Artificial Intelligence and Constitutional Governance in India: Ensuring Fundamental Rights in the Digital Age

Shreya Dhingra
Assistant Professor, Faculty of Law, Tantia University
doi.org/10.64643/IJIRTV12I6-191393-459

I. INTRODUCTION

India is now one of the world's largest digital democracies, with close to a billion internet users and rapidly expanding deployment of artificial intelligence (AI) across governance, markets, and everyday life. AI systems shape what information citizens see, how they access welfare, whether they receive loans or jobs, and even how law enforcement and courts make decisions. Yet India's constitutional and statutory framework was designed in the mid-twentieth century, long before algorithmic decision-making, deepfakes, or platform-mediated public discourse.

This paper argues that while India's Constitution remains normatively rich, current legal instruments do not adequately address AI's impact on fundamental rights such as equality, free expression, privacy, and dignity. The central claim is that India needs a forward-looking, constitutionally grounded governance structure for AI, anchored in a Digital Bill of Rights and supported by a dedicated Digital Constitutionalism Commission. Drawing selectively on comparative developments such as the EU's General Data Protection Regulation (GDPR)[1] and Artificial Intelligence Act, and on domestic initiatives like NITI Aayog's Responsible AI roadmap, the paper proposes a framework that institutionalizes transparency, fairness, due process, and accountability in AI-enabled environments.

II.CONSTITUTIONAL FOUNDATIONS AND AI

The Indian Constitution places fundamental rights at the heart of democratic governance. For AI regulation, three provisions are especially significant: Articles 14, 19, and 21.

Article 14 guarantees equality before the law and equal protection of the laws. It proscribes arbitrary state action and prohibits discrimination, particularly on grounds such as caste, religion, gender, and disability. The principle has evolved from a thin notion of formal equality to a more substantive understanding that requires the state to address structural disadvantage and indirect discrimination.

Article 19(1)(a) guarantees freedom of speech and expression, including the right to receive information and participate in public discourse. At the same time, Article 19(2) permits reasonable restrictions on specified grounds such as sovereignty, public order, and decency. In the digital context, this dual structure requires the state both to refrain from unjustified censorship and to design regulatory frameworks that enable a plural, informative, and fair public sphere.

Article 21 protects the right to life and personal liberty. Judicial interpretation has expanded Article 21 to encompass privacy, dignity, informational self-determination, and the guarantee that any deprivation of liberty must observe just, fair, and reasonable procedures. This capacious reading provides the primary constitutional basis for digital rights in India.

Traditionally, these provisions have been interpreted as binding the "State," raising what is often described as the "state action" problem: large private platforms and AI developers exercise quasi-public power over speech, association, and access to essential services, yet fall outside classic public-law categories. Recent jurisprudence, however, has started to erode this rigid dichotomy.

[1] European Union, *General Data Protection Regulation*, Regulation (EU) 2016/679.

In K.S. Puttaswamy v. Union of India[2], the Supreme Court recognized privacy as a fundamental right intrinsic to life and liberty. The Court explicitly acknowledged that privacy threats arise not only from state surveillance but also from data-hungry private corporations, creating conceptual space for constitutional scrutiny of private digital harms. In Shreya Singhal v. Union of India[3], the Court struck down section 66A of the Information Technology Act (IT Act), affirming that free speech guarantees extend fully to online expression and that restrictions on digital speech must conform to the strict standards of Article 19(2). In Anuradha Bhasin v. Union of India[4], the Court treated internet access as integral to free speech and the freedom to practice a profession, insisting that network shutdowns must be transparent, necessary, and proportionate.

Taken together, these decisions show a judiciary willing to adapt constitutional doctrine to digital conditions. Yet they still leave a doctrinal gap: private platforms that control visibility, remove content, profile users, and shape information flows wield powers similar to those of traditional public authorities but do not easily fit within existing state-action tests. AI-driven content moderation, recommendation, and profiling thus occur in a regulatory vacuum, affecting fundamental rights without clear avenues for constitutional accountability.

