Journal of Legal Subjects
Vol: 05, No. 02, July-Dec 2025
ISSN: 2815-097X, DOI: 10.55529/jls.52.1.12

Research Paper

Artificial intelligence governance in India: a critical legal framework for accountability, transparency, and regulatory compliance

Mukund Kumar Pandey*
*Ph.D. Research Scholar, Faculty of Law, Kalinga University, Raipur, Chhattisgarh, India.

Article Info

Article History:
Received: 02 March 2025
Revised: 13 May 2025
Accepted: 19 May 2025
Published: 05 July 2025

Keywords:
Artificial Intelligence
Governance
Algorithmic Accountability
Digital Personal Data Protection Act
Niti Aayog
Regulatory Compliance

ABSTRACT

In India, Artificial Intelligence (AI) is now making a significant presence in the fields of governance, finance, healthcare, policing and judicial administration. However, India does not have a specific AI legislation, but it does have existing legislations like the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023 and non-binding policy guidance. The paper examines the emerging governance framework of AI in India, discusses the constitutional principles, legislative initiatives and soft-law frameworks such as the European Union's AI Act, the OECD's Principles of AI and the sectoral approach in the United States. There are significant regulatory gaps identified in the study based on four analytical criteria: accountability, transparency, enforceability, and remedies to algorithmic harms, such as the lack of a dedicated AI authority, mandatory algorithmic impact assessments, a robust explainability framework and effective remedies to harms. The paper demonstrates through examples of face recognition and automated decision-making that it is not sufficient to have voluntary governance based on principles that lead to the protection of fundamental rights for high-risk AI systems. It proposes a five-layer governance framework, comprising constitutional protection, statutory measures, soft law, sectoral oversight and a special accountability mechanism, such as the creation of a statutory AI authority and compulsory audits. Concluding the paper, it is suggested that the pro-innovation stance adopted by the Indian government fosters technological progress, but it is necessary to have more robust and legally enforceable protection mechanisms to safeguard the constitutional values of equality, privacy and due process in the era of Artificial Intelligence.

Corresponding Author:
Mukund Kumar Pandey
Ph.D. Research Scholar, Faculty of Law, Kalinga University, Raipur, Chhattisgarh, India.
Email: maddy.muku2@gmail.com

Journal of Legal Subjects (JLS)

Copyright © 2025 The Author(s). This is an open access article distributed under the Creative Commons Attribution License, (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Previously human decisions are being made by Artificial Intelligence systems from creditworthiness assessment, insurance underwriting and criminal risk scoring and welfare eligibility to recruitment screening and even preliminary judicial research. India is one of the largest digital-public-infrastructure ecosystems in the world with a national policy for adoption of DPI and increasing number of services run on AI means India is a national user and a mass consumer of AI. From digital identity to agriculture advisory services, government initiatives are increasingly incorporating machine learning capabilities, and private platforms have been using recommendation engines, biometric authentication and generative AI tools that impact hundreds of millions of users.

Compared to a corresponding development in the legal structure that regulates it, the speed of the diffusion of AI has been significantly faster. There has been an extensive EU Artificial Intelligence Act [1] with a risk-based classification approach, but India still does not have a comprehensive legislation on AI. The Information Technology Act, 2000 [2] continues to be the main legislation governing electronic transactions and cybercrime, but that legislation was not intended to control cyber autonomous decision making systems. The Digital Personal Data Protection Act, 2023 [3] is the nation's first comprehensive data protection act, but does not directly regulate the operation, output or risk classification of an AI system; it only limits the use of personal data that is commonly used in AI systems. Normative context for assessing the harms of AI is provided by the constitutional guarantees of equality (Article 14), freedom of expression (Article 19) and the right to privacy as recognised as part of Article 21 [4] in Justice K.S. Puttaswamy v. Union of India [5] but as background principles.

The Government of India have been playing with soft-law measures without any binding rules. The National Strategy for Artificial Intelligence [6] had mentioned about economic growth and social inclusion in its vision of 'AI for All' and the two parts of the Approach Document [7], [8] spoke about seven 'Principles for Responsible AI', which include (a) Safety, (b) Equality, (c) Inclusivity, (d) Privacy, (e) Transparency, (f) Accountability, and (g) The protection of positive human values. On this, the Ministry of Electronics and Information Technology (MEITY) has issued periodic advisories to intermediaries regarding the generation of AI content [9] and recently introduced India's AI Governance Guidelines in November 2025 [10] which is a risk-based and balanced approach, but not an AI stand-alone law. While these instruments have become numerous, none are specific enough, enforceable enough, or as detailed as they need to be to protect rights, India's approach to AI is also driven by political economy factors, such as the goal of becoming a global AI hub [11], [12].

