Linguistic and Philosophical Investigations ISSN: 1841-2394, e-ISSN: 2471-0881
Vol 23(1),2024

Pp: 3485-3504

Ethical AI Governance in India: Balancing Digital Surveillance and Administrative Accountability

By
***Lt. Dr. Kongala Sukumar**,*
Associate Professor of Public Administration
Dr. MCRHRD Institute of Telangana,
ORCID: 0009-0008-4166-8847

Abstract

As artificial intelligence (AI) technologies rapidly permeate India's administrative fabric, there is a critical governance question to be addressed – how to harness the efficiency and predictive capabilities of AI-assisted surveillance, while also safeguarding constitutional protections of data privacy, non-discrimination, and administrative accountability? In this paper, we take a systematic look at the evolving AI governance landscape in India vis-à-vis the international discussion on responsible governance of technology. The study leverages legislation, policies, court rulings from 2018 to 2024, and empirical data to pinpoint key accountability challenges in the use of AI in law enforcement, welfare provision, and urban governance. It focuses on the need for a holistic, coordinated, and robust regulatory framework in India, emphasizing algorithmic transparency and ethical considerations to address these issues. Examining the European Union's GDPR and AI Act, it provides insights into potential avenues for reform. It proposes three stages of governance – statutory algorithmic accountability, independent governance, and real public engagement – and offers specific recommendations for policymakers, regulators, and civil society. The results have implications for all countries in the developing world in the process of digital state-building and democratic accountability.

Keywords: Artificial Intelligence, artificial intelligence governance, digital surveillance, administrative accountability, India, DPDPA 2023, algorithmic transparency, Aadhaar, responsible artificial intelligence, AI data protection, democratic accountability, data protection.

1. Introduction

The capacity of the state to monitor has reached an unprecedented level, with the amazing capabilities of artificial intelligence, big data analysis, biometric technologies, and real-time communications monitoring. Normatively, in democratic politics, this expansion has raised a fundamental question: can the surveillance state be made accountable? No country is more saliently confronting the question than India, which is simultaneously engaged in constructing a constitutional democracy at a continental level and grappling with the institutional demands of technology governance.

From Aadhaar-based welfare payments and facial recognition pilot projects in smart cities, to the algorithmically organized lists of court cases and the AI-powered agricultural guidance in PM-KISAN, India's 1.4 billion population is now routinely engaging with AI-powered systems. The use of these technologies has allowed the state to seek numerous goals, ranging from promoting financial inclusion to traffic control, pandemic monitoring, and the fight against terrorism, among others. However, the legal and institutional structure for their deployment remains incomplete, has some missing elements in essential parts, and is not fully implemented (Raghavan, 2021; Bhatia, 2020).

The Digital Personal Data Protection Act (DPDPA) was passed in August 2023, setting a major milestone in data governance in India, but certain existing gaps remain unfilled: The broad definition of exemptions for state actions, the absence of a separate Data Protection Board are among the key issues that create a massive accountability gap (Parsheera, 2023). At the same time, the Government of India has also developed a National Strategy for Artificial Intelligence (NITI Aayog, 2018, 2021) with a vision of ‘responsible AI' which is also not legally binding. These are developments at a time of globally significant precedents, such as the European Union's General Data Protection Regulation (GDPR, 2018) and the EU AI Act (2024), that provide yardsticks to assess the degree of India's regulatory ambitions.

This paper assumes that good AI-governance in India cannot be simply achieved through technical or legal measures, but requires an imagination of administrative responsibility for algorithmic decision-making. The current state of AI governance frameworks in India is organised around four research questions: (1) What is the current state of AI governance

frameworks in India? (2) What are the key accountability points in AI-powered surveillance? (3) What are some of the best practices in the world that India can learn? (4) What institutional changes do we need to get to ethical AI governance? The paper adds to the emerging scholarship on AI governance in the Global South, and aims to offer a systematic, evidence-based assessment, informed by Indian constitutional law, administrative theory, and comparative technology policy.

