



# The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

## ALGORITHMIC BIAS AND CONSTITUTIONAL MORALITY: EXAMINING THE NEED FOR AN AI ACCOUNTABILITY FRAMEWORK IN INDIA

*Aradhya Singh*

### ABSTRACT

Driven by rapid technological integration, Artificial Intelligence (AI) has transitioned from an efficiency tool into a primary driver of public administration and private enterprise. However, its widespread deployment has exposed a critical vulnerability: algorithmic bias. When machine learning models rely on historical data, they risk codifying and amplifying systemic inequalities, converting societal prejudices into automated decisions.<sup>1</sup> In India, this technological mismatch poses a severe challenge to constitutional morality—the foundational ethos of justice, liberty, and equality enshrined in the Constitution. Left unchecked, biased algorithms directly undermine fundamental rights, particularly the guarantees of equality before the law and non-discrimination under Articles 14 and 15, as well as the right to life and personal liberty under Article 21.<sup>2</sup> This research article examines the growing friction between automated decision-making and constitutional protections in the Indian context. Utilizing a doctrinal and comparative methodology, the study analyses emerging domestic jurisprudence alongside global AI governance benchmarks, such as the European Union's AI Act. The findings reveal deep-seated regulatory gaps within India's current legal framework, particularly the Information Technology Act, which lacks the specificity required to address algorithmic opacity, proxy discrimination, and developer accountability. Furthermore, the research demonstrates that treating technology as entirely neutral masks its disparate impact on marginalized communities. Ultimately, this paper argues that preserving constitutional

<sup>1</sup> See Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671, 677–80 (2016).

<sup>2</sup> INDIA CONST. arts. 14, 15, 21.

morality in the digital age requires moving beyond voluntary ethical guidelines. It proposes a robust, legally binding AI Accountability Framework for India that mandates algorithmic auditing, clear data-provenance standards, and independent oversight bodies to ensure that AI development remains tethered to democratic values and constitutional guarantees.

## **INTRODUCTION: THE RISE OF ALGORITHMIC GOVERNANCE**

A citizen is denied a critical government benefit because an Artificial Intelligence (AI) system automatically classifies him as "high risk." He never receives an explanation. No human official takes responsibility, pointing instead to the code. Has technology merely automated bureaucracy, or has it created an entirely new form of constitutional injustice?

This hypothetical is rapidly becoming a reality as India transitions into an era of algorithmic governance. Over the last decade, AI has evolved from an experimental efficiency tool into the administrative backbone of both the public and private sectors.<sup>3</sup> From automated welfare distribution, facial recognition in policing, and judicial case management to private sector hiring, credit scoring, and health diagnostics, algorithms increasingly dictate the allocation of rights, resources, and liberties. This widespread shift has birthed the phenomenon of algorithmic decision-making, where complex mathematical models replace human discretion. In the pursuit of optimization, state and private actors have placed an unprecedented reliance on predictive technologies systems designed to analyse past behaviours to forecast future risks, worthiness, or criminality.<sup>4</sup>

However, this rapid integration is built upon a deeply flawed premise: the myth of technological neutrality. For too long, software and mathematical models have been viewed as objective, infallible arbiters free from human prejudice. In reality, AI systems are not decoupled from society; they are trained on historical data that reflects existing systemic inequalities.<sup>5</sup> When machine learning models ingest biased datasets, they inevitably codify, automate, and amplify those prejudices under a veneer of mathematical objectivity. Far from being neutral, algorithmic bias transforms historical societal inequities into automated, self-replicating decisions.<sup>6</sup>

---

<sup>3</sup> See Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 Harv. J.L. & Tech. 353, 356–59 (2016).

<sup>4</sup> See Karen Yeung, *Algorithmic Regulation: A Critical Assessment*, 11 Regulation & Governance 505, 508–12 (2018).

<sup>5</sup> See Tarleton Gillespie, *The Relevance of Algorithms*, in *Media Technologies: Essays on Communication, Materiality, and Society* 167, 172–75 (Tarleton Gillespie et al. eds., 2014).

<sup>6</sup> See Barocas & Selbst, *supra* note 1, at 677–80.

In India, this technological mismatch creates a profound friction with the core tenets of the legal system, bringing into sharp focus the urgent relevance of constitutional values. The Indian Constitution is not a static text but a living instrument governed by "constitutional morality"—the foundational ethos of justice, liberty, and equality.<sup>7</sup> Automated decisions that operate as an impenetrable "black box" directly clash with these values. When an algorithm discriminates against a marginalized group or denies a citizen a right without recourse, it strikes at the heart of Article 14 (Equality before the law) and Article 15 (Prohibition of discrimination).<sup>8</sup> Furthermore, the lack of transparency and human accountability inherent in automated systems undermines the "procedure established by law" guaranteed under Article 21 (Right to life and personal liberty),<sup>9</sup> reducing due process to a digital lottery.

## UNDERSTANDING ARTIFICIAL INTELLIGENCE AND ALGORITHMIC BIAS

### A. WHAT IS ARTIFICIAL INTELLIGENCE?

To understand how automated systems intersect with legal systems, it is essential to first demystify the technical landscape of Artificial Intelligence (AI). AI is not a singular technology but an umbrella term encompassing various systems capable of performing tasks that historically required human intelligence.<sup>10</sup> At its foundation lies *machine learning* (ML), a method of data analysis that automates analytical model building. Instead of being explicitly programmed to perform a task, ML systems use computational algorithms to learn iteratively from data, identifying underlying patterns to make decisions with minimal human intervention.

