
ALGORITHMIC ACCOUNTABILITY AND HUMAN RIGHTS: RECONCILING ARTIFICIAL INTELLIGENCE GOVERNANCE WITH INTERNATIONAL AND INDIAN LEGAL FRAMEWORKS

Divya Singh, Ph.D. Research Scholar, Faculty of Law, University of Lucknow, Lucknow,
Uttar Pradesh

Deo Pratap Singh, Ph.D. Research Scholar, History Department, Ram Manohar Lohiya
National Law University, Lucknow, Uttar Pradesh.

ABSTRACT

Artificial intelligence (AI) has transformed decision-making across various sectors creating unprecedented opportunities for efficiency while simultaneously raising complex legal and human rights concerns affecting equality, privacy, dignity and access to justice arising out of their opacity. Existing international human rights instruments, including the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, the UN Guiding Principles on Business and Human Rights, together with the European Union's AI Act, provide an emerging normative framework for responsible AI governance. However, significant regulatory gaps like algorithmic transparency, explainability, attribution of liability and effective remedies persist resulting in discriminatory outcomes. Within India, constitutional guarantees under Articles 14 and 21, judicial recognition of informational privacy and the Digital Personal Data Protection Act, 2023 establish an important foundation, although a comprehensive legal framework addressing algorithmic discrimination remains absent. Adopting a doctrinal and comparative legal methodology, this study analyses international human rights standards, constitutional jurisprudence, statutory and judicial developments to evaluate the adequacy of existing legal safeguards. It further examines the conceptual foundations, challenges and compatibility of AI governance with constitutional values. The paper advocates a rights-based regulatory model founded upon transparency, human rights due diligence, oversight, algorithmic impact assessments and strengthened institutional accountability. It argues that embedding constitutional principles of equality, dignity, fairness and procedural justice within AI regulation is essential to ensuring that technological innovation advances democratic governance without compromising fundamental human rights.

Keywords: Artificial Intelligence, Human Rights, Algorithmic Bias, Constitution, International law, Discrimination.

INTRODUCTION

Artificial Intelligence (AI) is a booming technological mechanism which analyses and produces information in seconds at the order of a human being. Artificial Intelligence is everywhere ([Boden 2016](#)) and its development, deployment and use is moving forward rapidly and contributing to the global economy. AI has many benefits (e.g., improvements in creativity, services, safety, lifestyles, helping solve problems) and yet at the same time, raises many anxieties and concerns (adverse impacts on human autonomy, privacy and fundamental rights and freedoms).¹

Algorithms, commonly defined as computational procedures that use certain inputs in order to generate answers,² have the ability to carry out automated decisions by adapting to and learning from historical data.³ Algorithmic decision-making systems are computational models that use algorithms to analyze data and make decisions or predictions. Algorithmic bias refers to systematic and unfair disparities in results that arise from the design, implementation or application of algorithms. These biases can perpetuate and amplify existing inequalities and discrimination based on factors such as race, gender, ethnicity, socioeconomic status and also add up to being a technological error which is justified under the garb of technology neutrality. The lack of algorithmic transparency is a bubbling issue which forms a major legal discussion on AI.⁴

Algorithmic governance ensure the rights to equality, dignity and non-discrimination as inalienable moral values according to the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR)⁵ and International Convention on the Elimination of All Forms of Racial Discrimination (ICERD). United Nations Guiding Principles on Business and Human Rights (UNGPs)⁶ and the OECD AI Principles make State and non-state actors bound by international human rights. Artificial Intelligence Act of the

¹ OECD 2019.

² J. M. Logg, Julia A. Minson & Don A. Moore, *Algorithm Appreciation: People Prefer Algorithmic to Human Judgment*, 151 *Organizational Behavior & Human Decision Processes* 90 (2019).

³ Felix Pethig & Julia Kroenung, *Biased Humans, (Un)Biased Algorithms?*, 183 *J. Bus. Ethics* 847 (2022), <https://link.springer.com/article/10.1007/s10551-022-05071-8>

⁴ Rowena Rodrigues, *Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities*, 4 *J. Responsible Tech.* 100005 (2020), <https://www.sciencedirect.com/science/article/pii/S2666659620300053>

⁵ *International Covenant on Civil and Political Rights* (ICCPR), opened for signature 16 December 1966, 999 UNTS 171 (entered into force 23 March 1976).

⁶ U.N. Human Rights Council, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework*, U.N. Doc. A/HRC/17/31 (Mar. 21, 2011).

European Union takes a rights-centered and risk-tiered regulatory approach.

India's legislative and judicial responses feature a foundational basis on AI. Constitution of India reflects the aforesaid international ideals under Article 14⁷ and Article 21.⁸ According to the landmark case of *Justice K.S. Puttaswamy v. Union of India*,⁹ informational privacy is crucial to human dignity which ensures that algorithms don't reinforce inequality, reinforce bias or obfuscate accountability. Primary legislation like the Digital Personal Data Protection Act, 2023 has let India embrace digital governance.

The replacement of human discretion with automated oversight challenges long-standing notions of organizational justice, fairness and accountability. Making sure that algorithms' disguised blueprints don't turn into the new patterns of inequality is a supremely human concern that goes beyond basic technical or legal issues.