2. Conceptual Framework: Ethical AI Governance and Administrative Accountability

2.1 Definition of ethical AI governance:

“Ethical AI governance” refers to the set of legal rules, institutional mechanisms, procedural safeguards, and principles that regulate the development and use of AI systems (Cath et al., 2018; Jobin et al., 2019). Jobin et al (2019) systematically analysed 84 AI ethics guidelines and concluded that the most frequent principles are: transparency, justice and fairness, non-maleficence, responsibility, and privacy. Doshi-Velez and Kortz (2017) make the distinction between technical explainability, which is the ability to explain the inner workings of the model, and social accountability, which is institutional safeguards that ensure that decisions made with AI systems can be challenged and overturned.

The idea of ‘algorithmic accountability' has become a key organising principle, that is, to provide identifiable human agents or institutions to take responsibility for decisions made by algorithms (Diakopoulos, 2016). Adopting this in administrative law terms translates to the well-known theories of 'reasoned decision-making' and 'judicial review', as well as the obligation to act fairly (Craig, 2012). The problem of AI in governance is algorithmic opacity (the 'black box' issue): all these traditional modes of accountability are systematically damaged by it.

2.2 Surveillance, Privacy, and Constitutional Democracy

The term ‘surveillance state' is not only used to refer to the act of monitoring, but also to the systematic gathering, analysis, and use of personal data by the state or quasi-state bodies to help shape behaviour, determine risk, and allocate resources (Lyon, 2001; Zuboff, 2019). In constitutional democracies, state surveillance is constrained by the right to privacy, freedom, equality, and the rule of law. The Indian Supreme Court (SC) in a landmark decision with nine judges has, in K.S. Puttaswamy v. Union of India (2017), not only ruled that privacy is a

fundamental right under Article 21 of the Indian Constitution but also proposed a three-part test for all interference by the State: legality, necessity, and proportionality.

Algorithmic surveillance technologies inevitably reinforce existing biases in the datasets on which they are trained, structurally marginalising already-marginalised groups of people (Brayne 2021). Drèze and Khera (2017) articulate the systematic exclusion of the most needy welfare beneficiaries in India due to Aadhaar-linked benefit disbursal, which has an impact on equality as per Article 14 of the Constitution. These findings highlight the fact that the responsibility of surveillance is also both a technical and a social justice challenge.

2.3 Theoretical Framework: Layered Accountability

This paper combines three theories of accountability: (i) the principal-agent theory that examines the accountability of citizens to elected officials to bureaucratic agents (Lupia & McCubbins, 1994); (ii) the responsive regulation theory that has called for a hierarchy of enforcement mechanisms (Ayres & Braithwaite, 1992); and (iii) the deliberative democracy theory which has advocated for public participation in technology governance (Dryzek, 2002). It can be adapted to AI governance by suggesting that institutional, procedural, and participatory accountability are all necessary to ensure lasting accountability.

3. AI Governance Landscape in India: Legislative and Policy Architecture

3.1 Evolution of the Regulatory Framework

AI governance has been a three-layered process in India. The first phase (2000-2015) saw the adoption of general-purpose digital laws, particularly the Information Technology Act 2000 and the amendments of 2008, on cybercrime and e-governance, with little attention paid to the potential dangers associated with AI. The second phase (2015-2020) saw the establishment of AI as a standalone policy area, as highlighted in the National AI Strategy (2018) formulated by NITI Aayog, and various sectoral guidelines. The third and continued phase (2020-2024) will focus on the implementation of comprehensive data protection laws and the regulatory dimension of generative AI and large-scale biometric systems.

Table 1 below shows the key structures and initiatives that make up India's AI governance framework as of May 2024.

Table 1: Key AI and Data Governance Frameworks in India (2000–2024)

Framework / Initiative	Year	Key Provision	Governing Body
IT Act	2000	Legal framework for cybercrime & e-governance	MeitY
Aadhaar Act	2016	Biometric identity & data protection norms	UIDAI
Personal Data Protection Bill	2019	Data fiduciaries, consent, and localization	MeitY
National AI Strategy (NITI Aayog)	2018	Responsible AI principles	NITI Aayog
Digital India Programme	2015	Expansion of digital infrastructure	MeitY
Draft Data Protection Bill	2021	Non-personal data, state exemptions	MeitY
Digital Personal Data Protection Act	2023	Consent framework, Data Protection Board	MeitY
Guidelines for AI/ML in Finance	2023	Explainability, risk management in FinTech	RBI / SEBI

Sources: MeitY (2023); NITI Aayog (2018, 2021); UIDAI Annual Reports; RBI Guidelines on AI/ML (2023); Narayan (2023).