III. EXISTING LEGAL FRAMEWORK: STRENGTHS AND GAPS

India's principal digital statute, the Information Technology Act 2000, predates contemporary AI and social media. Section 79 grants intermediaries limited liability for third-party content provided they follow due-diligence requirements and respond to takedown requests. The Act also offers a legal basis for electronic records and digital signatures, supporting e-commerce and online transactions. However, it is largely silent on algorithmic curation, automated decision-making, and AI-driven harms.

The IT Act does not require platforms to explain or disclose how recommendation systems work, to assess algorithmic bias, or to provide users with procedural safeguards against erroneous or discriminatory automated decisions. Its focus remains on "illegal" content rather than on systemic issues like disinformation amplification, filter bubbles, or opaque ranking that may be constitutionally harmful even when individual pieces of content are lawful.

The Digital Personal Data Protection Act 2023 (DPDPA) marks a significant step forward in privacy regulation. It defines personal data, creates rights to access, correction, and erasure for "data principals," and imposes duties of purpose limitation, data minimization, and security on "data fiduciaries." It also establishes a Data Protection Board empowered to investigate violations and impose monetary penalties.

Yet the DPDPA's focus is primarily on data protection, not on AI governance as such. It does not directly regulate automated profiling, algorithmic decision-making, or platform curation practices that may have profound equality and free-speech implications even if they comply with narrow data-processing rules. Nor does it mandate algorithmic transparency, human-in-the-loop safeguards, impact assessments, or bias audits. Non-personal and anonymized datasets, which frequently serve as training data for high-impact AI models, remain largely outside its scope.

Sector-specific regulations partially address AI in fields such as health, finance, and criminal justice, but the coverage is fragmented and inconsistent. Predictive policing tools, automated welfare distribution systems, and AI-assisted diagnostics are being deployed without uniform statutory standards for transparency, explainability, or equality. This patchwork approach risks leaving individuals exposed to serious rights violations depending on the sector and technology involved.

IV. ALGORITHMIC GOVERNANCE AND THE DIGITAL PUBLIC SPHERE

Modern platforms do not merely host user content; they govern attention through algorithmic curation. Ranking, recommendation, and personalization systems determine which posts or videos are

[2] (2017) 10 SCC 1.
[3] (2015) 5 SCC 1.
[4] (2020) 3 SCC 637

displayed, which accounts grow in influence, and which topics gain or lose visibility. These systems rely on a combination of engagement metrics, inferred user preferences, advertiser interests, and platform norms.

Because most of these algorithms operate as "black boxes," their logics are opaque even to many engineers, let alone to ordinary users or regulators. This opacity has at least three constitutional implications.

First, disinformation and manipulation are structurally incentivized when engagement serves as the primary optimization metric. False or incendiary content often generates more reactions than accurate but less emotionally charged information, and ranking systems can therefore amplify low-quality or misleading material. In a democracy, where citizens require access to reliable information to exercise political judgment, such systemic amplification undermines the conditions for meaningful free speech under Article 19.

Second, filter bubbles and polarization arise when recommendation engines primarily feed users content similar to what they have previously engaged with. Over time, individuals are exposed to narrower informational ecosystems, reinforcing existing views and isolating them from competing perspectives. This narrows the effective marketplace of ideas, inhibits deliberative democracy, and can exacerbate social cleavages, implicating both Article 19 and the equality guarantee of Article 14.

Third, suppression of marginalized voices can occur through algorithmic demotion or biased visibility. Creators from minority communities may receive systematically lower reach or engagement due to algorithmic profiling or skewed training data, even when they comply with platform rules. The resulting asymmetry in visibility reproduces offline hierarchies in the digital realm, undermining substantive equality and equal participation in public discourse.

In parallel, automated content-moderation systems remove or downrank posts at massive scale. While such systems are necessary to manage volume and address genuine harms, they frequently operate without clear notice, reasons, or accessible appeals. Users whose content is removed or whose accounts are suspended often receive generic explanations, if any, and may have no meaningful way to contest decisions. From a constitutional perspective, this raises procedural due-process concerns. Under Maneka Gandhi v. Union of India[5], the Court made it clear that any deprivation of rights must follow just, fair, and reasonable procedures. Although that case concerned state action, the underlying principle—that decisions affecting core interests must be transparent and contestable—provides a useful template for assessing the fairness of platform governance in an AI-mediated public sphere.