In this paper, three related research questions are asked. What are the current laws and regulations that cover the use of AI in India, and how do they align with the need for accountability, transparency, and redress in high-risk AI systems? Second, what are the advantages and disadvantages of India's approach as compared to EU's risk based statutory approach, liability based sectoral approach in the USA and soft law approach in OECD? Third, what would be India's balancing act between the constitutional ideals of equality, privacy and due process and the aspiration of an AI-powered economy?

This regulatory vacuum is anything but hypothetical or theoretical. Biometric identity systems are used to rely on a person's eligibility for welfare schemes, even though its usage is not explicitly authorised by law in India, credit-scoring models have become more and more machine-learning based, and law-enforcement is used in several States with facial recognition systems with no specific authorising law.In India, the concept of automated eligibility verification through biometric identity systems, increasing machine-learning based credit-scoring models, and law-enforcement using facial recognition systems without the presence of any specific authorising law are increasing trends in the country's welfare architecture. Each deployment poses the same issue as to who is liable for the errors, biases and

infringement upon rights of an automated outcome and how such victim is to be able to get redress, but different portions of the regulatory morass set out above currently fall on these deployments [13]. This cut-and-paste approach itself is a failure of regulation, involving the setting of uneven standards of accountability to regulate a similar function across different sectors and harmonisation by courts to settle individual disputes well after the damage has been caused and not a simple and comprehensive design of the regulatory ex-ante.

With its scale, this need is magnified for India. India's utilization of AI systems rests on payments infrastructure (UPI) and identity systems which, at the population level of citizens, are already influential in shaping rights, the degree of bias of which has a parallel impact on the rights of citizens. A “cheap” regulatory policy is therefore not equal: it is redistributive, and costs the individual citizen, who is the least likely to be able to point it, document it or take ordinary civil action against it.

The paper has the following organization. In Section 2, the existing academic and policy literature on the general governance of AI as well as on the general and specific policy trajectory of India is reviewed, which is broken down into three subsections; India's policy trajectory, accountability and transparency theory, and comparative regulator trajectories in AI. In Section 3, it's revealed the doctrinal and comparative methodology used in this study, the primary sources and secondary sources material sources used in this study and the four criteria used in the analysis all throughout the paper. The results of the comparative and gap analysis are summarized in two comparative tables in section 4, followed by a discussion of gaps in accountability and explainability, a proposed 5-layer governance structure (Figure 1) and a discussion on the practical implications of the compliance of the gap analysis. The paper concludes with the policy recommendations and further research opportunities in Section 5.

2. RELATED WORK

2.1 India's Policy Trajectory on AI Governance

While varied, all the studies of the Indian trajectory of AI governance have a common theme: The lack of a rights-protective logic, as compared to an industrial policy logic, in the governance of AI in India. By focussing on the 'data state agenda', Joshi's [11] work helps to situate the policy interventions in place in India in a larger political economy that is seeking to afford open public digital infrastructure and broad access to this set of data, not only in order to minimise individual risks but to enable the rapid production of AI. Overall, writing in the Carnegie Endowment, Mohanty and Sahu observe that India's regulatory journey has been fragmentary, beginning with a “hands-off” approach to the regulation and gradually moving towards a more structured, but still non-prescriptive, and risk-based approach as seen in the recent number of advisories by the Ministry of Electronics and Information Technology (MeitY), and the 2025 AI Governance Guidelines [12].

This body of work is supplemented with doctrinal commentary on the Digital Personal Data Protection Act, 2023, which is not a specific legislation related to AI, but also the closest in India to an "accountability" measure concerning algorithmic aspects since most of the AI systems operate with or on personal data. [3], [14] situate the DPDPA in the collection of the world's data protection laws and contend that the architecture of DPDPA is akin to the General Data Protection Regulation (GDPR) of the European Union, but has been weaker with regard to the architecture of independent surveillance and oversight. In a doctrinal analysis [15] recognizes that there are gaps in the consent, exemption and cross-border transfer provisions in the DPDPA, documents which seem remarkably similar to those of both the EU General Data Protection Regulation and the (then proposed) American Privacy Rights Act, and offers 'rules of thumb' and a strategy of interpretation to improve its practical application.