A specialized and more complex subset of machine learning is *deep learning*. Modelled loosely on the human brain's neural networks, deep learning utilizes multi-layered artificial neural networks to process unstructured data—such as image, video, and natural language text.<sup>11</sup> Because these networks contain hidden layers, they can autonomously extract high-level features from raw input, though this often creates a "black-box" phenomenon where the system's exact reasoning path becomes hidden from human view.

In public administration and corporate sectors, these systems are primarily deployed through *predictive analytics*. Predictive analytics utilizes historical data, statistical algorithms, and machine learning techniques to identify the likelihood of future outcomes based on historical

---

<sup>7</sup> See Constituent Assembly Debates, Vol. VII, Nov. 4, 1948 (statement of Dr. B.R. Ambedkar).

<sup>8</sup> INDIA CONST. arts. 14, 15.

<sup>9</sup> INDIA CONST. art. 21.

<sup>10</sup> See Scherer, *supra* note 3, at 356.

<sup>11</sup> See *id.*

trends.<sup>12</sup> When these predictions are operationalized to execute actions, they culminate in *automated decision-making* (ADM). ADM systems ingest user data, apply predictive models or pre-defined rule sets, and autonomously generate an output—such as approving a loan, shortlisting a job candidate, or flagging a citizen for tax fraud—effectively replacing human discretion with automated execution.

## B. WHAT IS ALGORITHMIC BIAS?

Algorithmic bias refers to systematic and repeatable errors within a computer system that create unfair outcomes, typically by privileging one arbitrary group of users over others.<sup>13</sup> From a scholarly perspective, algorithmic bias is rarely understood as a purely technical glitch; rather, it is viewed as a socio-technical phenomenon where existing human prejudices are translated into mathematical code. Scholars note that algorithms do not operate in a vacuum; they reflect the biases embedded within the societies that build and train them.<sup>14</sup>

In legal and sociological scholarship, algorithmic bias is defined as the automated reproduction of systemic discrimination. Because mathematical model's optimization processes are designed to find patterns within past data, they treat historical inequalities not as systemic failures to be corrected, but as neutral metrics to be replicated.<sup>15</sup> Consequently, algorithmic bias is the mechanism by which historical, societal, and institutional prejudices are codified, institutionalized, and masked behind a veneer of computational objectivity.

## C. SOURCES OF ALGORITHMIC BIAS

The manifestation of bias within an AI system can occur at multiple stages of its lifecycle, originating from distinct technical and cognitive sources:

- **DATA BIAS:** This occurs when the historical discrimination present in society is accurately captured by, and reflected in, the training datasets. If a historical dataset reflects systemic prejudice—such as lower credit approval rates for women or higher arrest rates for marginalized castes—the algorithm will interpret these skewed patterns as objective truths, thereby automating and perpetuating historical discrimination into future decisions.<sup>16</sup>

---

<sup>12</sup> See Yeung, *supra* note 4, at 508.

<sup>13</sup> See Barocas & Selbst, *supra* note 1, at 677.

<sup>14</sup> See Gillespie, *supra* note 5, at 172–75.

<sup>15</sup> See Barocas & Selbst, *supra* note 1, at 694–98.

<sup>16</sup> See *id.*

- **SAMPLING BIAS:** This arises when the data used to train a machine learning model does not accurately represent the actual population the model will encounter in the real world. For example, if a facial recognition system is trained primarily on datasets consisting of lighter-skinned faces, its accuracy drops significantly when deployed against diverse populations, leading to high rates of misidentification and systemic exclusion.
- **DESIGN BIAS:** This form of bias is introduced by the programmers, engineers, and data scientists during the structural development of the algorithm. It stems from human assumptions built directly into the code, such as how a problem is framed, which variables are selected as metrics for success,<sup>17</sup> and how different data points are weighted. If an engineer incorrectly assumes a specific proxy variable represents merit, that subjective assumption becomes hardcoded as an objective rule.
- **FEEDBACK LOOP BIAS:** Also known as self-reinforcing discrimination, this occurs when an algorithmic output influences real-world actions, which then generate new data that validates the initial biased output. For instance, a predictive policing algorithm may flag a marginalized neighbourhood as "high risk" based on flawed historical data. Police forces are subsequently concentrated in that area, leading to more arrests. These new arrests are fed back into the system, "proving" the algorithm's original biased prediction correct.
- **AUTOMATION BIAS:** This represents a cognitive source of bias where human operators blindly trust algorithmic outputs, favouring automated suggestions over their own professional judgment or contradictory evidence.<sup>18</sup> When judges, police officers, or welfare administrators treat an algorithmic risk score as infallible, they lose their capacity for critical human oversight, allowing automated biases to be executed without scepticism or constitutional scrutiny.

## CONSTITUTIONAL MORALITY: THE CONSTITUTIONAL COMPASS FOR EMERGING TECHNOLOGIES

### A. MEANING AND EVOLUTION

---

<sup>17</sup> See Sandra Wachter, Brent Mittelstadt & Chris Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, 41 Computer L. & Security Rev. 1, 12–15 (2021).

<sup>18</sup> See Linda J. Skitka, Kathleen Mosier & Mark Burdick, Accountability and Automation Bias, 52 Int'l J. Hum.-Comput. Stud. 701, 703–05 (2000).

The concept of constitutional morality in India provides a vital framework for reconciling rapid technological change with foundational democratic principles. The lineage of this doctrine traces back to the framing of the nation, explicitly anchored in Dr. B.R. Ambedkar's vision during the Constituent Assembly Debates.<sup>19</sup> Drawing from the classical historian George Grote, Dr. Ambedkar defined constitutional morality as a "paramount reverence for the forms of the Constitution," coupled with a disposition to obey constitutional authorities and act within the strictures of legal procedures.<sup>20</sup> He cautioned that a democratic text is only as effective as the spirit of those administering it, famously observing that democracy in India was merely "a top-dressing on an Indian soil, which is essentially undemocratic."<sup>21</sup> For Ambedkar, constitutional morality was the tool required to bridge the gap between political equality (one vote, one value) and deep-seated social and economic inequalities.