Finally, algorithmic profiling affects access to jobs, credit, insurance, and key services. Systems trained on biased historical data may deny opportunities to individuals from certain communities, neighborhoods, or social backgrounds. When such patterns track constitutionally suspect classifications such as caste, religion, or gender, they amount to indirect discrimination in violation of Article 14's substantive equality mandate, as elaborated in cases like Navtej Singh Johar v. Union of India.[6]

V. POLICY INITIATIVES AND COMPARATIVE INSIGHTS

Recognizing the ethical and societal risks of AI, India's policy think-tank NITI Aayog has developed a Responsible AI roadmap[7]. The framework emphasizes principles such as safety, transparency, fairness, accountability, privacy, participation, inclusivity, and innovation. These principles resonate strongly with constitutional values of equality, liberty, and dignity and provide a valuable normative guide. However, the framework is advisory and lacks binding force, enforcement mechanisms, or remedies for affected individuals.

In contrast, the European Union's Artificial Intelligence Act[8] adopts a comprehensive, risk-based regulatory approach. It categorizes AI systems into prohibited, high-risk, and lower-risk tiers, with corresponding regulatory obligations. Unacceptable-risk systems, such as certain forms of manipulative or social-scoring AI, are banned outright. High-risk systems in areas like employment,

[5] (1978) 1 SCC 248.
[6] (2018) 10 SCC 1.
[7] NITI Aayog, *Responsible AI for All* (2021).
[8] European Union, *Artificial Intelligence Act* (2024).

education, law enforcement, credit, and essential services must undergo conformity assessments, implement risk-management systems, ensure human oversight, and maintain detailed documentation. Some categories of AI must also undergo fundamental rights impact assessments prior to deployment, obliging developers and deployers to identify and mitigate rights risks in advance. Enforcement is backed by significant penalties and oversight institutions.

India cannot and should not simply transplant the EU model, given differences in constitutional text, institutional capacity, and developmental priorities. Nonetheless, several features of the EU approach are instructive: differentiating regulation by risk level, focusing explicitly on fundamental rights rather than purely economic harms, embedding ex ante impact assessments, and mandating human oversight for high-stakes decisions. These elements can be adapted to Indian conditions and integrated into domestic constitutional reasoning.

VI. THE DEEPFAKE CHALLENGE

Deepfake technologies illustrate how AI can undermine both individual dignity and democratic processes. By generating highly realistic synthetic audio-visual content, deepfakes can depict individuals saying or doing things they never did. In political contexts, fake speeches or fabricated scandalous material can be used to discredit candidates, activists, or journalists, chilling speech and distorting electoral choice. For targeted individuals, especially women, non-consensual deepfake pornography inflicts severe psychological and reputational harm, implicating Article 21's protection of dignity and bodily integrity. Existing legal tools—such as defamation provisions in the Indian Penal Code and certain IT-related offences—address some aspects of deepfake harms but are inadequate. Detection is technically challenging, attribution is difficult, and post-facto criminal prosecution rarely provides timely relief. Victims often face an uphill battle in securing swift takedown, compensation, or restoration of reputation.

A constitutionally sensitive regulatory approach would impose affirmative obligations on platforms to detect and label synthetic media, provide users with authentication tools, rapidly remove non-consensual sexual content without requiring burdensome victim reporting, and maintain procedures that balance the fight against disinformation with the protection of legitimate political and journalistic speech. Such duties would align with the state's responsibility to secure an environment in which Article 19 rights can be meaningfully exercised and Article 21 dignity is respected.

VII. A DIGITAL BILL OF RIGHTS FOR INDIA

To move beyond fragmented and reactive regulation, this paper proposes a Digital Bill of Rights that extends constitutional principles to AI and platform governance. Six core principles are central to this framework.