The one thing that they have in common is that the regulatory sequencing of India has been different to that of the EU. India is a long laggard behind the EU which had a data protection statute covering the general data and followed it with an AI-specific statute, but India did not have an AI-specific statute and had its comprehensive data protection statute just adopted in 2023 when across areas of the economy, AI is already a big part of everyday business. One area that the commentators point to as problematic is the absence of a sequencing in the DPDPA, which fails to sufficiently consider the risk of algorithmic outputs,

algorithmic behaviour and systemic risk. This observation is important as an assumption in the gap analysis that takes place later in this paper.

2.2 Accountability and Transparency in AI Governance Theory

Another body of literature focuses on the conceptual framework of the accountability of AI, without relying on a specific jurisdiction. In a recent paper, [16] maintain that the concept of accountability in AI is often used but poorly understood and suggest it should be understood as "answerability", which includes the concepts of authority recognition, interrogation and limitation of power, with seven features of answerability being operationalised: context, range, agent, forum, standards, process, and implications. The downside of the framework is that accountability is not a dichotomy, it does not have to be an existing structure or standard that has gained traction in a jurisdiction, so that even if a principle of accountability is adopted that doesn't imply that the forums and/or processes in which that principle can be acted upon are.

Tied up closely with transparency is another value, discussed also by [17], which they refer to as the "transparency ideal"—that is, their argument goes, even if we open up an algorithm's internal workings or code, that alone does not necessarily give rise to accountability, as opacity depends as much on the complexity of the infrastructure and organisation as it does on a single "hidden variable". This influential taxonomy [18] distinguishes three kinds of opacity: intentional secrecy, technical illiteracy, an opacity that is inherent to the specific machine-learning methods and scale at which they are applied, which gain immediate significance in the design of an explainability mandate, as a means to direct the required regulatory measures. [19] ground-breaking survey on 'Explainable Artificial Intelligence' (XAI) maps the technical approaches for lowering the opacity of models, offering a technical language increasingly used in the enactment of laws and the making of legal obligations on explainability.

On a lower level, national AI ethics documents around the world have followed the call of [20] for an AI4People model, and in turn the five principles model put forward [21] (beneficence, non-maleficence, autonomy, justice, explicability), which coincide closely with NITI Aayog's principles. The overall impression gleaned from this literature is that the vocabulary of concepts around accountability, transparency and explicability has been well developed; the ongoing difficulty, as this paper suggests, is to move from the realm of concepts to the obligations of the law [22].

2.3 Comparative Regulatory Models

The most advanced statutory framework today is that provided for the European Union's Artificial Intelligence Act, Regulation (EU) 2024/1689 which classifies AI systems into three tiers of risk based on their ability to affect fundamental rights and give rise to specific risks, subjecting "high risk" systems to ex-ante conformity assessment, documentation and human oversight requirements, banned AI practices, and only partially regulated "minimal risk" ones. Another benchmark, which is more principle-based and softer, and that has influenced, as a reference framework, measures adopted by countries has been OECD/LEGAL/0449 Stakeholder Handbook on Artificial Intelligence published in 2022. As observed by [23] comparing international AI regulation: where different jurisdictions are aiming to fall, social and legal landscapes tend to align themselves on a spectrum from the EU pushing for rights-based, ex-ante approach, to the U.S. leaning towards a liability-based model of ex-post protection. The Asia regulatory space has seemed to be located far more towards the latter extreme of the spectrum, although all Asian nations are pushing rhetorically towards rights protection.

Lastly, the ways in which the use of high-risk artificial intelligence functions have been subject of greater Indian scholarly scrutiny on a comparative constitutional level. Mustafa and Leo's study of the National Automated Facial Recognition System, on the other hand, focuses on the proportionality doctrine enunciated in Puttaswamy to examine the constitutionality of the use of live facial recognition in the deployment of policing devices and found that there is no specific statute in the parliament to authorize the deployment of live facial recognition and thus its constitutional viability was questionable. As shown in this case-specific literature, the question of accountability and transparency is of practical concern on a larger

scale as part of the debate, and has an important real-world footing to support the doctrinal examination analysed at length in this paper.