Rather than serving as a static rule of law, constitutional morality has evolved in Indian jurisprudence as an active mechanism to protect minority interests, individual autonomy, and institutional integrity.<sup>22</sup> It dictates that state actions must be evaluated not just against the literal text of the Constitution, but against its underlying conscience. It directly challenges majoritarian impulses, stipulating that the popular or social morality of a given historical moment must yield to the supreme values of the constitutional framework.

## **B. DEVELOPMENT THROUGH JUDICIAL INTERPRETATION**

The contemporary revival of constitutional morality has been driven by the Supreme Court of India, which transformed it from an administrative virtue into a robust tool for judicial review:

- **GOVERNMENT OF NCT OF DELHI V. UNION OF INDIA, (2018) 8 SCC 501:**  
In this seminal federalism dispute, the Supreme Court utilized constitutional morality to define the parameters of institutional power.<sup>23</sup> The Court held that constitutional morality requires public officials to act with institutional civility, fidelity to constitutional processes, and a commitment to collaborative governance. Justice D.Y. Chandrachud noted that constitutional morality requires filling in constitutional

---

<sup>19</sup> See Constituent Assembly Debates, Vol. VII, Nov. 4, 1948 (statement of Dr. B.R. Ambedkar).

<sup>20</sup> *Id*

<sup>21</sup> *Id*

<sup>22</sup> See Afreen Afshar Alam & Gautam Bhatnagar, Constitutional Morality: The New Instrument of Justice?, 2025 *Revista de Drept Constituțional* 79, 83–85 (2025).

<sup>23</sup> *Government of NCT of Delhi v. Union of India*, (2018) 8 S.C.C. 501 (India).

silences to enhance and complete the spirit of the document, ensuring that public institutions survive through shared, accountable sovereignty.<sup>24</sup>

- **NAVTEJ SINGH JOHAR V. UNION OF INDIA, (2018) 10 SCC 1:** Decriminalizing consensual same-sex relations by reading down Section 377 of the Indian Penal Code,<sup>25</sup> the five-judge Bench established the absolute supremacy of constitutional morality over social or popular morality. The Court declared that the fundamental rights of a marginalized community cannot be compromised based on populist sentiments or traditional prejudices.<sup>26</sup> Judges are bound by an oath to uphold the constitutional values of dignity, liberty, and equality, using constitutional morality as an interpretive resource to foster social self-renewal and individual emancipation.

### C. CORE ELEMENTS

When deconstructed, constitutional morality comprises a matrix of interconnected principles that collectively drive "transformative constitutionalism"—the doctrine that the Constitution aims to progressively transform society from an unequal social order into a vibrant, rights-respecting democracy.<sup>27</sup>

- **JUSTICE & EQUALITY:** Moving past formal equality, these elements demand a substantive commitment to dismantling systemic disparities. They protect individuals from arbitrary exclusion and discriminatory stratification, ensuring fair distribution of rights and resources.
- **LIBERTY & DIGNITY:** These values guarantee individual autonomy, bodily privacy, and the right to live free from non-consensual state or corporate surveillance. Dignity is treated as an absolute, non-negotiable core of human existence.
- **ACCOUNTABILITY & RULE OF LAW:** This requires that all exercises of power—whether by human bureaucrats or automated systems—must be transparent, explainable, and bound by predictable, reviewable legal procedures, rather than arbitrary discretion.

### D. THE NORMATIVE FOUNDATION FOR AI GOVERNANCE

---

<sup>24</sup> See id.

<sup>25</sup> Navtej Singh Johar v. Union of India, (2018) 10 S.C.C. 1 (India).

<sup>26</sup> See Id

<sup>27</sup> See Id

As algorithmic systems increasingly replace human discretion, **constitutional morality must become the normative foundation for AI governance.** Current global conversations often rely heavily on corporate "AI Ethics" or voluntary principles, which lack enforcement mechanisms<sup>28</sup> and carry no democratic legitimacy. Grounding AI governance in constitutional morality shifts the focus from voluntary corporate compliance to mandatory constitutional obligations.

When an algorithm executes an automated decision—whether allocating welfare benefits, assessing credit risk, or deploying predictive policing—it operates as a functional proxy for the state or a powerful commercial entity. If that algorithm relies on historical datasets that embed societal prejudices, it perpetuates systemic discrimination. Under a framework of constitutional morality, an algorithmic "black box" that produces biased or unexplainable outputs is not treated merely as a technical error; it is recognized as a direct violation of Articles 14, 15, and 21.<sup>29</sup>

By utilizing constitutional morality as a compass, India can construct a legally binding accountability framework. This framework treats technological systems not as neutral tools, but as socio-technical instruments that must be explicitly engineered to preserve dignity, guarantee substantive equality, and submit to the rule of law.

## **ALGORITHMIC BIAS AND ARTICLE 14: THE CONSTITUTIONAL CHALLENGE OF EQUALITY**

### **A. EQUALITY BEFORE LAW**

Article 14 of the Constitution of India guarantees to all person's equality before the law and the equal protection of the laws within the territory of India.<sup>30</sup> Historically, the jurisprudence governing this fundamental right was confined to the "doctrine of reasonable classification." Under this traditional test, differential treatment by the State does not violate Article 14 provided it satisfies two conditions: it must be founded on an intelligible differentia that distinguishes persons or things grouped together from others left out, and that differentia must bear a rational nexus to the object sought to be achieved by the challenged action.