The study adopts a doctrinal and comparative legal research methodology, analysing international human rights instruments, Indian constitutional provisions, statutory frameworks, judicial precedents and scholarly literature to evaluate algorithmic accountability in artificial intelligence governance. The study primarily examines algorithmic accountability in the context of public governance and constitutional rights, with limited discussion of sector-specific challenges arising in private domains such as finance, healthcare and employment.

ALGORITHMIC ACCOUNTABILITY AND HUMAN RIGHTS LAW

An artificial neural network (ANN) is a conceptual framework for developing AI algorithms. It's a human brain model made up of an interconnected network of neurons connected by weighted communication channels.¹⁰ AI uses various algorithms to find complex non-linear correlations in massive datasets (analytics). Machines learn by correcting minor algorithmic errors (training), thereby boosting prediction model accuracy (confidence).¹¹ The key

⁷ INDIA CONST. art. 14.

⁸ INDIA CONST. art. 21.

⁹ *K.S. Puttaswamy v. Union of India*, (2017) 10 S.C.C. 1.

¹⁰ Naik, N., B.M.Z. Hameed, D.K. Shetty, D. Swain, M. Shah, R. Paul, K. Aggarwal, S. Ibrahim, V. Patil, K. Smriti, S. Shetty, B.P. Rai, P. Chlosta & B.K. Somani, *Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?*, 9 *Frontiers in Surgery* 862322 (2022), <https://doi.org/10.3389/fsurg.2022.862322>

¹¹ Rong G, Mendez A, Bou Assi E, Zhao B, Sawan M. *Artificial intelligence in healthcare: review and prediction case studies. Engineering*. (2020) 6:291301. doi: 10.1016/j.eng.2019.08.015.

David D. Miller & Eric W. Brown, *Artificial Intelligence in Medical Practice: The Question to the Answer?*, 131 *American Journal of Medicine* 129 (2018), <https://doi.org/10.1016/j.amjmed.2017.10.035>

intersection between algorithmic accountability and human rights require a contemporary description today. According to traditional legal theory, accountability requires an identifiable actor, a person or legal entity that has the ability to influence, make decisions and so be held accountable for harm. The issue is more so normative: how can human rights law, which is based on the idea of rational human agency, govern systems that use mathematical reasoning instead of moral judgment to make decisions?¹² There can be two prominent aspects which depict that there is a symbiotic relation between algorithm management and Human Rights-

1) Transitioning from sole responsibility to collective responsibility- The primary documents laying down it's foundation are the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR). Instead of outright violating algorithmic systems operated by private or hybrid enterprises, the state usually supports or controls them. Because of this power imbalance, multi-stakeholder accountability a framework where states, businesses and developers share responsibility for preventing and correcting algorithmic bias must replace state-centric accountability.¹³ This change is formalized by the UN Guiding Principles on Business and Human Rights 2011, which are built around the "Protect, Respect and Remedy" framework that says States must protect human rights, businesses must respect human rights (by conducting human rights due diligence) and victims of business-related human rights abuses must have access to effective remedies.¹⁴ AI systems used for public purposes like welfare targeting, surveillance or predictive policing¹⁵ must adhere to the equality, justice and reasonableness requirements enshrined in the constitution.

2) Algorithmic Bias as a Human Rights Violation - Algorithmic bias is a structural distortion of human rights protections rather than just a computational aberration.¹⁶ The principle of non-discrimination under Article 2 of the UDHR¹⁷ and Articles 14–15 of the Indian Constitution is violated when algorithms systematically disfavor people based on their race, gender, caste or socio-economic background. The Constitution, in Article 14, requires equal legal protection,

¹² Srinivas M.K., *Algorithmic Bias and Human Rights Protection: A Comparative Study of International Instruments and Indian Legal Regimes*, 8(6) International Journal of Law Management & Humanities 739.

¹³ Rodrigues, *Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities*, supra note 4, at 2.

¹⁴ United Nations Development Programme, *Human Rights Due Diligence: An Interpretive Guide* (2021), https://www.undp.org/sites/g/files/zskgke326/files/202210/HRDD%20Interpretive%20Guide_ENG_Sep%202021.pdf

¹⁵ Felix Pethig & Julia Kroenung, *Biased Humans, (Un)Biased Algorithms?*, supra note 3, at 2.

¹⁶ Amal Singh Patel, *Algorithmic Bias and the Quest for Equal Justice in India*, Virtuosity Legal (2025), <https://virtuositylegal.com/algorithmic-bias-and-the-quest-for-equal-justice-in-india>

¹⁷ Universal Declaration of Human Rights art. 2, G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948).

but algorithmic systems can treat similarly situated people differently due to the biases embedded in the training data.¹⁸ Human Rights Committee's General Comment No. 18 views equality as an intrinsic aspect of dignity. This dignity is thus denied by algorithmic decisions that degrade, exclude or misclassify individuals.

Algorithmic opacity violates the Indian Constitution's Article 21 right to procedural fairness and Article 2(3)¹⁹ ICCPR's right to an effective remedy. The procedural legitimacy of the law itself is jeopardized when people are forced to accept algorithmic determinations without being given the opportunity to appeal or seek an explanation. As stated in *Modern Dental College v. State of Madhya Pradesh*²⁰, the Indian judiciary's offers a normative model for AI governance: algorithmic decision-making needs to be reasonable, open and defensible in a democracy.