Overall, the literature reviewed in this section reveals a threefold gap that the present contribution addresses. First, the DPDPA and NITI Aayog's principles have so far been analysed separately, with few studies comparing them against a common set of accountability criteria. Second, although a conceptual framework exists, little research applies it to India's institutional structure—identifying which "forum," in Novelli, [16] terms, already exists or is missing. Third, the comparative surveys by [23], which place India within a global taxonomy, do not go as far as suggesting architectures for reform, as this paper does in Section 4. The present paper therefore aims at providing a structured analysis and a criterion-based synthesis, which can bring together India-specific doctrinal analysis, general's accountabilities theory and comparability of regulatory design in one analysis.

3. METHODOLOGY

This study adopts a doctrinal research method combined with structured comparative law analysis. The key legal resources currently used to regulate AI in India were identified, interpreted and systematized through doctrinal analysis, including the Constitution of India (Articles 14, 19, 21); the Information Technology Act, 2000 ; the Digital Personal Data Protection Act, 2023; and the following advisories issued by NITI Aayog's National Strategy for Artificial Intelligence and its two-part Approach Document on Principles for Responsible AI of 2021, advisories issued by MeitY on due diligence for usage of AI generated content the India AI Governance Guidelines issued in November 2025. The following primary judicial sources were examined to calculate parameters for the constitutional benchmark for testing AI-powered state action: Justice K.S. Puttaswamy v. Union of Indiaand Shreya Singhal v. Union of India [24] which deals with the principles of legality, legitimate aim, and proportionality.

Secondary sources included peer-reviewed literature on the theory of AI accountability, literature on transparency/explainability, and literature on comparative regulation of AI, from law and policy journals, as well institutional reports from The Carnegie Endowment for International Peaceand the OECD [22]. Four criteria were used to develop an analytical framework for each AI instrument to consider: (i) accountability: if an AI system or the development of an AI system results in a decision regarding a right, interest or life of an individual, who may be held responsible for the action of the system and to before whom this responsibility can be taken up; (ii) transparency: the disclosure of the logic, data used for training or the criteria which inform decision making of an AI system; (iii) a system of enforcement: if an AI system or development of such an AI system leads to a decision regarding a right, interest or life of an individual, concerns about the accountability of the relevant party and the existence of a jurisdiction before which the party is accountable; (iv) the means of redress: existence of a channel for challenging a negative decision in the case of the impact of an AI system on an individual's rights, interests or life.

It was then followed by a step of comparison with three models, namely EU's risk-tiered statutory approach... NIST AI Risk Management Framework (as representative EU soft law regulation) and US model... (Through Colorado AI Act, a state law and NIST's model respectively as representative state, federal laws)." China's licensing-based approach was contrasted with the three models. This comparative exercise is listed below in Table 1. The 4 criterion analytical framework was then applied to each instrument in India in particular to get the gap analysis as presented in Table 2. Finally, the findings of both analyses were combined, resulting in a proposed governance structure (see Figure 1) that merges constitutional, statutory, soft-law, sectorial and new proposed levels of accountability. This study, of a qualitative and doctrinal nature, does not make any specific factual or quantitative statements about the levels of compliance or enforcement effectiveness; it is not a prediction nor a statistical claim, but an analysis of a policy debate that is ongoing [24].

4. RESULTS AND DISCUSSION

4.1 Comparative Regulatory Landscape

Table 1 is a comparison of the India with the European Union, the United States, the OECD and China based on the four analytical criteria in this study.

Table 1. Comparative Analysis of AI Governance and Regulatory Frameworks across Jurisdictions

Jurisdiction	Governing Instrument	Regulatory Approach	Enforceability	Redress Mechanism
India	IT Act 2000; DPDPA 2023; NITI Aayog Guidelines; MeitY Advisories	Advisory, sectoral, self-regulatory	Largely non-binding for AI-specific harms	Fragmented; Data Protection Board only for data claims
European Union	Regulation (EU) 2024/1689 (AI Act)	Risk-tiered, statutory, ex-ante	Binding with tiered penalties	National market surveillance authorities; judicial review
United States	Sector-specific guidance; NIST AI RMF; state statutes (e.g., Colorado AI Act)	Ex-post, liability-driven, fragmented federalism	Binding at state level; voluntary federally	Tort litigation; state attorney-general enforcement
OECD (Principles)	OECD/LEGAL/0449 Recommendation	Soft-law, principle-based	Non-binding	None; relies on member implementation
China	Algorithmic recommendation & generative AI regulations	Ex-ante licensing and content control	Binding, state-enforced	Administrative sanction; content takedown

The framework in use in India is situated in the middle ground, having only a statutory data protection instrument (DPDPA), but no obligations related to the various tiers of risk as in the EU AI Act, and a relatively minimal enforcement framework which is in contrast to the US model, which is fragmented and tort-oriented, which at least provides a mechanism for redress after the fact through tort action. Despite India having a much larger population of citizens that are exposed to algorithmic harms and the potentially far-reaching consequences these can have on public welfare, India's main AI-specific tools such as NITI Aayog principles, [7] and policies, [8] continue to maintain their advisory nature, distinguishing its approach from the European one most significantly, and aligning it more closely with the OECD's soft-law posture [22].