---

<sup>28</sup> See Sandra Wachter & Brent Mittelstadt, A Right to Reason to an Alternative in the Western Legal Tradition: Institutionalizing the Black Box of All-Algorithms, 14 Law, Innovation & Tech. 211, 215–18 (2022).

<sup>29</sup> INDIA CONST. arts. 14, 15, 21.

<sup>30</sup> INDIA CONST. art. 14.

However, as India's constitutional landscape expanded, the Supreme Court recognized that a purely formal classification test was inadequate to address sophisticated or systemic forms of state discrimination.<sup>31</sup> This realization led to the evolution of equality as a dynamic, substantive concept. In *E.P. Royappa v. State of Tamil Nadu*, Justice P.N. Bhagwati famously emancipated Article 14 from its doctrinaire limits, declaring that "equality<sup>32</sup> is a dynamic concept with many facets and dimensions and it cannot be 'cribbed, cabined and confined' within traditional and doctrinaire limits." The Court established that equality and arbitrariness are "sworn enemies"—where an act is arbitrary, it is implicit that it is unequal both according to political logic and constitutional law, thereby violating Article 14.<sup>33</sup>

## B. MANIFEST ARBITRARINESS DOCTRINE

The non-arbitrariness standard reached its modern analytical peak through the formulation of the "manifest arbitrariness" doctrine. While early jurisprudence occasionally struggled to define the precise boundaries of what made a state action arbitrary, the Supreme Court provided definitive structural clarity in the landmark case of *Shayara Bano v. Union of India*.<sup>34</sup> In this case, which challenged the validity of instant triple talaq (*talaq-e-biddat*), the Court elevated manifest arbitrariness into a distinct, freestanding ground for invalidating both executive actions and legislative enactments under Article 14.

Writing for the majority, Justice Rohinton Nariman clarified that a state action is manifestly arbitrary if it is done capriciously, irrationally, or without adequate determining principles.<sup>35</sup> An act is considered manifestly arbitrary when it is excessive, disproportionate, or completely untethered from reason, logic, or public interest. Under this doctrine, the State can no longer defend a measure simply by proving the presence of an intelligible differentia; if the underlying operation or outcome of that measure is fundamentally irrational or unfair, it fails the constitutional mandate of Article 14.<sup>36</sup>

## C. AI AND ARBITRARY DECISION-MAKING

When automated systems are deployed in public governance and high-stakes private sectors, their operational profiles frequently intersect with the exact mischief prohibited by the manifest

---

<sup>31</sup> *State of West Bengal v. Anwar Ali Sarkar*, A.I.R. 1952 S.C. 75 (India).

<sup>32</sup> *E.P. Royappa v. State of Tamil Nadu*, (1974) 4 S.C.C. 3, ¶ 85 (India).

<sup>33</sup> *Id.*

<sup>34</sup> *Shayara Bano v. Union of India*, (2017) 9 S.C.C. 1 (India).

<sup>35</sup> *Id.*

<sup>36</sup> *Id.*

arbitrariness doctrine. Rather than making decisions based on tailored, individualized circumstances, Artificial Intelligence (AI) and machine learning models process human subjects through rigid, data-driven patterns. This institutionalizes systemic discrimination across critical socio-economic sectors:<sup>37</sup>

- **HIRING SYSTEMS:** Automated recruitment algorithms are frequently trained on historical resume data supplied by corporations. Because these historical datasets reflect legacy workplace disparities, the machine learning models identify gender, age, or socioeconomic proxies as implicit indicators of merit. Consequently, the system systematically downgrades female applicants or candidates from marginalized backgrounds, converting past occupational exclusion into an automated rule.
- **CREDIT SCORING:** Financial institutions increasingly utilize predictive analytics to evaluate loan eligibility and creditworthiness. These algorithmic engines ingest non-traditional data points, including geographic location, smartphone usage patterns, and community spending habits. In a deeply stratified society like India, these variables function as mathematical proxies for caste, economic vulnerability, and historical deprivation. The algorithm effectively locks marginalized communities out of the formal financial ecosystem based on structural disadvantages, rendering the denial of credit unequal and discriminatory.<sup>38</sup>
- **WELFARE DISTRIBUTION:** In public administration, the state utilizes automated decision-making systems to verify identity, assess poverty metrics, and distribute essential welfare benefits. When these algorithmic fraud-detection models flag a legitimate beneficiary as "fraudulent" or "high-risk" due to missing data points or biometric mismatches, the citizen is summarily excluded from life-sustaining entitlements. This automation replaces holistic administrative review with cold, computational estimation.<sup>39</sup>

#### **D. THE BLACK BOX PROBLEM AND OPAQUE STATE ACTION**

At the heart of these discriminatory outcomes lies the "black box problem"—the inherent opacity of advanced deep learning architectures. Because these networks optimize across

---

<sup>37</sup> See Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 *Cardozo L. Rev.* 1671, 1685–92 (2020).

<sup>38</sup> See Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 *Calif. L. Rev.* 671, 694–98 (2016).

<sup>39</sup> See Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* 174–78 (2018).

millions of shifting parameters, the precise mathematical pathway leading to a specific algorithmic output is opaque to human observers, including the data scientists who engineered the system.

This total lack of transparency fundamentally disrupts the rule of law and procedural due process. When an individual is subjected to an adverse algorithmic decision, they are denied the ability to understand *why* they were targeted, classified, or excluded. This opacity renders the fundamental principles of natural justice, such as the right to a fair hearing (*audi alteram partem*),<sup>40</sup> impossible to exercise, as a citizen cannot meaningfully challenge a rationale that cannot be articulated.