1. Algorithmic transparency and explainability

Individuals should be able to understand, at least at a high level, how algorithms affecting their rights operate and what factors they consider. Platforms deploying significant content-ranking, recommendation, or moderation systems should publish accessible explanations of their logic, disclose key design choices, and provide aggregate transparency reports regarding removals, demotions, and appeals. In high-impact domains such as employment, credit, and welfare, affected persons should receive individualized explanations sufficient to contest adverse decisions. Transparency enables the exercise of the right to receive information under Article 19(1)(a) and supports reasonableness review under Article 14.

2. Due process and right to appeal

Decisions that significantly affect users' interests—such as account suspensions, demonetization, or automated rejection of benefits and credit—should follow fair procedures. At a minimum, users should receive timely notice, clear reasons for the decision, and a genuine opportunity to appeal to a human reviewer. Scalable mechanisms such as tiered review and independent ombuds processes can operationalize this principle without halting innovation. Embedding due-process requirements into AI governance aligns with Article 21's insistence that any deprivation of liberty or livelihood be just, fair, and reasonable.

3. Non-discrimination and algorithmic fairness

AI systems must not perpetuate or amplify discrimination on grounds such as caste, religion, gender, disability, or sexual orientation. Before deployment, developers of high-risk systems should conduct fairness assessments, including tests for disparate impact on protected groups. After deployment, platforms and deploying entities should engage in ongoing monitoring, auditing, and recalibration. Individuals who suffer harm from discriminatory algorithms should have access to effective remedies, including correction, retraining, or compensation. This principle translates Article 14's equality mandate into the algorithmic age.

4. Data minimization and purpose limitation
Personal data should be collected and processed only to the extent necessary for specified, legitimate purposes, and not repurposed in ways that undermine privacy or autonomy. Explicit consent for using personal data in AI training should be meaningful and revocable, and retention periods should be proportionate. Governance of ostensibly anonymized or aggregated datasets should address re-identification risks. This principle deepens the DPDPA's protections and gives concrete content to Article 21's privacy jurisprudence.

5. Public participation in AI governance
Decisions about high-impact AI systems that shape public discourse or access to essential services should not be left solely to technocrats or corporate actors. Civil society organizations, affected communities, and democratic institutions should have a voice in setting norms, assessing risks, and designing safeguards. Participatory processes—such as public consultations, multi-stakeholder forums, and citizen panels—can help ensure that AI governance reflects constitutional commitments to democracy, association, and expression under Articles 19(1)(a) and 19(1)(c).

6. Platform accountability and independent oversight
Platforms should not be able to disclaim responsibility by invoking algorithmic autonomy. Clear lines of accountability must be established, identifying which entities—developers, deployers, or platform operators—bear obligations and potential liability for algorithmic harms. Regular transparency reporting, mandatory audits for large or high-risk systems, and structured data access for independent researchers (with privacy safeguards) can enable democratic scrutiny. This principle concretizes the Constitution's vision of limited, accountable power in a context where private actors increasingly perform quasi-public functions.

VIII. INSTITUTIONAL IMPLEMENTATION: A DIGITAL CONSTITUTIONALISM COMMISSION

To give legal force to the Digital Bill of Rights, India should consider creating a specialized Digital Constitutionalism Commission. This body would not replace existing regulators but coordinate and complement them, ensuring that AI and platform governance remain anchored in constitutional values.
The Commission's mandate could include: issuing binding standards and guidelines for algorithmic transparency, fairness, and due process; conducting or overseeing fundamental rights impact assessments for designated high-risk AI systems; receiving and investigating complaints from individuals and communities affected by algorithmic harms; facilitating independent audits by researchers and civil society; and advising Parliament and the executive on emerging digital-rights challenges such as foundation models, immersive environments, and new forms of automated surveillance.
Composition would be interdisciplinary, including constitutional lawyers, computer scientists, social scientists, technologists, and representatives of civil society and marginalized communities. Such pluralism is necessary to capture the technical, legal, and social dimensions of AI and to secure democratic legitimacy.
By coordinating with the Data Protection Board, competition authorities, sectoral regulators, and courts, a Digital Constitutionalism Commission could help avoid regulatory fragmentation, close gaps in protection, and ensure that India's transition to an AI-enabled society is governed by coherent, rights-respecting norms.