It is telling that the relatively splintered United States model this paper does not present as a model to follow at least offers two features that are missing from India's framework; namely, a litigable body of general tort law, and, in specific states such as Colorado, obligatory tiers of responsibility for AI developers and deployers based on the type and severity of the "high-risk" AI system. The doctrine of torts in India, in its general principles, could accommodate in principle claims of 'algorithmic negligence' or 'deficient service'; but from a precedent perspective, there is a relative lack of development in the area of AI claims, because no such evidence is mandated as evidence – and the lack of documentation or disclosure requirements makes it extremely difficult at the threshold evidentiary level to find out how an automated decision was arrived at. This clearly north-south imbalance between the parties impacted and those developing AI is already a regulatory challenge that a disclosure and audit-trail could help solve, as it is in the EU AI Act's [1] high-risk provisions.

4.2 Gap Analysis of India's Domestic Instruments

Table 2 breaks down the same four criteria for the instruments used in India to take a closer look and sees that out of the four criteria, there is no instrument in existence that meets all four.

Table 2. Comparative Assessment of AI Governance Mechanisms in India

Governance Criterion	IT Act, 2000	DPDPA, 2023	NITI Aayog Principles	MeitY Advisories / Guidelines
Statutory force	Yes (partial; intermediary rules)	Yes	No (policy only)	No (executive advisory)
AI-specific scope	No	Indirect (personal data only)	Yes	Yes
Accountability of deployers	Limited	Limited to data fiduciaries	Aspirational	Conditional, self-declared
Transparency / explainability mandate	Absent	Absent	Recommended	Recommended, not enforced
Independent oversight body	CERT-In (cyber incidents only)	Data Protection Board (data only)	None	None
Individual redress for algorithmic harm	Indirect (civil/criminal law)	Limited to data grievances	None	None

An attempt to cover the substantive fairness of the output from an AI system is made in Table 2 through the Information Technology Act, 2000 [2] itself as well as the Indian Computer Emergency Response Team established under that act. The Indian Computer Emergency Response Team is a statutory entity without jurisdiction over substantive fairness or accuracy of an AI system's output however it does have jurisdiction over cybersecurity incidents and intermediary due diligence. Even though it extends statutory force [3] and introduces an adjudicatory body in the form of the Data Protection Board of India, the Board's jurisdiction is limited to cases involving personal data: therefore, an AI system that generates discriminatory results with anonymous or aggregated or synthetic data is completely excluded from the Board's powers, in a situation which is increasingly common in credit scoring and insurance underwriting. The Principles of EWLAI [7], [8] issued by NITI Aayog and the advisories issued by the Ministry of Education and Technology [9], [10] are more focused on AI and talk of the fundamental principles of transparency and accountability as 'normative commitments' without any binding or enforcement powers and also lack an independent forum which an aggrieved party can raise its grievance before.

This is no gap of heads in the clouds. The states that deployed on NCRP have no specific law authorizing the establishment of the system, relying only on the generic policing powers, and executive notifications [13]. Applying the principle of proportionality (as elaborated in Puttaswamy [5] legality, legitimate aim, necessity, and proportionality stricto sensu) to such deployments brings to fore the need for an accessible and predictable application of the law; to that end, it is hard to see how a scheme that relies on executive notifications would meet the 'legality' requirement when concerned with biometric monitoring of the general public. Indian legislation does not currently distinguish between 'permissible' and ‘prohibited' use scenarios, and between ‘prohibited' and ‘narrowly permissible' (as in the European Union, where treatment of the scanning of fingerprints for remote verification is separately and more strictly defined) independent authorisation requirements for various use cases.