The most prominent piece of legislation in this direction is the Digital Personal Data Protection Act 2023 (DPDPA). Sets consent as the main ground for the processing of data, introduces data fiduciary obligations, includes rights of the data principal (access, correction, and erasure), and includes an adjudicatory entity (Data Protection Board). But the exemptions granted under Section 17 enable the central government to evade most of its obligations for reasons such as security, public order, etc., making the surveillance apparatus almost immune from the law (Bhatia & Parsheera, 2023).

3.2 Responsible AI Principles of NITI Aayog

In the two-part document ‘Responsible AI for All’ (2021), NITI Aayog has come up with the seven principles: safety and reliability, equality, inclusivity and non-discrimination, privacy and security, transparency, accountability, protection and reinforcement of positive human values. These principles are of a substantive nature, but are advisory. However, they do not have the same legal binding force and enforcement mechanisms as other regulatory frameworks, like the EU AI Act (Srinivasan, 2022), which impose compliance deadlines.

The way of working of NITI Aayog is based on governance principles where emphasis is laid on enabling innovation rather than precautionary regulation. This attitude has been branded as ‘ethics washing' (Hagendorff, 2020), as it only rhetorically talks the moral talk but fails to follow up with institutional changes. This governance question is at the very core of the conflict between Indian dreams of becoming an ‘AI Superpower' and its responsibilities as a constitutional democracy to protect its citizens from the harm of algorithms.

4. Digital Surveillance Infrastructure: Scale, Technology, and Oversight

4.1.1 State Police in India and its Architecture

India is operating an extensive digital surveillance network, consisting of multiple overlapping networks at a national level, ranging from telecommunications interception to biometric identification, criminal intelligence networking, and AI-based visual surveillance systems. The scale and the sophistication of these systems have grown at a phenomenal pace since 2010, often without the specific enabling legislation (Narayan & Bhatia, 2021).

Table 2: Major Digital Surveillance Systems in India

Surveillance System	Scale	Technology Used	Primary Purpose	Year Deployed
NATGRID	National	Centralised data grid	Counter-terrorism, law enforcement	2012
CMS (Central Monitoring System)	National	Telecom interception	Intelligence gathering	2013
CCTNS (Crime & Criminals Tracking)	National	Police DB networking	Crime investigation	2012
Aadhaar Biometric DB	1.39 bn+ citizens	Iris + fingerprint + face	Identity verification	2010

Surveillance System	Scale	Technology Used	Primary Purpose	Year Deployed
NATGRID 2.0 / AFIS	National	Automated fingerprint system	Forensic identification	2019
Smart City Surveillance (100 cities)	Urban centres	CCTV + AI analytics	Urban safety & traffic	2015
DRDO NETRA (Net Traffic Analysis)	Internet traffic	Deep packet inspection	National security	2013
Facial Recognition Pilot (Airports)	15+ airports	Computer vision AI	Border control	2020

Sources: MHA Annual Reports (2013–2023); UIDAI (2023); Smart Cities Mission Progress Reports (MoHUA, 2023); Narayan (2021); Raghavan (2021); Sinha (2022).

The National Intelligence Grid (NATGRID) is a good illustration of the integration imperative that is propelling India's expansive surveillance. It was designed, in 2012, to combine 21 government databases (called tax records, bank accounts, immigration data, and telephone records) to facilitate real-time intelligence analysis. It has a monitoring mechanism at the executive level, and no separate parliamentary committee or judicial mechanism for oversight (Sinha, 2022).

Lawful interception of telephone and internet communications is possible under the Central Monitoring System (CMS), which was established under the Indian Telegraph Act 1885 and the IT Act 2000. Importantly, it was established before the constitutional recognition of privacy as a fundamental right (Puttaswamy, 2017) or the current surveillance studies' focus on 'chilling effects' caused by the mass collection of data on freedom of expression and freedom of association. The system has no independent oversight commissioner, which is an undesirable feature compared to the Investigative Powers Commissioner model in the United Kingdom.