THEORETICAL AND CONCEPTUAL FOUNDATIONS

Artificial Intelligence is defined as a 'machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.'²¹

Algorithmic accountability emphasizes transparency, explainability and responsibility is essential for ensuring fairness in algorithmic decision-making²²-

- 1) In order to be transparent, oversight organizations and impacted parties must have access to the architecture, reasoning, and data sources of AI systems
- 2) explainability requires that algorithmic outputs be interpretable. The right of the individual to seek an effective remedy is rendered unreal in the absence of interpretability.
- 3) Responsibility means that the results of algorithmic judgments are the responsibility of both

¹⁸ Amal Singh Patel, *Algorithmic Bias and the Quest for Equal Justice in India*, supra note 16, at 5.

¹⁹ International Covenant on Civil and Political Rights art. 2(3)(a), Dec. 16, 1966, 999 U.N.T.S. 171.

²⁰ Casemine Editorial Desk, *Establishing Fairness and Merit in Private Medical Admissions: The Landmark Ruling in Modern Dental College v. State of Madhya Pradesh*, CaseMine (2016), <https://www.casemine.com/commentary/in/establishing-fairness-and-merit-in-private-medical-admissions:-the-landmark-ruling-in-modern-dental-college-v.-state-of-madhya-pradesh/view>

²¹ Stuart J. Russell, Karine Perset & Marko Grobelnik, *Updates to the OECD's Definition of an AI System Explained* (OECD AI Policy Observatory, 2023), <https://oecd.ai/en/work/ai-system-definition-update>

²² Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale Univ. Press 2021).

state and non-state actors.

This underscores the need for clear frameworks that protect employees from unfair practices and ensure transparency in algorithmic decision-making.²³ Therefore, this strategy is operationalized by the European Union's AI Act, 2024, proactive risk assessment and mitigation is similar to the UNGPs' duty of due diligence, Digital Personal Data Protection Act, 2023 offers a partial framework through requirements of data minimization, purpose limitation and consent.

Algorithmic Justice refers to the creation and application of algorithms in a way that is fair, transparent, accountable and considerate of human rights. The main goal is to detect and mitigate biases in automated decision-making to prevent technology from perpetuating or worsening already existing social inequalities in areas like law enforcement.

Algorithmic bias refers to the systematic distortions and unfair preferences or prejudices²⁴ that are embedded within the design, implementation or application of algorithms, leading to discriminatory outcomes. These distortions often replicate or amplify pre-existing human prejudices embedded in the data. AI models rely on training datasets, coding structures and design decisions made by developers. When those inputs reflect social inequalities, the outputs mirror such inequalities in ways that impact decision-making in justice delivery and governance.²⁵ These biases can manifest in various forms, such as favoring certain groups over others, reinforcing existing inequalities or perpetuating stereotypes.²⁶

Digital Discrimination means the unfair, unethical or differential treatment of individuals based on the automated processing of their personal data. It frequently reproduces and amplifies social inequalities²⁷ resulting in discriminatory outcomes against marginalized groups. The Federal Communications Commission (FCC), in a related context, has established a legal definition that is highly relevant, defining digital discrimination to include both

²³ Konjen Hiram, *Algorithmic Management and the Future of Human Work: Implications for Autonomy, Collaboration, and Innovation*, arXiv (2025), <https://arxiv.org/html/2511.14231v1>

²⁴ Anh Dao, *Human Dignity, the Right to Be Heard, and Algorithmic Judges*, British Yearbook of International Law (2020), <https://doi.org/10.1093/bybil/braa009>

²⁵ Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 Calif. L. Rev. 671 (2016).

²⁶ Malcolm Langford, *Taming the Digital Leviathan: Automated Decision-Making and International Human Rights*, American Journal of International Law (2020), <https://doi.org/10.1017/ajil.2020.18>

²⁷ Criado Natalia, *Digital Discrimination* (ResearchGate), https://www.researchgate.net/publication/336792693_Digital_Discrimination
DOI:10.1093/oso/9780198838494.003.0004 In book: Algorithmic Regulation (pp.82-97)

disparate treatment (intentional discrimination) and *disparate impact* (actions that result in discriminatory effects, even if unintentional), where such effects are not justified by genuine issues of technical or economic feasibility.²⁸

RECONCILING TECHNOLOGICAL INNOVATION WITH HUMAN RIGHTS

Human rights law aims to regulate technological progress within the parameters of justice and human dignity, not to oppose it outright.²⁹ Unfairness (Smith 2017)³⁰, bias (Courtland 2018)³¹ and discrimination (Smith 2017)³² repeatedly pop up as issues and have been identified as a major challenge related to the use of algorithms and automated decision-making systems. Therefore, the challenge is to create "anticipatory accountability" legal standards. In order to ensure ex ante consideration of risks to equality, privacy and due process, algorithmic impact assessments (AIAs) must be incorporated as a legislative obligation prior to implementation.