4.3 Accountability and Explainability Deficits

In India, the stand taken by Novelli, Taddeo, and Floridi [16] on answerability is reflected in only some of its attributes, notably ‘context' and ‘standards', with very few concrete policies addressing the

remaining attributes, particularly ‘forum' and ‘process'. There is no current statutory mechanism other than the general judicial system where a developer or deployer of a high-risk AI system will owe the invitation to explain a biased or wrong response from the AI system. Having a technical disclosure obligation without a body to listen, assess and respond to what they hear in a disclosure will have little accountability effects in practice, as Ananny and Crawford [17] describe.

The explainability deficit only exacerbates this issue. There are no current instruments in India that require the sort of model documentation, algorithmic impact assessment, right to meaningful explanation which would enable an affected individual to understand why an automated system denied them a loan, classified them as high-risk or excluded them from a welfare scheme. However, the lack of an explainability mandate in Indian policy may not be a technical limitation, as Arrieta and colleagues' [19] taxonomy of explainability techniques suggests meaningful explanation is technically possible even with complex models. The India AI Governance Guidelines of November 2025 [10] suggest a risk-based approach that in theory might allow for increased explanatory requirements for more critical use cases, but in their current form, are still interpretive and non-binding; which leaves room for individual ministries and regulators to interpret them differently as they implement.

4.4 Toward a Five-Layer Governance Architecture

Combining the comparative and gap analyses above, this paper proposes a five-layer AI governance framework, shown in Figure 1.

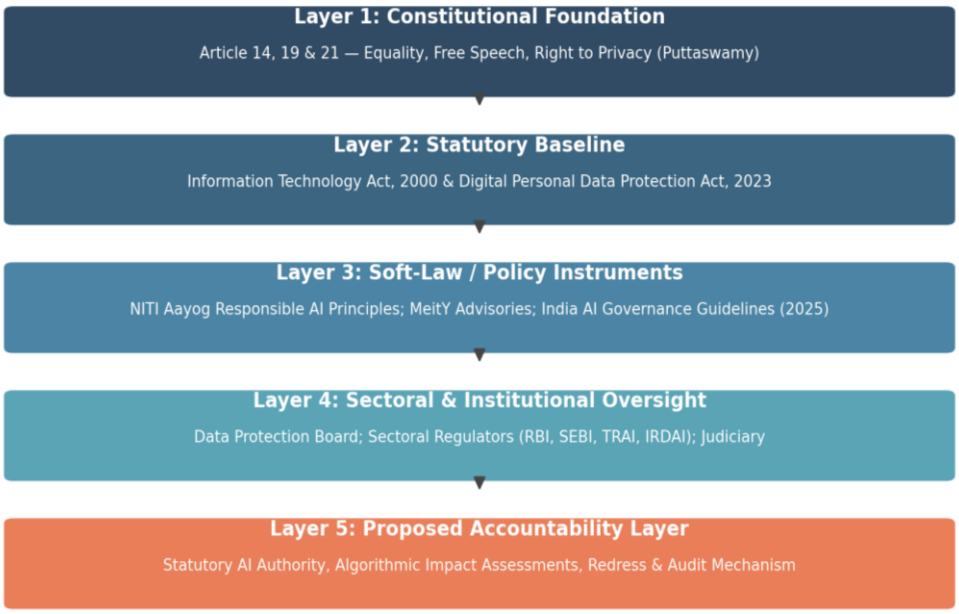


Figure 1. Proposed Five-Layer AI Governance Architecture for India

The initial two 'layers' of the proposed architecture, constitutional foundations [4] and the statutory baseline (IT Act [2] and DPDPA [3]) are already in place and do not require any additional legislation other than the consistent interpretation of the law by the judiciary. The third layer consists of NITI Aayog's principles [6], [7] and MeitY's advisories which exist already and need not undergo any structural change other than further fine tuning of the policies. The fourth and fifth layers are the key layers in the proposed critical intervention that is put forward in this paper. The fourth layer would institutionalise the supervisory role that is currently informal, where sectoral regulators like the Reserve Bank of India, the Securities and Exchange Board of India, the Telecom Regulatory Authority of India and the Insurance Regulatory and Development Authority of India provide their own specific guidance on AI in their respective sectors, with a common taxonomy of AI risk categories. The fifth and the proposed new layer would create a dedicated statutory body having the power to declare high risk AI systems, carry out Impact Assessment of algorithms before deployment and audit such algorithms after deployment, and

resolve individual complaints related to automated decisions, which currently don't exist in any one AI-specific Indian institution.

