rights Alignment

a) Equality & Non-Discrimination (Art.14, Art.15): AI systems must uphold Article 14's promise of "equality before the law" and Article 15's ban on caste/religion/sex discrimination. Policy should prohibit algorithms from using caste, religion, or gender proxies unless explicitly required by law for positive measures. As commentators note, unchecked AI often "replicate[s] and amplif[ies] discrimination against people who've borne the brunt of historical oppression". Under the Constitution and ICCPR/UDHR equality norms, any AI output showing unlawful bias (e.g. over-policing of a community) must trigger legal accountability.

b) Right to Life, Liberty and Privacy (Art.21): Article 21 guarantees life, liberty and (by Puttaswamy v. Union of India) informational privacy. Any AI intrusion on privacy or liberty must meet strict tests: it must be lawful, necessary, proportionate and carry procedural safeguards. For instance, mass surveillance or predictive policing infringe privacy unless explicitly authorized, narrowly targeted, and transparent. Our model integrates these principles by requiring legal mandates and oversight for AI uses affecting bodily security or reputation.

c) Economic and Social Rights (Art.21 read with Art.21A/Art.25 etc): AI policies must also respect socio-economic rights (e.g. right to food, healthcare). Automated welfare exclusion threatens the right to food under international law. Aligning with UDHR and ICESCR, regulators should ensure AI in public services does not deprive anyone of essentials. For example, welfare algorithms should default to inclusion, with the burden of proof on the system to justify exclusions.

5.3. Safeguards for Marginalized Communities

a) Inclusive Data & Design: Mandate collection of disaggregated data on caste, gender, disability, and religion to test algorithms. AI developers must involve community representatives from SC/ST, OBC, Adivasi, Muslim, Christian, and other groups in design reviews. This helps detect "proxy" biases (e.g. locality = religion) that entrench marginalization. Metrics should be tracked separately for each group to ensure equitable outcomes.

b) Affirmative Audit Focus: Require that high-risk AI (policing, credit scoring, hiring, benefits distribution) pass stringent fairness audits. For example, algorithms that allocate social benefits must show error rates are uniform across castes and gender. Any significant disparity triggers a halt and redesign.

c) Accessibility and Alternatives: Ensure AI systems do not bar those with limited tech literacy or resources. For instance, if biometric verification (fingerprint/face) is used in welfare, alternatives (PIN, SMS OTP) must be offered so that the elderly, manual laborers or people without smartphones are not excluded. This prevents the 'coerced inclusion' noted by rights groups.

d) Grievance Redressal: Create dedicated grievance cells (within Data Protection Authority or local government offices) where citizens can report AI harms. Victims should not be left "stuck in a bureaucratic maze, with little to no access to remedy".

Instead, a clear appeal channel (like a consumer court for algorithms) must be established, with stipulated response times and oversight.

5.4. Case Studies: AI Bias in India & Model Remedies

a) *Fig: Government memo on Telangana food-subsidy processing.* In Telangana's Samagra Vedika welfare system, linked databases automatically flagged "ineligible" beneficiaries. From 2014–2019 the state cancelled over 1.86 million subsidized food cards, often without notice. Investigations found thousands of wrongful exclusions: e.g. 67-year-old widow Bismillah Bee was falsely tagged as a car-owner and denied rations. Under our model, every automated welfare check would require an audit trail and human review before action. Agencies would have to verify AI rejections with field officers and allow citizens to rebut algorithmic errors (with proof) before benefits are cut. This would catch data mistakes like Bee's (where algorithmic "Syed Ali" was misidentified) and protect citizens' right to food and dignity.

b) *Fig: Telangana resident displays her ration card after Samagra exclusion.* The Samagra case starkly shows how opaque AI can harm the poor. The proposed policy fixes this by insisting on transparency: affected families must be informed of the AI decision and grounds for exclusion, and given a chance to appeal. Enshrining a formal "right to explanation" (as in GDPR/DPA law) means algorithmic denials are reviewable. (It also echoes Art.21 due process: no person should lose a basic entitlement without a fair hearing.)