Consequently, **opaque algorithms have become a modern form of arbitrary state action.** When the state abdicates its discretionary duties to a mathematical model, it effectively insulates its decisions from constitutional scrutiny behind a shield of technical complexity. An unexplainable, biased algorithmic output is the literal definition of manifest arbitrariness: it is a decision executed without transparent determining principles, functioning as an unreviewable exercise of absolute power. To preserve the integrity of Article 14 in the digital age, technology can no longer be permitted to operate as a zone of constitutional exception; the "black box" must be subordinated to the non-arbitrariness mandates of the Constitution.

## AI, DISCRIMINATION AND ARTICLES 15 & 16

### A. DIRECT AND INDIRECT ALGORITHMIC DISCRIMINATION

Articles 15 and 16 of the Constitution of India form the bedrock of the country's anti-discrimination jurisprudence. Article 15(1) prohibits the State from discriminating against any citizen on grounds *only* of religion, race, caste, sex, place of birth, or any of them. Article 16(1) and 16(2) extend this protection specifically to matters of public employment, ensuring equal opportunity and prohibiting exclusion based on these protected characteristics.<sup>41</sup>

When applied to artificial intelligence, discrimination manifests in two distinct modalities: direct and indirect. *Direct discrimination* occurs when an algorithm explicitly utilizes a protected characteristic—such as gender, caste, or religion—to make a disadvantageous decision. While system developers rarely hardcode these categories explicitly into an optimization model, direct discrimination occurs through the use of explicit proxies. For

<sup>40</sup> See Sandra Wachter, Brent Mittelstadt & Chris Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, 41 Computer L. & Security Rev. 1, 12–15 (2021).

<sup>41</sup> INDIA CONST. arts. 15–16.

example, algorithmic credit models or hiring tools may ingest geographical data (such as specific pin codes or neighbourhoods) or surnames. In the Indian social fabric, where residential areas and surnames are deeply segregated along caste and religious lines, the algorithm's use of these variables functions as direct discrimination, intentionally or systematically filtering out individuals based on their identity.<sup>42</sup>

Conversely, *indirect discrimination* presents a more insidious challenge. It occurs when an algorithm uses an apparently neutral criterion that applies to everyone equally, but has a disproportionate, adverse impact on a protected group. Machine learning models excel at finding hidden statistical correlations across thousands of seemingly unrelated data points. For instance, an automated system might penalize employment candidates who have multi-year gaps in their resumes. While a "resume gap" appears to be a neutral metric of professional continuity, it disproportionately penalizes female applicants who take time off for maternity or caregiving duties. Similarly, algorithms that reward candidates based on the possession of high-end consumer technology or historical family zip codes indirectly institutionalize socio-economic and caste-based stratification, violating the core mandates of substantive equality.<sup>43</sup>

## **B. PUBLIC EMPLOYMENT CONCERNS AND AUTOMATED RECRUITMENT**

The integration of AI within public employment introduces severe friction with Article 16. Across both the public sector and corporate domains, organizations are deploying automated screening systems to manage large volumes of employment applications. These AI recruitment systems use natural language processing (NLP) to parse resumes, rank candidates, and predict job performance based on historical patterns of successful employees within the organization.

This reliance on historical baselines poses an immediate constitutional risk. If a government department or public sector undertaking historically favoured or hired predominantly male candidates due to historical biases, the machine learning model will identify maleness—or technical terms more frequently found on male resumes—as an objective marker of competence. Consequently, automated screening scales up structural exclusion. By delegating the initial filtration of candidates to an unreviewable mathematical formula, the state denies

---

<sup>42</sup> See Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671, 694–98 (2016).

<sup>43</sup> See Sandra Wachter, Brent Mittelstadt & Chris Russell, Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI, 41 Computer L. & Security Rev. 1, 12–15 (2021).

citizens a fair, individualized evaluation of their qualifications. This arbitrary exclusion directly undermines the guarantee of "equal opportunity" enshrined in Article 16(1).<sup>44</sup>

### **C. DISCUSSION: CAN ALGORITHMIC BIAS BE TREATED AS CONSTITUTIONAL DISCRIMINATION?**

A critical doctrinal question emerges: can algorithmic discrimination be treated as constitutional discrimination under Articles 15 and 16? Historically, the state often escaped liability under Article 15 by arguing that a measure was not based *only* on a prohibited ground, but relied on secondary criteria like efficiency, merit, or geographic distribution.

However, India's anti-discrimination jurisprudence underwent a profound shift in *Lt. Col. Nitisha v. Union of India*. In this landmark judgment, the Supreme Court explicitly recognized the doctrine of indirect discrimination within the Indian constitutional framework.<sup>45</sup> The Court held that an intention to discriminate is not necessary to establish a violation of Articles 15 and 16; rather, the judiciary must look at the *disparate impact* of a measure. If a facially neutral rule disproportionately burdens a protected class and cannot be justified by an objective, necessary purpose, it amounts to constitutional discrimination.

This structural framework is directly applicable to AI governance. Because algorithmic bias operates through systemic, disparate impacts rather than overt animus, the *Nitisha* precedent bridges the technical-legal divide. When an automated system systematically penalizes women or marginalized castes under the guise of statistical optimization, its output cannot be insulated from judicial review. Under the lens of transformative constitutionalism, the "black box" must submit to the impact test. If an algorithmic framework creates a disproportionate barrier for a protected class without a compelling, non-discriminatory justification, it constitutes a clear violation of Articles 15 and 16, transforming a technical defect into an actionable constitutional wrong.

## **PRIVACY, SURVEILLANCE, AND ARTICLE 21 IN THE AGE OF AI**

### **A. AI AND DATA COLLECTION: FUELLING ALGORITHMIC POWER**

Article 21 guarantees that no person shall be deprived of life or personal liberty except according to procedure established by law. In the digital era, personal liberty directly collides with advanced AI architectures—specifically deep learning and predictive modelling—which

---

<sup>44</sup> INDIA CONST. art. 16.