AI systems use faulty or insufficient training data, are badly designed, or function in extremely complex socio-technical environments, the legal risk is increased.³³ These biases manifest as actual deprivations of liberty, unfair treatment in the criminal justice system, a decline in public confidence in legal institutions, detect political opposition, monitor online speech, censor civic involvement, online hate speech or AI-driven amplification of misinformation, implementation of facial recognition and biometric technologies, DNA analysis. The deployment of AI technologies without strong protections directly affects the equality of access to fundamental rights, including privacy, data protection, justice, education, health³⁴, housing and public services. The risk of exclusion or injury is disproportionately high for vulnerable groups as suggested by an European Parliament Report of 2017.³⁵ The reliance on biased data and flawed

²⁸ *FCC Digital Discrimination Rules Will Affect Landlords, Building Owners, Contractors*, Cooley LLP (Feb. 7, 2024), <https://www.cooley.com/news/insight/2024/2024-02-07-fcc-digital-discrimination-rules-will-affect-landlords-building-owners-contractors>

²⁹ Sue Anne Teo, *Artificial Intelligence and Its "Slow Violence" to Human Rights*, 5 AI & Ethics 2265 (2024), <https://doi.org/10.1007/s43681-024-00547-x>

³⁰ *Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making*, Future of Privacy Forum (Dec. 11, 2017), <https://fpf.org/2017/12/11/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making/>

³¹ Davide Castelvetti, *Bias Detectives: The Researchers Striving to Make Algorithms Fair*, Nature (June 20, 2018), <https://doi.org/10.1038/d41586-018-05469-3>

³² *Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making*, supra note 30, at 7.

³³ D.M. Douglas, J. Lacey & D. Howard, *Ethical Risk for AI*, 5 AI & Ethics 2189 (2025), <https://doi.org/10.1007/s43681-024-00549-9>

³⁴ C. J. Colón-Rodríguez, *Shedding Light on Healthcare Algorithmic and Artificial Intelligence Bias*, Office of Minority Health, U.S. Department of Health & Human Services (2023), <https://minorityhealth.hhs.gov/news/shedding-light-healthcare-algorithmic-and-artificial-intelligence-bias>

³⁵ "because of the data sets and algorithmic systems used when making assessments and predictions at the different stages of data processing, big data may result not only in infringements of the fundamental rights of individuals,

algorithms can result in the unjust surveillance, harassment, and profiling of marginalized communities, violating individuals' rights to equal treatment and due process under the law.

Current human rights frameworks have difficulties in actual enforcement in the age of artificial intelligence.³⁶ First, the implementation of rights is made more difficult by the opacity of socio-technical systems; the processes for challenging judgments are frequently not well-suited to algorithmic complexity. Second, traditional rights frameworks might not adequately account for new harms like algorithmically mediated persuasion, secret social manipulation, or automated exclusion from public goods because they were created in pre-digital environments.

From a normative perspective, the adoption of AI in society calls for a reexamination of fundamental legal principles: responsibility, liability, and evidential requirements need to be extended beyond human actors to socio-technical systems that make important choices. The dignity, autonomy, and equality of every person should be protected by legal measures that guarantee technology is used as a tool for inclusion rather than exclusion.³⁷

INTERNATIONAL INSTRUMENTS AND INDIAN ETHICAL-LEGAL FRAMEWORK ON ALGORITHMIC BIAS

A dynamic overlapping between international human rights treaties and India's developing jurisprudence is revealed by the comparative analysis.³⁸

The Universal Declaration of Human Rights 1948 recognised the 'inherent dignity and of the equal and inalienable rights of all members of the human family'³⁹ as the foundational starting point of the human rights framework. Other prominent aspects like privacy (Article 12

but also in differential treatment of and indirect discrimination against groups of people with similar characteristics, particularly with regard to fairness and equality of opportunities for access to education and employment, when recruiting or assessing individuals or when determining the new consumer habits of social media users," European Parliament Resolution of 14 March 2017 on Fundamental Rights Implications of Big Data: Privacy, Data Protection, Non-Discrimination, Security and Law Enforcement, 2016/2225(INI), 15 (Mar. 14, 2017).

³⁶ N. Naik, *Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?* Supra note 10, at 4.

³⁷ Pedro Vitor Marques Nascimento et al., *The Future of AI in Government Services and Global Risks: Insights from Design Fictions*, 13 European Journal of Futures Research (2025), <https://doi.org/10.1186/s40309-025-00253-9>

³⁸ Jillian Rogers, *Artificial Intelligence Risk & Governance*, Wharton Human-AI Research (2023), <https://ai.wharton.upenn.edu/white-paper/artificial-intelligence-risk-governance/>

³⁹ Universal Declaration of Human Rights pmbl., G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948)

UDHR⁴⁰; Article 17 ICCPR⁴¹), equality (Articles 1⁴² and 2⁴³ UDHR; Article 26 ICCPR⁴⁴) and effective remedy (Article 8 UDHR⁴⁵; Article 2(3) ICCPR⁴⁶) are enshrined in the documents.