This proposed architecture has been designed intentionally to not endorse wholesale replacement of the existing instruments in India. Instead, it offers the DPDPA [3], the IT Act and the NITI Aayog principles [7] and as building blocks of an overall architecture, reflecting the literature's finding that most jurisdictions construct a regime of AI governance in layers on top of existing data protection and consumer protection laws. Furthermore, the proposed fifth layer here implements the risk-based approach to AI governance preferred by the Indian state, which is likewise echoed in the Indian guidelines for AI governance 2025 (which are merely a restatement of India's stated preferences). The fifth layer here brings a binding element to the risk-based approach to AI governance that the Indian policy already in principle recognises: audit, assessment and redress obligations [25].

4.5 Policy Implications for Regulatory Compliance

In practice, India's current compliance requirements should be treated as a floor, not a ceiling, for AI developers and deployers. Meeting the DPDPA's consent, purpose-limitation, and data-minimisation requirements only addresses the data-input dimension, not the fairness, accuracy, or explicability of an AI system's outputs. Organisations that embrace algorithmic impact assessments, document their models in line with the cataloguing framework of Arrieta and colleagues and have internal grievance-redress procedures for decisions influenced by AI are more likely to be better equipped to reduce the reputational and litigation risks of algorithmic decision-making, and to prepare for the binding obligations that this paper suggests will be a future phase of India's regulatory journey in the years to come, as the 2025 Guidelines [10] are sectoralised by rule-making.

5. CONCLUSION

India's current AI governance regime mixes robust constitutional safeguards, data protection regulations, cybersecurity measures, and non-obligatory ethical principles, but lacks a specific legal instrument regulating high-risk AI systems in fields such as employment, finance, welfare and law enforcement. This analysis reveals that India's approach is more aligned with OECD principles than the EU's legally binding, risk-based AI governance framework, using doctrinal and comparative methods. The paper calls for a five-layer governance structure, such as a statutory AI authority, compulsory algorithmic impact assessments for high-risk AI systems, and an efficient grievance redress mechanism, among others, to address these gaps. These steps are designed to complement, not impede, India's ongoing innovation-based initiative in the field of AI, ensuring accountability, transparency, and public trust. The study also recognizes that the doctrinal approach adopted in the study does not look at the outcomes of enforcement, and India's AI Governance Guidelines are yet to be fully implemented. Future research should empirically evaluate how these guidelines are applied by sectoral regulators, explore experiences of implementing the EU AI Act, and explore current methods of redress against harmful automated decisions by citizens. Building accountability and protecting constitutional principles of equality, privacy, and due process will play a key role in India's future aspirations to become a trusted global leader in AI.

Acknowledgments

The authors have no specific acknowledgments to make for this research.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mukund Kumar Pandey	✓	✓	✓	✓		✓		✓	✓	✓	✓		✓	✓

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Informed Consent

All participants were informed about the purpose of the study and their voluntary consent was obtained prior to data collection.

Ethical Approval

Not applicable.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] D. Joshi, "AI governance in India - law, policy and political economy," Commun. Res. Pr., vol. 10, no. 3, pp. 328-339, Jul. 2024, doi.org/10.1080/22041451.2024.2346428

[2] F. Mustafa and U. Leo, "On facial recognition and fundamental rights in India: A law and technology perspective," SSRN Electron. J., 2021, doi.org/10.2139/ssrn.3995958

[3] K. Sundara and N. Narendran, "The Digital Personal Data Protection Act, 2023: analysing India's dynamic approach to data protection," Comput. Law Rev. Int., vol. 24, no. 5, pp. 129-141, Oct. 2023, doi.org/10.9785/cri-2023-240502

[4] P. Naithani, "Analysis of India's Digital Personal Data Protection Act, 2023," Int. J. Law Manag., vol. 67, no. 5, pp. 543-553, Aug. 2025,. doi.org/10.1108/IJLMA-05-2024-0174

[5] C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: what it is and how it works," AI Soc., vol. 39, no. 4, pp. 1871-1882, Aug. 2024, doi.org/10.1007/s00146-023-01635-y

[6] M. Ananny and K. Crawford, "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability," New Media Soc., vol. 20, no. 3, pp. 973-989, Mar. 2018, doi.org/10.1177/1461444816676645

[7] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," Big Data Soc., vol. 3, no. 1, p. 205395171562251, Jan. 2016, doi.org/10.1177/2053951715622512

[8] Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Inf. Fusion, vol. 58, pp. 82-115, Jun. 2020,. doi.org/10.1016/j.inffus.2019.12.012