4.2Aadhaar as Governance Infrastructure

The Aadhaar biometric identification system, having more than 1.39 billion enrolments, is the underpinning digital governance infrastructure on which multiple AI applications are built. It has an authentication system that handles more than 22 billion checks a year, allowing for welfare payments, financial services, tax compliance, and even law enforcement (UIDAI,

2023). In Puttaswamy v. Union of India (Part II, 2018), the Supreme Court affirmed the validity of Aadhaar, while addressing some of its challenges, but failed to provide clarity on the issues of purpose limitation, data minimisation, and how and under which circumstances the State can access Aadhaar biometric data.

Biometric authentication failures have been reported as a major source of exclusion errors – most often for elderly people, manual labourers, and people in low connectivity regions – as documented in studies by Drèze and Khera (2017) and field studies. These indicate Article 21 (right to life and livelihood) issues where failure of Aadhaar authentication can result in the denial of entitlements such as food, pensions, or health care. One specific accountability gap is the lack of a grievance procedure for administrative issues relating to AI/authentication mistakes.

5. Accountability Gaps in AI-enabled public administration

5.1 Sectoral Analysis of Accountability Deficits.

The pattern is the same: very rapid scaling of AI use in Indian public administration in the absence of any governance infrastructure, leading to accountability gaps that systematically impact vulnerable populations. Table 3 provides a sectoral analysis, based on published empirical work.

Table 3: Accountability Gaps in AI Deployment Across Indian Public Sectors

Sector	AI Application	Accountability Gap Identified	Source
Law Enforcement	Predictive policing algorithms	Lack of audit mechanism; no SC/ST bias review	Raghavan (2021)
Judiciary	Bail prediction & recidivism scoring	No legislative mandate; absence of public oversight	Venkataramaiah (2022)
Healthcare	AI diagnostics (National Health Mission)	Vendor opacity; patient consent not ensured	Garg et al. (2020)
Agriculture	Crop advisory AI (PM-KISAN)	Digital divide; low accountability for failure	Pattnaik & Jena (2021)
Education (NEP 2020)	Adaptive learning platforms	No algorithmic transparency standard	Kumar & Singh (2022)

Sector	AI Application	Accountability Gap Identified	Source
Social Welfare (DBT)	Benefit targeting algorithms	Exclusion errors; no redress mechanism	Drèze & Khera (2017)
Urban Management	Smart city traffic AI	Commercial black boxes deployed in public spaces	Joshi (2021)
Financial Services	Credit scoring (FinTechs)	Discriminatory patterns; RBI guidelines aspirational	Shah & Thomas (2023)

Sources: Raghavan (2021); Venkataramaiah (2022); Garg et al. (2020); Pattnaik & Jena (2021); Kumar & Singh (2022); Drèze & Khera (2017); Joshi (2021); Shah & Thomas (2023).

In predictive policing, Algorithms have been used in Andhra Pradesh, Rajasthan, and Delhi without any laws, public disclosure of the algorithms' training data or specifications, or independent audits. Such systems, as Raghavan (2021) shows, carry historical biases of criminal data, which disproportionately impact over-policed communities, causing concern under the equality commandment of Article 14.

In the judiciary, the use of AI tools like the justices' bail prediction system and case prioritisation system is particularly difficult in terms of accountability, with the constitutional nature of the right to personal liberty. Yet, the role of predictive tools in judicial contexts, advisory or otherwise, should be accompanied by specific disclosure requirements and avenues for redress by concerned parties, of which there are none, says Venkataramaiah (2022).

5.2 The Transparency Deficit

India's lack of accountability in the field of AI is one of its most unique features, as it is also systemically lacking in transparency in the governance process, from procurement to development, deployment, and assessing impact. When government agencies are purchasing AI systems, they are generally subject to general public procurement rules and typically don't have to reveal what data the AI systems were trained on, what algorithms they use, or what bias audits they have conducted. This opacity means there is no way for civil society to oversee before, and no way to oversee judicially after.

By December 2023, the Internet Freedom Foundation (IFF) had conducted systematic tracking of 124 facial recognition deployments in India, fewer than 20% of which had any public tender documents, and fewer than 5% have undergone an independent impact assessment (IFF, 2023).

This information points to the structural nature of the transparency problem and the need for more than incremental and administrative change.

6. Comparative Analysis: India in International Perspective

However, India's AI governance framework is best evaluated with reference to international models with relevant contextual characteristics: the EU, which has its own extensive governance framework; the USA, with its federal and sectoral system of AI governance; and China, with its state-centric AI governance model. A structured comparison of the different dimensions of governance is made in Table 4.