IX. CONCLUSION

India's constitutional framework is not obsolete in the face of AI; its core values of liberty, equality, and dignity are precisely what is needed to guide

technological development. The real challenge lies in updating legal doctrines, statutes, and institutions so that they can respond effectively to new forms of power and harm. Opaque algorithms that shape public discourse, biased profiling systems that reproduce structural discrimination, and deepfakes that corrode trust and dignity all raise issues squarely within the Constitution's normative domain.

This paper has argued that existing instruments—the IT Act, DPDPA, and scattered sectoral rules—are insufficient to meet these challenges. Comparative developments, especially the EU's risk-based AI regulation, and domestic policy work like NITI Aayog's Responsible AI principles, offer important lessons but must be adapted to Indian conditions. A Digital Bill of Rights, supported by a Digital Constitutionalism Commission, would provide a systematic, rights-based framework for governing AI and platform power. Ultimately, the task is to ensure that the deployment of AI in India strengthens, rather than erodes, the constitutional promise of a democratic, just, and humane social order.

REFERENCES

[1] Anuradha Bhasin v. Union of India, (2020) 3 SCC 637. Supreme Court of India.

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

[3] Citron, D. K., & Chesney, R. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753-1820.

[4] Digital Personal Data Protection Act, 2023 (No. 22 of 2023). Government of India.

[5] European Parliament and Council of the European Union. (2016). *General Data Protection Regulation (GDPR)*, Regulation (EU) 2016/679. Official Journal of the European Union, L119.

[6] European Parliament and Council of the European Union. (2024). *Artificial Intelligence Act*, Regulation (EU) 2024/1689. Official Journal of the European Union, L 248.

[7] Information Technology Act, 2000 (No. 21 of 2000). Government of India.

[8] K.S. Puttaswamy v. Union of India, (2017) 10 SCC 1. Supreme Court of India.

[9] Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis, 10*, 113-147.

[10] Maneka Gandhi v. Union of India, (1978) 1 SCC 248. Supreme Court of India.

[11] Ministry of Electronics and Information Technology (MeitY). (2023). *India AI Mission*. Government of India.

[12] NITI Aayog. (2021). *#AIForAll: Responsible AI - A Roadmap for Artificial Intelligence*. Government of India.

[13] Navtej Singh Johar v. Union of India, (2018) 10 SCC 1. Supreme Court of India.

[14] Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

[15] O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing.

[16] Parthasarathi, A., & Sundaresan, S. (2022). Platform power and constitutional rights: The case for horizontal application. *National Law School of India Review, 34*(1), 45-78.

[17] Png, M. T. (2020). The impact of privacy regulation on AI development. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 427-433.

[18] Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33-44.

[19] Shreya Singhal v. Union of India, (2015) 5 SCC 1. Supreme Court of India.

[20] Telecom Regulatory Authority of India (TRAI). (2023). *The Indian Telecom Services Performance Indicators Report*. Government of India.

[21] Ullah, F., et al. (2022). Deepfake detection challenges, datasets, and detection approaches. *IEEE Access, 10*, 31807-31834.

[22] Vyas, N., & Dhingra, S. (2024). Constitutional limits on algorithmic governance: Lessons from India. *Journal of Indian Law and Technology, 21*(1), 112-135.

[23] World Bank. (2023). *Digital India: Technology to transform a connected nation*. World Bank Group.

[24] Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

[25] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 1-15.

[26] D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

[27] Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

[28] Fourcade, M., & Gerry, R. (2020). Just algorithms? Governing artificial intelligence. *Sociological Theory, 38*(4), 278-302.

[29] Ganguly, G. (2023). Deepfakes and democratic backsliding: Evidence from India. *Asian Journal of Law and Society, 10*(2), 345-367.