<sup>45</sup> *Lt. Col. Nitisha v. Union of India*, (2021) 15 S.C.C. 1 (India).

depend on massive data ingestion. To map trends or forecast behaviour, these engines harvest vast quantities of personal information, including transactional history, location coordinates, metadata, and biometric identifiers. This pervasive extraction blurs the line between public and private spheres, institutionalizing a system of persistent tracking that individuals cannot realistically opt out of.<sup>46</sup>

## B. PROFILING AND PREDICTIVE GOVERNANCE

This systemic data harvest enables advanced behavioural profiling. AI systems aggregate disparate data points to generate dynamic risk scores, estimating an individual's likelihood to default on a loan, violate regulations, or commit a crime. This shifts public administration from reactive enforcement to predictive governance. Instead of responding to verified legal infractions, the State increasingly relies on algorithmic inferences to pre-emptively manage crowds, deny welfare entitlements, or restrict movement. This approach substitutes individual agency with statistical probability, subjecting Article 21 liberties to automated determinations.

## C. THE RIGHT TO PRIVACY AND THE PUTT SWAMY FRAMEWORK

In *Justice K.S. Putt swamy v. Union of India*, a nine-judge Bench of the Supreme Court held that the right to privacy is an integral facet of Article 21, encompassing informational privacy, individual autonomy, and spatial privacy.<sup>47</sup> Under the *Putt swamy* doctrine, any State intrusion into privacy must clear a strict three-prong proportionality test:

- **Legality:** Backed by an explicit, valid legislative enactment.
- **Legitimate Aim:** Serving a necessary state purpose like public order or national security.
- **Proportionality:** Demonstrating a rational nexus via the least intrusive means available, balancing the infringement against the state's objective.

Advanced AI deployments routinely fail this standard. The open-ended, continuous data collection required to train machine learning models fundamentally violates the core constitutional tenets of data minimization and purpose limitation.

## D. AI SURVEILLANCE AND PREDICTIVE POLICING

---

<sup>46</sup> See Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* 63–75 (2019).

<sup>47</sup> *Justice K.S. Puttaswamy v. Union of India*, (2017) 10 S.C.C. 1 (India).

This friction is highly evident in the deployment of automated facial recognition systems (FRFS), biometric monitoring, and predictive policing software across Indian municipalities. Predictive policing models process historical crime logs to forecast future criminal hotspots. However, because these systems ingest biased historical data, they systematically direct law enforcement to concentrate patrols in historically marginalized neighbourhoods. This creates a self-fulfilling feedback loop that intensifies surveillance and enforcement actions against vulnerable communities.<sup>48</sup>

## **E. THE PREDICTIVE SHIFT: FROM OBSERVATION TO CONTROL**

Ultimately, AI transforms surveillance from observation into prediction, creating unprecedented threats to individual autonomy. Traditional surveillance was limited by human resource constraints. AI removes these physical boundaries by automating mass tracking at scale. More critically, by analysing past behaviours to forecast future choices, AI shifts the role of surveillance from watching what citizens *have done* to predicting what they *might do*. This predictive ecosystem undermines individual autonomy and the presumption of innocence. Subordinating human behaviour to algorithmic optimization compromises the core protections of Article 21, turning personal liberty into a variable managed by the State.

## **HUMAN DIGNITY AND THE RIGHT TO EXPLANATION**

### **A. DIGNITY AS A CONSTITUTIONAL VALUE**

Article 21 jurisprudence establishes that the right to life extends beyond mere animal existence to encompass absolute human dignity. In *Maneka Gandhi v. Union of India*, the Supreme Court mandated that any deprivation of liberty must be just, fair, and reasonable.<sup>49</sup> This was solidified in *Francis Coralie Mullin v. Administrator, UT of Delhi*, which declared that life<sup>50</sup> includes the right to live free from degrading practices. Under this paradigm, dignity shields individuals from dehumanizing administrative actions, ensuring the State treats citizens as ends in themselves rather than mere regulatory objects.

### **B. AUTOMATED DECISION-MAKING AND DEHUMANIZATION**

The deployment of automated decision-making (ADM) systems challenges this dignity-centric framework. By delegating administrative power to predictive algorithms, public institutions

---

<sup>48</sup> See Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* 174–78 (2018).

<sup>49</sup> *Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248 (India).

<sup>50</sup> *Francis Coralie Mullin v. Adm'r, Union Territory of Delhi*, (1981) 1 S.C.C. 608 (India).

reduce complex human lives to static data points. This computational reductionism strips away subjective, contextual realities and evaluates citizens solely against historical statistical aggregates, replacing individual identity and human agency with automated risk and eligibility metrics.

### **C. THE RIGHT TO EXPLANATION AND PROCEDURAL FAIRNESS**

To counter this automated isolation, a "right to explanation" is necessary to uphold natural justice (*lex naturalis*). Procedural fairness requires *audi alteram partem* (the right to a fair hearing) and "speaking orders"—decisions anchored in clear, reasoned explanations that permit meaningful judicial review. When opaque "black-box" models deny welfare or flag citizens without a comprehensible rationale, they violate these principles. Without knowing the underlying data points and variables driving an outcome, citizens cannot meaningfully contest decisions hidden behind proprietary code.<sup>51</sup>

### **D. CONSTITUTIONALIZING "DIGITAL DUE PROCESS"**

To bridge this regulatory gap, "digital due process" must become a constitutional requirement under Article 21. Traditional due process mechanisms are structurally ill-equipped to handle algorithmic opacity. Digital due process expands classical protections into the digital sphere, mandating that automated public systems be "fair by design." This requirement forces absolute algorithmic transparency, independent human oversight, and mandatory human intervention, ensuring technology remains subordinate to human reason and constitutional obligations.