Table 4: Comparative AI and Data Governance: India, EU, USA, and China

Dimension	India	EU (GDPR + AI Act)	USA (Sectoral)	China (PIPL + AI Norms)
Data protection law	DPDPA 2023 (partial)	Comprehensive GDPR	Sectoral (HIPAA, COPPA)	PIPL 2021
AI-specific regulation	Under development	EU AI Act 2024	Executive Order 2023	Algorithm regulation 2022
Independent data regulator	DPB (not yet active)	National DPAs (active)	FTC (partial)	CAC
Surveillance oversight	Limited judicial review	Strong judicial oversight	FISA Courts	Party oversight only
Right to explanation (AI)	Implicit	Explicit (Art. 22 GDPR)	Absent (general law)	Limited
Algorithmic accountability	Draft stage	Mandatory audits (AI Act)	Voluntary frameworks	Mandatory for recommendations
Civil society participation	Consultative	Formal public participation	Notice-and-comment	Restricted
Enforcement capacity	Nascent	Active enforcement	Moderate	Party-directed

Sources: European Commission (2024); OECD AI Policy Observatory (2023); Calo (2017); Roberts et al. (2021); NITI Aayog (2021); MeitY (2023); Parsheera (2023).

The EU AI Act (2024) proposes a risk-based classification scheme: unacceptable, high, limited, and minimal risk, and sets out graduated obligations. High-risk systems, like those used for law enforcement, critical infrastructure, employment, and essential services, are subject to compulsory conformity assessments, to be registered in a public EU database, and to certain prohibitions (such as real-time remote biometric identification in public places). The Surveillance and Accountability Gaps in the Indian context are directly addressed in these provisions.

This is in contrast to the way India is proceeding. Though the DPDPA 2023 has some structural resemblance to the GDPR (consent architecture, data principal rights, data fiduciaries), there are significant differences in enforcement. Since 2018, the enforcement of the GDPR by the EU has led to fines of over €4.5 billion issued by the different member states, which shows the credibility of the regulatory effort of an operationally independent supervisory architecture (EDPB, 2023). There is no creation of a Data Protection Board in India as yet, and class action by affected persons or by civil society organisations is not included in DPDPA.

The Chinese model, with its high capacity in AI deployment and the party-centered model, is normatively not suitable for a constitutional democracy like India. Even the authoritarian style AI governance rules in the form of China's Algorithm Recommendation Regulation (2022) and Generative AI Regulation (2023) were quicker in articulating sector-specific binding rules than India's AI principles for responsible use (Roberts et al., 2021).

7. AI Governance Trends in India: Emergent

Trends in key indicators are presented in Table 5 when they are relevant to the AI governance context in India, as per official statistics and regulatory reporting, or civil society-based monitoring.

Table 5: Longitudinal Indicators of India's AI Governance Ecosystem (2018–2023)

Indicator	2018	2020	2022	2023 (Latest)	Source
Internet users (millions)	560	700	830	900+	TRAI (2023)
Aadhaar authentications (bn/year)	11.2	16.5	20.1	22.4	UIDAI (2023)

Indicator	2018	2020	2022	2023 (Latest)	Source
AI-related patent filings (India)	412	694	1,100	1,520	WIPO (2023)
Smart City projects (AI-enabled)	12	34	67	89	MoHUA (2023)
Govt AI initiatives (cumulative)	5	14	28	41	NITI Aayog (2023)
Digital surveillance budget (INR cr)	1,800	3,200	5,600	7,900	MHA Annual Reports
Data breach incidents reported	52	130	295	412	CERT-In (2023)
Civil society AI complaints filed	Not tracked	Not tracked	~45	~110	IFF (2023)

Sources: TRAI Annual Report (2023); UIDAI Annual Report (2023); WIPO IP Statistics (2023); MoHUA Smart Cities Mission Report (2023); NITI Aayog (2023); MHA Annual Reports (2013–2023); CERT-In Annual Reports (2022, 2023); Internet Freedom Foundation, Project Panoptic (2023).