AI systems' potential to both advance and jeopardize human rights was first acknowledged substantially in the UN Human Rights Council's Resolution 41/11 (2019) on *"New and Emerging Digital Technologies and Human Rights."*⁴⁷ It highlighted the duty of governments to guarantee that AI is created and applied in accordance with current human rights standards. Resolution 47/23 (2021) subsequently reaffirmed the demand for an HRBA (human rights-based approach) to AI, emphasizing responsibility, transparency and non-discrimination as crucial tenets. A focus paper from the EU Agency for Fundamental Rights (FRA 2018) outlines the potential for discrimination against marginalized groups⁴⁸ via algorithms and states that "the principle of non-discrimination, as enshrined in Article 21 of the Charter of Fundamental Rights of the European Union, needs to be taken into account when applying algorithms to everyday life" (FRA 2018).⁴⁹

UN Guiding Principles on Business and Human Rights (UNGPs) expanded the role of non-state actors in human rights obligations by mandating that companies, including tech companies, perform human rights due diligence (HRDD) across their operations. HRDD refers to evaluating datasets, training procedures and decision results for discriminatory effects in an algorithmic environment. In support of inclusive, open and accountable AI systems, the UNESCO Recommendation on the Ethics of Artificial Intelligence 2021 and the OECD Principles on Artificial Intelligence (2019)⁵⁰ further solidified this growing agreement. The United Nations Global Principles for Information Integrity recommends five principles in order to uphold information integrity in the digital age. The five principles – on societal trust and

⁴⁰ Universal Declaration of Human Rights art. 12, G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948).

⁴¹ International Covenant on Civil and Political Rights art. 17, Dec. 16, 1966, 999 U.N.T.S. 171.

⁴² Universal Declaration of Human Rights art. 1, G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948).

⁴³ Universal Declaration of Human Rights art. 2, G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948).

⁴⁴ International Covenant on Civil and Political Rights art. 26, Dec. 16, 1966, 999 U.N.T.S. 171.

⁴⁵ Universal Declaration of Human Rights art. 8, G.A. Res. 217 A (III), U.N. Doc. A/810, at 71 (Dec. 10, 1948).

⁴⁶ International Covenant on Civil and Political Rights art. 2(3)(a), Dec. 16, 1966, 999 U.N.T.S. 171.

⁴⁷ U.N. Human Rights Council, *New and Emerging Digital Technologies and Human Rights: Draft Resolution*, U.N. Doc. A/HRC/RES/47/L.9 (Geneva, 2021).

⁴⁸ European Union Agency for Fundamental Rights, *Bias in Algorithms: Artificial Intelligence and Discrimination* (2022)

⁴⁹ European Union Agency for Fundamental Rights, *Fundamental Rights Report 2018* (2018), https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-fundamental-rights-report-2018_en.pdf

⁵⁰ Organisation for Economic Co-operation and Development, *OECD Principles on Artificial Intelligence* (2019). Artificial Intelligence underscore that transparency, accountability and human oversight are non-negotiable when deploying AI in governance and law.

resilience, promoting healthy (business) incentives, public empowerment, independent, pluralistic and free media and transparency and research, all point towards enabling a healthier information environment.⁵¹

Artificial Intelligence Act (AI Act 2024) of the European Union. The Act establishes statutory requirements such as algorithmic transparency, human oversight and *ex ante* conformance assessments by classifying AI systems into acceptable, high, limited and low risk tiers. It is significant that the Act's preamble specifically cites the EU Charter of Fundamental Rights, establishing human rights as part of the AI legal framework⁵². Also the General Data Protection Regulation (GDPR) incorporates rights of data subjects such as the right to information, erasure, rectification and the right not to be subjected to automated decision-making.⁵³ In turn, those processing data have duties to comply with the principles of data protection – including that of lawfulness, transparency, purpose limitation, accuracy and accountability amongst others.⁵⁴

2) Indian Ethical and Legal Framework

Ethical Principles for Algorithm Deployment- The three ethical aspects in algorithm deployment are- Firstly, fairness and non-discrimination, emphasizing that algorithms should be meticulously crafted to treat all individuals impartially, irrespective of their race, gender or socioeconomic status. Secondly, transparency and explainability i.e. decision-making processes must be within appropriate bounds for ensuring trust⁵⁵ and explainable to relevant stakeholders, including end-users and those affected by the system. This not only fosters accountability⁵⁶ but also empowers individuals to comprehend and challenge algorithmic outcomes, promoting trust and legitimacy. Thirdly, accountability and responsibility i.e. taking

⁵¹ United Nations, *Global Principles for Information Integrity*, <https://www.un.org/en/information-integrity/global-principles>

⁵² Yavuz Selim Balcioglu, Ahmet Alkan Çelik & Erkut Altındağ, *A Turning Point in AI: Europe's Human-Centric Approach to Technology Regulation*, 23 *Journal of Responsible Technology* 100128 (2025).

⁵³ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), arts. 12–23, 2016 O.J. (L 119) 1.

⁵⁴ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), art 5, 2016 O.J. (L 119) 1.

⁵⁵ Nicholas Diakopoulos, *Automating the News: How Algorithms Are Rewriting the Media* 58–61 (2019); Nicholas Diakopoulos & Michael Koliska, *Algorithmic Transparency in the News Media*, 66 *Digital Journalism* 810, 810–12; Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For*, 16 *Duke L. & Tech. Rev.* 18, 39 (2017); Andrew Tutt, *An FDA for Algorithms*, 69 *Admin. L. Rev.* 83, 110–11 (2017).