[9] L. Floridi et al., "AI4People-an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," Minds Mach. (Dordr.), vol. 28, no. 4, pp. 689-707, Nov. 2018, doi.org/10.1007/s11023-018-9482-5

[10] L. Floridi and J. Cowls, "A unified framework of five principles for AI in society," Harv. Data Sci. Rev., vol. 1, no. 1, Jul. 2019, 10.1162/99608f92.8cd550d1

[11] S. Alanoca, S. Gur-Arieh, T. Zick, and K. Klyman, "Comparing apples to oranges: A taxonomy for navigating the global landscape of AI regulation," in Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, New York, NY, USA: ACM, Jun. 2025, pp. 914-937. doi.org/10.1145/3715275.3732059

[12] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," Nat. Mach. Intell., vol. 1, no. 9, pp. 389-399, Sep. 2019. doi.org/10.1038/s42256-019-0088-2

[13] B. Mittelstadt, "Principles alone cannot guarantee ethical AI," Nat. Mach. Intell., vol. 1, no. 11, pp. 501-507, Nov. 2019. doi.org/10.1038/s42256-019-0114-4

[14] M. Veale and F. Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act - Analysing the good, the bad, and the unclear elements of the proposed approach," Comput. Law Rev. Int., vol. 22, no. 4, pp. 97-112, Aug. 2021. doi.org/10.9785/cri-2021-220402

[15] S. Wachter, B. Mittelstadt, and L. Floridi, "Why a right to explanation of automated decision-making does not exist in the general data protection regulation," Int. Data Priv. Law, vol. 7, no. 2, pp. 76-99, May 2017. doi.org/10.1093/idpl/ipx005

[16] A. D. Selbst and J. Powles, "Meaningful information and the right to explanation," Int. Data Priv. Law, vol. 7, no. 4, pp. 233-242, Nov. 2017. doi.org/10.1093/idpl/ipx022

[17] I. D. Raji et al., "Closing the AI accountability gap," in Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, New York, NY, USA: ACM, Jan. 2020. doi.org/10.1145/3351095.3372873

[18] M. Mitchell et al., "Model cards for model reporting," in Proceedings of the Conference on Fairness, Accountability, and Transparency, New York, NY, USA: ACM, Jan. 2019, pp. 220-229. doi.org/10.1145/3287560.3287596

[19] T. Gebru et al., "Datasheets for datasets," Commun. ACM, vol. 64, no. 12, pp. 86-92, Dec. 2021,. doi.org/10.1145/3458723

[20] N. A. Smuha, "From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence," Law Innov. Technol., vol. 13, no. 1, pp. 57-84, Jan. 2021, doi.org/10.1080/17579961.2021.1898300

[21] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial intelligence and the 'good society': The US, EU, and UK approach," Sci. Eng. Ethics, vol. 24, no. 2, pp. 505-528, Apr. 2018, doi.org/10.1007/s11948-017-9901-7

[22] H. Roberts, "The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation," AI Soc, vol. 36, no. 1, pp. 59-77, Mar. 2021. doi.org/10.1007/s00146-020-00992-2

[23] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," ACM Comput. Surv., vol. 54, no. 6, pp. 1-35, Jul. 2022, doi.org/10.1145/3457607

[24] T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," Minds Mach. (Dordr.), vol. 30, no. 1, pp. 99-120, Mar. 2020, doi.org/10.1007/s11023-020-09517-8

[25] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nat. Mach. Intell., vol. 1, no. 5, pp. 206-215, May 2019. doi.org/10.1038/s42256-019-0048-x

How to Cite: Mukund Kumar Pandey. (2025). Artificial intelligence governance in India: a critical legal framework for accountability, transparency, and regulatory compliance. Journal of Legal Subjects (JLS), 5(2), 1-12. https://doi.org/10.55529/jls.52.1.12

BIOGRAPHY OF AUTHOR



Mukund Kumar Pandey, is a Ph.D. Research Scholar at the Faculty of Law, Kalinga University, Raipur, Chhattisgarh, India. His research focuses on technology law, with particular emphasis on the legal and regulatory dimensions of Artificial Intelligence governance, data protection, and digital accountability in India. He holds prior academic qualifications in law and has presented and published work examining the intersection of emerging technologies with constitutional and statutory frameworks. His current doctoral research examines the adequacy of India's regulatory instruments

	in securing accountability and transparency in algorithmic decision-making systems. Email: maddy.muku2@gmail.com