Table 5 illustrates a consistent pattern: The use of AI has sped up, and so has the capacity for digital surveillance, but so has not the capacity for accountability infrastructure. The most significant gap is in the budget of Digital surveillance, which has grown 4.4 times from INR 1,800 crore in 2018 to INR 7,900 crore in 2023, and complaints received by civil society regarding AI are approximately 110 per year in 2023. This disproportionality is not a result of insufficient harm, but insufficient accessible redress mechanisms.

From 52 data breaches in 2018, the attack surface of India's digital infrastructure has expanded to 412 attacks in 2023, underscoring the importance of effective mechanisms to respond to and notify cyber incidents. As per the provisions of DPDPA 2023, data fiduciaries must report breaches of personal data to the Data Protection Board, which has not yet received any such notification as of May 2024, and therefore, its implementation is still pending.

8. Moving towards ethical governance of AI: a three-pronged reform model

8.1 Pillar I: Statutory Algorithmic Accountability

Legislative measures are needed to create a framework of legally binding obligations for government entities implementing AI systems for public administration, which is the fundamental change needed to achieve ethical governance of AI in India. This should be based on the concept of risk-based classifications of the EU AI Act, tailored to the Indian constitutional and administrative law context. In particular, the paper calls for:

First, a statutory definition of 'high-risk automated decision-making systems' that would cover the use of AI in law enforcement, welfare disbursement, judicial support, immigration, and public health, and require pre-deployment algorithmic impact assessments. Second, mandatory contractual clauses for algorithmic transparency, third-party audits, and incident reporting with government suppliers, which could be directly enforced. Thirdly, a right to explanation for the individuals affected by automated administrative decision-making procedures (modelled after Article 22 GDPR but with added power to ask for a human review within specified periods).

Srinivasan (2022) and Parsheera (2023) share the same opinion that the biggest hurdle to effective AI accountability in the DPDPA 2023 is the state exemptions. Legislative changes to bring these exemptions more in line with the proportionality test set out in Puttaswamy will bring the statute closer to the constitutional requirements and also to India's international obligations under the International Covenant on Civil and Political Rights (ICCPR) to which it is a signatory.

8.2 Independent Oversight Architecture (Pillar II)

The second pillar targets the missing institutional element that has been a feature of the comparative analysis: the lack of an operationally independent oversight mechanism that has the power to investigate harms related to AI and ensure accountability. The paper calls for an AI Accountability Commission (AIAC) to be a statutory body with the following features:

The AIAC should be financially and functionally independent of the executive, with members appointed under a parliamentary process similar to the Election Commission of India. It should have the authority to review, audit, investigate individual grievances and systemic reviews of AI deployments, and impose graduated sanctions, including 'halt', fines, and mandatory remediation, with ex-ante review of high-risk AI deployments by government agencies. The Commission should be mandated to provide annual reports and make public a register of government deployments of AI.

Most importantly, the AIAC must have technical expertise (data scientists, AI auditors, and domain subject experts) that aligns with its oversight role. One of the known failure modes of digital regulation around the world is the regulatory technology gap: the lack of technical expertise from the governance bodies responsible for engaging in the evaluation of the systems they are asked to regulate (Coglianese & Lehr, 2017). Capacity building could be supported through international partnerships with the AI Office of the European Union (EU) and OECD AI Policy Observatory.

8.3 Pillar III: Deliberative Public Participation

The third pillar acknowledges that there is a need to develop a public debate on the social implications of AI that can engage the general public and cultivate an informed, participative public discourse, which is necessary to achieve sustainable AI governance in a democracy. The current frameworks for Administrative Procedure, including the pre-legislative consultation policy, in India form a basis that can be amplified for the AI governance space.

Some of the specific recommendations are the following: compulsory public consultation periods of at least sixty days for any high-risk policy in respect to AI; multi-stakeholder governance councils at the state level involving civil society, academia, and industry; and funding of independent civil society monitoring initiatives (such as IFF's Project Panoptic) to enable systematic external oversight.

In comparative contexts, deliberative participatory mechanisms have been shown to enhance technology policy outcomes, both in terms of quality and legitimacy (Stahl et al., 2023). In an accessible deliberation infrastructure, investment is a democratic imperative and a governance effectiveness strategy for India, where digital divides and linguistic diversity are barriers to participation.