### **THE ACCOUNTABILITY DEFICIT: WHO ANSWERS FOR ALGORITHMIC HARM?**

The rise of automated governance has created a severe responsibility problem, distributed across a complex network of developers, corporations, government agencies, and data providers. When a biased or erroneous algorithm causes tangible harm—such as wrongfully denying welfare benefits or misidentifying a suspect—the diffusion of labour across these actors creates an accountability vacuum. Because no single entity retains full control over the technical architecture or its downstream outputs, traditional tort and public law liability mechanisms struggle to assign blame, leaving victims without clear paths to legal recourse.<sup>52</sup>

---

<sup>51</sup> See Sandra Wachter & Brent Mittelstadt, A Right to Reason to an Alternative in the Western Legal Tradition: Institutionalizing the Black Box of All-Algorithms, 14 *Law, Innovation & Tech.* 211, 215–18 (2022).

<sup>52</sup> Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 *Harv. J.L. & Tech.* 353, 356–59 (2016).

This structural failure introduces profound public law challenges, paralyzing administrative accountability and neutralizing constitutional remedies. Under established principles of administrative law, public authorities cannot unlawfully delegate their statutory, discretionary powers to third parties, let alone to opaque mathematical equations.<sup>53</sup>

Consequently, the core question of modern public law emerges: **Can the State evade constitutional responsibility by delegating decisions to algorithms?** Under the doctrine of state action and constitutional morality, the answer must be a definitive no. The state cannot use technological complexity to insulate itself from judicial review. When an automated system acts as a functional proxy for public administration, its outputs remain fully attributable to the state, requiring strict compliance with fundamental rights.<sup>54</sup>

## **INDIA'S EXISTING AI GOVERNANCE FRAMEWORK: PROGRESS WITHOUT ACCOUNTABILITY?**

### **A. THE LEGISLATIVE DEFICIT: IT ACT AND DPDP ACT**

India's current statutory regime handles advanced Artificial Intelligence (AI) through a fragmented approach. The aging Information Technology Act, 2000 serves as the primary legislation for cyber activities, yet it features structural limitations. Written long before the emergence of deep learning, it lacks the technical vocabulary or legal mechanisms needed to regulate algorithmic opacity, automated decision-making, or automated discrimination.<sup>55</sup>

Conversely, the Digital Personal Data Protection (DPDP) Act, 2023—supplemented by the DPDP Rules, 2025—introduces a strict, consent-centric framework for processing data. While its strengths lie in robust data-minimization mandates and penalties reaching up to ₹250 crore for breaches, its weaknesses regarding AI are pronounced. The Act focuses on data handling rather than algorithmic execution. Consequently, it fails to address downstream algorithmic harms, automated profiling, or the inherent biases generated inside machine learning models after data ingestion.<sup>56</sup>

### **B. POLICY ASPIRATIONS: NITI AAYOG AND INDIAAI MISSION**

In the absence of dedicated statutory laws, India's approach relies on policy initiatives. NITI Aayog has established foundational policy through its *Responsible AI Principles*, advocating

<sup>53</sup> *A.K. Kraipak v. Union of India*, (1969) 2 S.C.C. 262 (India).

<sup>54</sup> *Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248 (India).

<sup>55</sup> Information Technology Act, No. 21 of 2000, Acts of Parliament, 2000 (India).

<sup>56</sup> Digital Personal Data Protection Act, No. 22 of 2023, Acts of Parliament, 2023 (India).

for non-discrimination, transparency, and accountability in automated systems. Similarly, the government's flagship *IndiaAI Mission* accelerates infrastructure development, utilizing a massive budget allocation to deploy over 38,000 GPUs to democratize compute access.<sup>57</sup>

However, a critical evaluation reveals that this ecosystem prioritizes economic and technological progress without creating enforceable accountability. NITI Aayog's principles operate purely as voluntary ethical guidelines, lacking enforcement powers. While the IndiaAI Mission builds sovereign compute capacity, it leaves deployment parameters largely unchecked. India has developed clear policy aspirations and technical infrastructure, but it lacks binding, statutory accountability mechanisms, creating an environment where algorithmic harms can occur without clear legal recourse.<sup>58</sup>

## COMPARATIVE PERSPECTIVES: GLOBAL MODELS OF AI REGULATION

### A. GLOBAL REGULATORY ARCHETYPES

The international landscape of Artificial Intelligence (AI) regulation reflects distinct geopolitical priorities, divided into four primary regulatory models:

- **THE EUROPEAN UNION (RISK-BASED MODEL):** Anchored by the groundbreaking EU AI Act, the Union utilizes a horizontal, risk-based classification system. It segregates AI applications into clear tiers (unacceptable, high, limited, and minimal risk). High-risk systems—such as automated screening or biometric evaluation—must meet strict transparency obligations, maintain precise logging, and implement mandatory human oversight frameworks before accessing the market.<sup>59</sup>
- **THE UNITED STATES (SECTOR-SPECIFIC MODEL):** Rejecting centralized legislation, the US relies on a decentralized, sector-specific approach. Regulatory bodies, such as the Federal Trade Commission (FTC), utilize existing consumer protection and anti-discrimination laws to penalize deceptive algorithmic practices, guided by high-level White House Executive Orders aimed at balancing safety with market dominance.
- **THE UNITED KINGDOM (PRINCIPLES-BASED MODEL):** The UK maintains a decentralized, principles-based framework. Rather than introducing new statutory

<sup>57</sup> NITI Aayog, *Responsible AI: Approach Document* (2021).

<sup>58</sup> IndiaAI Mission, Ministry of Electronics & Information Technology, Government of India (2024).