⁵⁶ Lorna McGregor, Daragh Murray & Vivian Ng, *International Human Rights Law as a Framework for Algorithmic Accountability*, 68 *Int'l & Comp. L.Q.* 309 (2019), <https://doi.org/10.1017/S0020589319000046>

responsibility involves addressing any harms that may arise, acknowledging the societal impact of algorithms, and actively attempting to mitigate adverse consequences. Fourthly, Privacy and data protection involves minimizing data collection, ensuring robust data security measures and obtaining informed consent for the use of personal data. Nissenbaum stated privacy as a form of ‘contextual integrity’ wherein information deemed appropriate in one context should be restricted in others.⁵⁷ Respecting individuals' privacy rights becomes paramount in safeguarding against potential mishandlings and building public trust in algorithmic technologies.

Indian Constitution Safeguards and Judicial Interpretation - The right to equality before the law and equal protection of laws under Article 14 forms the cornerstone of algorithmic fairness. In *State of West Bengal v. Anwar Ali Sarkar*⁵⁸ the Court held that equality demands not only equal laws but equal application. Algorithms reproducing structural inequalities would offend this settled principle. In *E.P. Royappa v. State of Tamil Nadu*⁵⁹, the Supreme Court interpreted arbitrariness as antithetical to equality, thereby embedding fairness as an essential component of access to justice directly resonates with the need to counter opaque algorithmic decision-making.

The right to life and personal liberty under Article 21 has evolved through judicial creativity to include privacy, dignity and informational autonomy. In *Justice K.S. Puttaswamy v. Union of India*, the Court recognised privacy as a fundamental right. In *Anuradha Bhasin v. Union of India*, the Court linked digital access and free expression under Article 19 to democratic participation, implicitly recognizing the human rights dimension of technological regulation.

Statutory and Policy Frameworks - The Digital Personal Data Protection Act (DPDPA) 2023 introduces principles of purpose limitation, data minimization and consent—each essential in mitigating bias and ensuring accountability in AI systems that process personal data. However, the Act remains silent on algorithmic discrimination, explainability or autonomous decision-making. NITI Aayog's “Responsible AI for All” (RAI4A)⁶⁰ framework of 2021 recognizes fairness, inclusivity and transparency as guiding values for AI development. Yet, being policy-oriented rather than legally binding, it lacks enforcement mechanisms. Regulatory bodies such

⁵⁷ Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 Wash. L. Rev. 119 (2004).

⁵⁸ *State of Madras v. V.G. Row*, AIR 1952 SC 196.

⁵⁹ *E.P. Royappa v. State of Tamil Nadu*, AIR 1974 SC 555.

⁶⁰ NITI Aayog, *AI for All: Driving Inclusive Growth Through Artificial Intelligence* (2022), https://www.niti.gov.in/sites/default/files/2022-11/Ai_for_All_2022_02112022_0.pdf

as the Ministry of Electronics and Information Technology (MeitY) and the Data Protection Board (under the DPDPA) are expected to play a greater role in overseeing algorithmic systems.

EMBEDDING HUMAN RIGHTS IN AI REGULATION

The synthesis of international and Indian approaches requires a dual commitment: to global human rights universality and local constitutional specificity.⁶¹ International law provides the normative grammar non-discrimination, transparency, remedy while Indian constitutionalism provides the moral lexicon dignity, reasonableness and social justice. Together, they can form a legal framework. To effectively regulate the dark sides of automated decision-making, Alston sets out a classic human rights approach, a Pareto-optimization logic⁶² in which no individual is made worse off through efficiency improvements. In the case of welfare, he argues that states must ensure that there is a legal basis for digital welfare reforms; promote digital literacy and non-digital access; maintain eligibility fairness and human dignity in procedures; protect civil rights through privacy constraints and limits on use of data to harass and surveil; democratize policy-making on digitalization and hold both public and private actors accountable.⁶³ The European Commission's High-Level Expert Group's 'Recommendation of the Council on Artificial Intelligence' (EU Recommendation) requires AI to include 'human-centred values and fairness' which are defined as 'freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice and internationally recognized labour rights'.⁶⁴

DOCTRINAL CHALLENGES FOR ALGORITHMIC REGULATION⁶⁵

Significant advances have been brought by AI but the key challenges still persist like

⁶¹ Rawa Almakky, *The Role of International Organisations in the Development of International Law: An Analytical Assessment of the United Nations*, 9 Law & World 40 (2023), <https://lawandworld.ge/index.php/law/article/view/422>

⁶² *Pareto Optimal Solution*, ScienceDirect Topics, <https://www.sciencedirect.com/topics/computer-science/pareto-optimal-solution>

⁶³ Report of the Special Rapporteur on Extreme Poverty and Human Rights, Philip Alston (Special Rapporteur on Extreme Poverty and Human Rights), *Report of the Special Rapporteur on Extreme Poverty and Human Rights*, U.N. Doc. A/74/48037, 8 (Oct. 11, 2019).