c) Facial Recognition Abuse: In recent Delhi riots, police used AI face-matching to arrest dozens of men at protests. Notably, those targeted were overwhelmingly Muslims in one case. Civil rights groups warn that without limits, FRT "criminalise minorities and channel[s] … bias to the rich". Our framework would prohibit unregulated use of public facial ID. Any surveillance AI must meet very high accuracy standards in the Indian context, and be operated only with warrants and strict audit logs. For instance, public CCTV face recognition would be banned unless there is court approval and continual accuracy monitoring, protecting privacy and preventing communal profiling.

d) Predictive Policing: Delhi's CMAPS system (and pilots in Jharkhand, Telangana) aimed to forecast crime but ended up surveilling Muslim and lower-caste neighborhoods disproportionately. In effect, it digitized historical police bias against Dalits and Adivasis. The policy model intervenes by declaring such high-risk policing AI impermissible without safeguards. All law-enforcement algorithms must undergo an

Algorithmic Impact Assessment for discrimination, and agencies must report arrests initiated by AI to an oversight committee. Independent panels (including judicial members) would review these reports for bias, and police would have to justify any AI-based targeting under Article 21 standards. This prevents a feedback loop of bias: if the AI starts singling out a community, it triggers review or shutdown.

e) Welfare Authentication: Even outside Samagra, biometric failures have denied aid. For example, daily-wage laborers with worn fingerprints get refused at ration shops. The new rules would ban single-factor ID checks; a lost fingerprint or face scan cannot by itself block access. Multiple authentication methods (Aadhaar PIN, head-of-household verification) must be available, ensuring poor or disabled people aren't unfairly excluded.

5.5. Governance & Oversight Structures

a) Independent AI Authority: Establish a statutory AI Oversight Commission (or empower the Data Protection Authority) to audit and regulate AI. Experts recommend setting up a dedicated "AI Safety Institute" to build technical capacity and guide policy. Such a body would develop standards, conduct market studies, and liaise with global peers. Crucially, it must be institutionalized (e.g. via Parliament) to ensure autonomy.

b) Algorithmic Audit Boards: Mandate periodic external audits of high-risk systems by accredited third parties. Similar to financial audits, auditors would check code and data for bias. OGP research suggests combining internal self-assessments with external review and public comment. In practice, the government could require any AI vendor to submit to an independent "AI conformity assessment" before sale to state agencies.

c) Cross-Agency Committees: Form an inter-ministerial AI Risk Committee (MeitY with NITI, Home, Finance, Health, etc.) to classify AI applications by risk. This mirrors the EU's approach of listing "high-risk" uses needing extra checks. The committee should coordinate training for regulators (RBI, TRAI, insurance authority) on AI risks. For instance, RBI could then issue guidelines on bias-testing in lending algorithms. A Parliamentary Standing Committee on IT could provide legislative oversight of this process.

d) Transparency and Registers: Require that all government AI deployments be logged in a public registry (by department). This allows citizens and watchdogs to know what AI tools are in use. Regular progress reports (inspired by a US Executive Order approach)

would compel agencies to report their AI systems and any harms detected. Such transparency builds trust and accountability, deterring secretive or abusive AI.

5.6. International Best Practices for India

a) EU AI Act – Risk-Based Rules: Adapt the EU's categorization of AI by risk. For "high-risk" systems (e.g. biometric ID, welfare allocation, law enforcement), enforce strict standards. Article 10 of the EU AI Act, for example, mandates high-risk AI be trained on datasets that are "free of errors" and "sufficiently representative". India can mirror this by requiring third-party certification that training data reflect India's diverse population (including caste, region, and socio-economic status).

b) Data Protection (GDPR-like) Safeguards: India's DPDP law should incorporate GDPR-style protections for automated decisions. Notably, it should bar any legal or similarly significant decision made solely by algorithm (echoing GDPR Article 22). For example, UK law (implementing GDPR) explicitly prohibits "significant decisions" without human review, and grants individuals the right to an explanation. Embedding such rights in Indian law ensures citizens can challenge faulty AI outputs.

c) Algorithmic Impact Assessments (AIA): Learn from jurisdictions (e.g. New York City's ADS Task Force) that require algorithmic impact statements. The AI Now Institute recommends that public agencies publish AIAs with each AI rollout, combining self-assessment and public input. India should mandate that before any public AI project, its AIA (covering fairness, privacy, security) be made public and subject to consultation.

d) Global Ethical Principles: Align with OECD and UNESCO AI ethics guidelines on fairness and human rights. For instance, the OECD's AI Principles emphasize transparency, accountability and non-discrimination. India should incorporate these into its policy lexicon to attract international collaboration and ensure its laws meet global standards.