<sup>59</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689).

bodies, it empowers existing regulators to interpret core principles—such as fairness and safety—within their respective fields to foster innovation.

- **CHINA (STATE-CENTRIC MODEL):** China enforces a highly targeted, state-centric model. Through specific regulations governing recommendation algorithms, deepfakes, and generative AI architectures, the state mandates strict security assessments and algorithm registries to ensure that technology aligns with state security and social control.

## **B. THE CASE FOR AN INDIAN CONSTITUTIONAL MODEL**

A comparative assessment reveals why India cannot simply copy these international frameworks. Purely economic or market-driven approaches, like those of the US or UK, treat algorithmic harms as transactional market failures, while security-centric models, like China's, prioritize state stability over individual protections.

In a deeply pluralistic and historically stratified society like India, algorithmic bias represents a profound civil rights issue rather than a minor technical glitch. When algorithms automate discrimination across caste, gender, or religious lines, they threaten fundamental rights. Therefore, India requires a native **constitutional model of AI governance**. This framework must be rooted directly in constitutional morality and the transformative guarantees of Articles 14, 15, and 21, ensuring that technological progress remains tethered to human dignity, equal protection, and democratic accountability.<sup>60</sup>

## **AN INDIAN AI ACCOUNTABILITY FRAMEWORK**

To prevent automated systems from operating in a legal vacuum, India requires a comprehensive statutory structure anchored in its unique constitutional jurisprudence. Integrating public administration and digital rights, this framework codifies seven operational pillars.

## **THE SEVEN PILLARS OF GOVERNANCE**

- **PILLAR 1: CONSTITUTIONAL IMPACT ASSESSMENT:** Prior to deployment, developers must conduct formal assessments ensuring compliance with Article 14 (Equality review against encoded socio-economic disparities) and Article 21 (Privacy and Dignity reviews to protect personal liberty).

---

<sup>60</sup> INDIA CONST. arts. 14, 15, 21.

- **PILLAR 2: MANDATORY ALGORITHMIC AUDITS:** Enforcement of periodic, independent technical reviews of deployed systems to catch historical bias replication and data drift errors.
- **PILLAR 3: EXPLAINABLE AI:** Guaranteeing citizens a "Right to Know" and a "Right to Explanation" when automated decisions impact their legal status, securing procedural due process.
- **PILLAR 4: HUMAN-IN-THE-LOOP GOVERNANCE:** Restricting full automation by mandating definitive human oversight for high-stakes welfare allocation, employment profiling, and criminal justice systems.
- **PILLAR 5: ALGORITHMIC TRANSPARENCY REGISTRY:** A centralized public ledger mandating disclosure of dataset origins, system accuracy, and proactive risk assessments.
- **PILLAR 6: AI OMBUDSMAN:** An independent, tech-neutral quasi-judicial regulator equipped with centralized complaint and dispute resolution mechanisms.
- **PILLAR 7: CONSTITUTIONAL TORT LIABILITY:** A strict liability regime enabling direct civil compensation for algorithmic harms, strictly limiting standard state defences like sovereign immunity.

### THE DOCTRINE OF "CONSTITUTIONAL ALGORITHMIC ACCOUNTABILITY"

This framework establishes the new legal doctrine of **Constitutional Algorithmic Accountability**. Rather than viewing algorithms as mere neutral business tools or inscrutable "black boxes," this doctrine establishes that whenever an algorithm exercises public functionality, acts as a conduit for state power, or fundamentally shapes civic rights, it must be bound by the same constitutional limits as the State itself.<sup>61</sup>

By grounding technical parameters directly within fundamental rights, the doctrine bridges the gap between digital automated systems and the rule of law. It explicitly shifts the legal burden of proof onto developers and deployers, demanding that they proactively demonstrate their systems do not discriminate or dehumanize citizens.

---

<sup>61</sup> Sandra Wachter & Brent Mittelstadt, *A Right to Reason to an Alternative in the Western Legal Tradition: Institutionalizing the Black Box of Algorithms*, 14 **Law, Innovation & Tech.** 211 (2022).

## **CONCLUSION: CONSTITUTIONALISING ARTIFICIAL INTELLIGENCE IN INDIA**

The widespread adoption of artificial intelligence across public administration and corporate enterprise makes it clear that algorithms are not neutral, objective tools. Instead, algorithmic bias introduces severe constitutional concerns, transforming encoded socio-technical disparities into automated, systemic discrimination. India's current legal framework remains fragmented; existing statutes like the IT Act and the DPDP Act focus primarily on data protection and cybersecurity, leaving profound downstream algorithmic harms, black-box opacity, and structural accountability gaps unaddressed. Therefore, India must reject a purely economic or market-driven regulatory approach. The nation should not merely regulate AI as a passing technological innovation, but explicitly govern it as a constitutional institution capable of deeply impacting fundamental rights.<sup>62</sup>

The future challenge before India is not whether artificial intelligence can make faster or more efficient decisions than human bureaucrats. The real challenge is whether those automated decisions can remain strictly consistent with the constitutional values of equality, dignity, liberty, and the rule of law. As machine learning models increasingly dictate access to socio-economic opportunities, welfare distribution, public employment, and criminal justice, the core doctrine of constitutional morality must become the foundational normative framework for AI governance. Ensuring that automated infrastructures submit to absolute transparency, digital due process, and strict judicial review is essential to protect marginalized communities from self-replicating prejudice. Ultimately, an accountable algorithm is not simply a technological necessity for the digital economy; it is a fundamental constitutional imperative required to preserve the democratic integrity of the Indian Republic.

---

<sup>62</sup> *Justice K.S. Puttaswamy v. Union of India*, (2017) 10 S.C.C. 1 (India).