⁶⁴ High-Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Trustworthy AI* 6 (Eur. Comm'n 2019), https://www.europarl.europa.eu/cmsdata/196377/AI%20HLEG_Ethics%20Guidelines%20for%20Trustworthy%20AI.pdf

⁶⁵ Srinivas M.K., *Algorithmic Bias and Human Rights Protection: A Comparative Study of International Instruments and Indian Legal Regimes*, supra note 12, at 4.

automation bias, selective adherence bias and algorithmic bias. According to Alon-Barkat and Busuioc, excessive trust in automated systems and the consequent ignorance of contradictory information can lead decision-makers to prefer algorithmic advice that confirms pre-existing stereotypes.⁶⁶ Conventional theories of equality, culpability and due process all of which were established in an era of human agency are being pushed to make room for autonomous and opaque technologies.⁶⁷ The underlying assumption of the law that culpability follows intention is becoming more and more incompatible with algorithmic probabilistic logic. The main theoretical conflicts that AI causes with regard to Indian and international human rights legislation are examined in this section.

- 1) The Doctrinal Challenges to Agency and Accountability- The dislocation of legal agency is at the core of the problem. The conclusions drawn by AI systems are difficult to attribute to a single human decision-maker. This opacity, also referred to as the "black box problem," compromises the legal connection between an act and its perpetrator that serves as the foundation for culpability.⁶⁸ As Pasquale argues in *The Black Box Society*, algorithmic opacity concentrates informational power in organizations while concealing decision logic from those it governs, eroding transparency and accountability.⁶⁹ Concerns include diminished autonomy, unclear evaluation criteria and psychological effects of working under constant algorithmic surveillance.⁷⁰ Similar issues have long been debated in Indian law through concepts like vicarious liability and constitutional torts. The Supreme Court emphasized the necessity of compensatory justice in *Nilabati Behera v. State of Orissa*,⁷¹ extending the state's accountability for basic rights abuses even in the absence of express intent. By this logic, a state should be held accountable under Articles 32 and 226 of the Constitution of India for violating fundamental rights if it implements an algorithmic system that results in exclusionary or discriminatory outcomes.

⁶⁶ Sharon Alon-Barkat & Madalina Busuioc, *Human-AI Interactions in Public Sector Decision Making: Automation Bias and Selective Adherence to Algorithmic Advice*, 33 J. Pub. Admin. Res. & Theory 153 (2023), <https://doi.org/10.1093/jopart/muac007>

⁶⁷ Alexandra Huneus & Mikael Rask Madsen, *Between Universalism and Regional Law and Politics: A Comparative History of the American, European, and African Human Rights Systems*, 16 Int'l J. Const. L. 136 (2018).

⁶⁸ *Black Box Problem in AI*, GeeksforGeeks, <https://www.geeksforgeeks.org/artificial-intelligence/black-box-problem-in-ai/>

⁶⁹ Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard Univ. Press 2015).

⁷⁰ *Algorithmic Management and the Future of Human Work: Implications for Autonomy, Collaboration, and Innovation*, arXiv (2025), <https://arxiv.org/html/2511.14231v1>

⁷¹ Unni Krishnan, J.P. v. State of Andhra Pradesh, AIR 1993 SUPREME COURT 2178.

- 2) Due Process in Algorithmic decision- In *Maneka Gandhi v. Union of India*⁷², the Supreme Court held that “procedure established by law” must be just, fair and reasonable. Use of opaque algorithms in judicial process would fall short of this constitutional requirement. The right to an effective remedy under Article 2(3) of the ICCPR⁷³ and the right to a reasoned decision, a component of procedural fairness, was acknowledged by the Indian Supreme Court to have been both violated by this opacity in *Kranti Associates v. Masood Ahmed Khan*.⁷⁴
- 3) Evidentiary Burdens and Algorithmic Opacity - Both domestic and international legal principles of evidence depend on the validity of documents and the dependability of human testimonies. Because algorithmic decision-making creates epistemic opacity, it becomes more challenging for claimants or courts to prove bias or causation. Since the deploying entity must prove that its algorithmic system conforms with equality and fairness standards, the burden of proof must change from the conventional "complainant-based" approach to shared burden.

Human rights jurisprudence is consistent with this inversion. In decisions such as *D.H. and Others v. Czech Republic*,⁷⁵ the European Court of Human Rights (ECtHR) acknowledged that statistical information may be sufficient to shift the burden of proof in matters involving discrimination. According to *Air India v. Nargesh Meerza*,⁷⁶ Indian courts have also recognized circumstantial evidence to infer prejudice. If these standards were applied to algorithmic contexts, the operator would need to provide audit data and algorithmic explainability to establish the system's neutrality when a prima facie incidence of bias is demonstrated, such as by statistical discrepancies or exclusionary consequences.

- 4) Doctrinal Lacunae- The classical debates on legal personhood recognizes natural and juristic persons—human beings and corporations—as rights-holders and duty-bearers. AI challenges this dichotomy by operating as a quasi-autonomous agent without moral or legal subjectivity. Granting AI systems independent personhood, as occasionally proposed in technological jurisprudence, risks diluting human accountability and undermining the moral purpose of law. Rather, the emphasis should be on functional liability—a doctrine that holds accountable the human or institutional entities that design, deploy or benefit from AI systems. This approach preserves the anthropocentric foundation of human rights law, which assumes them to be

⁷² *Maneka Gandhi v. Union of India*, AIR 1978 SC 597.

⁷³ ICCPR art. 2(3).