5.7. Implementation Roadmap for Regulators

a) MeitY (Technology Ministry): Empower MeitY to issue binding AI governance rules under the IT Act and forthcoming DPDP Act. Specifically, MeitY should finalize guidelines on automated decisions (in line with its draft AI framework) and mandate DPIAs/AIAs for any project using personal data. It can also set up a dedicated AI auditing arm. For example, following experts' advice, MeitY could either be strengthened to act as regulator or a new AI regulator can be created under its aegis.

b) NITI Aayog and PSA Office: Task NITI with integrating these rules into India's AI strategy. NITI should require that all publicly funded AI initiatives include bias mitigation plans (building on its Ethics Guidelines). It can also commission periodic market studies (as done for digital economy) to detect algorithmic harms. The Office of the Principal Scientific Adviser can coordinate between ministries to ensure a unified approach (similar to how it advises on emerging tech).

c) Sector Regulators & RBI/TRAI/IRDA: Issue sector-specific orders. For example, RBI could direct banks to test credit-scoring models for caste/gender bias before deployment. Election Commission and police boards should ban use of AI profiling in elections and policing without strict oversight. Consumer Protection bodies must treat AI fraud screening as a regulated activity.

By combining data-centric regulations with constitutional safeguards, this framework ensures AI in India serves all citizens fairly. Regular audits, public oversight bodies, and enforceable rights will prevent AI from perpetuating old injustices, while international best practices guide India's unique implementation.

6. CONCLUSION

The dualism in a potential set of benefits from and risks of artificial intelligence integration into core societal institutions is wedded. At the same time, these same AI technologies hold unprecedented promise to improve efficiency, increase accessibility and spur innovation in sectors ranging from health, government, law and education – yet these same technologies present major new risks to core human rights, when deployed without appropriate ethical, legal and social safeguards. Based on this research, we critically analyzed how algorithms bias strengthens the structural discrimination, how the open data collection processes put the privacy in peril and how the lack of legal accountabilities of harms resulting from AI puts us in an extreme accountability lacuna.

In the Indian context, these problems are complicated by a combination of entrenched caste and gender, as well as economic disparities and a lacuna of strong AI specific regulatory framework. Poor or bias data in the implementation of AI in various systems, affect primarily marginalized communities, Dalits, Adivasis, Muslims, women and rural communities. Cases show predictive policing system, credit score algorithms and welfare eligibility system

operating in exclusionary, surveillance or judgmental fashion lacking in due process and transparency, in examples as illustrated. This violates the constitutional guarantee under Articles 14, 15 and 21 of the Constitution, as also India's obligation of international human rights under instruments such as the UDHR and ICCPR.

Confronted by the simultaneous challenge to manage enabling AI innovation but without precedence to ensure the innovation does not occur at the cost of dignity, autonomy or equality, states around the world have attempted to strike a balance. Systemic regulation by the European Union is acting through the AI Act and GDPR which promote accountability, human oversight and data protection. India has issued a draft Digital Personal Data Protection Bill and the United States responded with a sectoral approach. However, we still lack a cohesive rights based legal framework on AI. This article contends that in the face of such new realities, the law must evolve.

The aim is not to stifle innovation but to direct it for artificial intelligence's sake and ours to advance human well being and not exacerbate social injustices. AI at its best is human centered – and rooted in the principles of dignity, equity and responsibility it must be. As we move toward a future governed by ever wiser smart machines, human rights watches should not grow dimmer. AI should be for people and the legal tools that we build today will decide whether AI is empowering or exclusionary.