⁷⁴ 8 September, 2010, available at <https://indiankanoon.org/doc/1304475/>

⁷⁵ 7 February, 2006, available at <https://hudoc.echr.coe.int/fre?i=001-83256>

⁷⁶ *Air India v. Nargesh Meerza*, AIR 1981 SC 1829

deeply embedded in contemporary legal frameworks,⁷⁷ while adapting it to new technological realities. The Indian Supreme Court's decision in *Common Cause v. Union of India*, which emphasized the sanctity of human autonomy in end-of-life decisions, underscores the jurisprudential centrality of human agency—a principle that must remain non-negotiable even in the age of algorithmic governance.

RECOMMENDATIONS

Analysing the cross-sector penetration of AI, certain improvement steps have been taken. The standard way forward as Laukyte suggests, to expand human rights frameworks to encompass 'algorithmic rights' that could include the right to algorithmic transparency and explainability, the right to be informed when interacting with an automated system and the right to choose a human over an algorithm.⁷⁸ An EU Parliament Science and Technology Options Assessment (STOA) study, [2019](#) outlined various policy options to govern algorithmic transparency and accountability, based on an analysis of the social, technical and regulatory challenges; each option addresses different aspect of algorithmic transparency and accountability: 1. awareness raising: education, watchdogs and whistle blowers; 2. accountability in public-sector by use of algorithmic decision-making; 3. regulatory oversight and legal liability; and 4. global coordination for algorithmic governance.⁷⁹ More specific solutions mooted to promote algorithmic transparency include algorithmic impact assessments⁸⁰ an algorithmic transparency standard⁸¹, counterfactual explanations, Local Interpretable Model-agnostic Explanations (LIME) etc. The Council of Europe's Convention on AI, Human Rights, Democracy, and the Rule of Law, which is the world's first binding international treaty on AI,⁸² could be a step forward in developing a rights-based approach to AI.⁸³

⁷⁷ Thomas W. Dornis, *Artificial Creativity: Emergent Works and the Void in Current Copyright Doctrine*, 22 Yale J.L. & Tech. 1 (2020), <https://doi.org/10.2139/ssrn.345148>

⁷⁸ Miglė Laukytė, *Averting Enfeeblement and Fostering Empowerment: Algorithmic Rights and the Right to Good Administration*, 46 Computer L. & Sec. Rev. 105718 (2022).

⁷⁹ European Parliament Resolution of 15 January 2019 on Autonomous Driving in European Transport, 2018/2089(INI) (Jan. 15, 2019).

⁸⁰ Danielle Reisman, Jason Schultz, Kate Crawford & Meredith Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability* (AI Now Inst., 2018), <https://ainowinstitute.org/aiareport2018.pdf>

⁸¹ IEEE Standards Association, *IEEE Standard P7001: Transparency of Autonomous Systems*, <https://standards.ieee.org/project/7001.html>

⁸² Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, C.E.T.S. No. 225 (opened for signature Sept. 5, 2024).

⁸³ Niamh Kinchin, *"Voiceless": The Procedural Gap in Algorithmic Justice*, 32 Int'l J.L. & Info. Tech. eaae024

Other recommendations that may follow are-

First, constitutionalizing AI governance by recognizing algorithmic fairness as an implicit component of Articles 14 and judicial acknowledgment of “algorithmic equality” as integral to the right to life and human dignity 21 of the Constitution of India.

Second, institutionalizing algorithmic oversight by making India aim towards an independent AI regulation and agency (like EU AI Office and the UK Centre for Data Ethics and Innovation) empowered to conduct mandatory algorithmic impact assessments (AIAs), audit AI systems used in public decision-making and provide accessible grievance mechanisms for affected individuals.

Third, embedding human rights impact assessments (HRIA) in public and private policy in line with the UN Guiding Principles on Business and Human Rights. These assessments would evaluate potential risks to equality, privacy and autonomy with findings subject to public disclosure and judicial review, ensuring both transparency and accountability.

Fourth, reimagining law for the algorithmic age where law must shift from ex post remedies to ex ante safeguards

Fifth, an algorithmic audit is a systematic and independent examination of an AI system to assess its fairness, transparency, bias, and efficacy. (4 stages- triage, assessment, mitigation, assurance)

Hence, the measures are promising in theory but their application would depend upon integrated efforts of international and domestic law mechanisms towards accountability.

CONCLUSION

The analysis demonstrates that algorithmic accountability has become an indispensable component of contemporary human rights governance as artificial intelligence increasingly influences public administration, justice delivery and socio-economic decision-making. While AI offers significant opportunities for efficiency and innovation, its deployment simultaneously generates concerns relating to opacity, discrimination, privacy, procedural

fairness and democratic accountability. The study establishes that international human rights instruments, the European Union's regulatory approach and India's constitutional jurisprudence collectively provide a normative foundation for ensuring transparency, explainability and institutional responsibility in algorithmic governance. The ultimate objective is to constitutionalize the algorithm's functioning, not to humanize it. It further argues that effective regulation must move beyond post hoc liability towards preventive mechanisms, including algorithmic impact assessments, independent oversight and human rights due diligence. Embedding constitutional values of equality, dignity and reasonableness within AI governance is essential to ensuring that technological advancement remains compatible with the rule of law. Ultimately, accountable algorithmic systems should reinforce public trust by promoting innovation that remains firmly anchored in human rights and constitutional ethics.