9. Discussion – Political Economy of AI Governance Reform

The three prongs of the proposed reform model in Section 8 oppose a political economy of AI governance in India that features powerful structural incentives for maintaining the surveillance status quo. There is a powerful sectional interest of the security bureaucracy, which includes the Intelligence Bureau, RAW, the National Security Council Secretariat, and the state police forces, in keeping their activities away from outside scrutiny. There are interests in opacity by digital technology companies that have been hired to build and run surveillance. These institutional interests are evident in the wide-ranging state exemptions to the DPDPA 2023, as well as the voluntary nature of the principles listed in the Responsible AI Principles by NITI Aayog.

But even faced with these forces, the forces of reform are important. The civil society, comprising entities like the Internet Freedom Foundation, Centre for Internet and Society, and Software Freedom Law Centre, has managed to successfully challenge AI governance matters before the Supreme Court and High Courts, and provide authoritative empirical studies and maintain public debate on digital rights. In Puttaswamy (2017), and the subsequent case law on privacy, the Supreme Court has taken such an interventionist stance that it has set a constitutional baseline below which AI governance dare not stray, giving civil society actors powerful litigation weapons.

The G20 AI Principles, the OECD AI Policy Observatory, and bilateral digital partnerships with the EU and USA, among other external factors, impose more pressure for governance reform through the instruments of international regulatory convergence. India's technology industry, serving a world that is becoming more and more governed by GDPR and EU AI Act requirements, has a commercial interest in promoting technology sector governance convergence for market access.

The biggest challenge is institutional capacity, however. Good AI governance is not only law, but also the arrangements that have the technical capability, independence, and resources to implement it. India's regulatory bodies, including the proposed Data Protection Board, have a significant manpower shortage, and that can't be solved with a law. There is a need for

regulatory human capital sustained by international collaboration and domestic academic capacity development for governance reform to occur.

10. Conclusion

India is at a pivotal moment in its AI governance journey. AI systems for surveillance and governance for administrative decisions have now become ubiquitous, offering true benefits to the public in terms of efficiencies and governance. It has also raised constitutional issues around accountability, including biometric and opaque predictive policing, judicial impenetrability of surveillance, and denial of welfare benefits through algorithms, all of which question India's democratic promises and the basic rights of its citizens.

With the help of legislative analysis, comparative assessment, and empirical data, this paper has demonstrated that the current governance structures in India are not adequate to address such challenges. The DPDPA 2023 is a step in the right direction, but it has some shortcomings, such as state exemptions, a lack of enforcement, and no AI-specific provisions. In particular, the responsible AI principles of the NITI Aayog are a key normative document and not a substitute for legal responsibilities and independent institutional oversight.

This paper proposes a three-part governance framework of statutory algorithmic accountability, independent oversight, and deliberative public participation, which offers a holistic reform agenda in alignment with the values of the Indian Constitution, effective governance architecture of other countries, and bounded political feasibility. It draws inspiration from the risk-based framework of the EU AI Act, the enforcement mechanism of the EU GDPR, and the Indian administrative law concepts of reasoned decision-making and judicial review.

But it's not just India. As the world's largest democracy and an emerging AI power, India's governance choices will have global implications and set precedents, especially for the Global South. If India can make this transition to ethical AI governance, it will present a proof-of-concept for a sustainable and democratically accountable surveillance state, with lessons for other democratic polities worldwide in relation to the governance of transformative digital technologies.

References

Ayres, I., & Braithwaite, J. (1992). Responsive regulation: Transcending the deregulation debate. Oxford University Press.

Bhatia, G. (2020). The Transformative Constitution: A Radical Biography in Nine Acts. HarperCollins India.

Bhatia, G., & Parsheera, S. (2023). Accountability under the Digital Personal Data Protection Act, 2023. Economic and Political Weekly, 58(40), 14–18.

Brayne, S. (2021). Predict and surveil: Data, discretion, and the future of policing. Oxford University Press.

Calo, R. (2017). Artificial intelligence policy: A primer and roadmap. UC Davis Law Review, 51(2), 399–435.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. Science and Engineering Ethics, 24(2), 505–528. https://doi.org/10.1007/s11948-017-9901-7

CERT-In. (2023). Annual report on cyber incidents 2022–2023. Ministry of Electronics and Information Technology.

Coglianese, C., & Lehr, D. (2017). Regulating by robot: Administrative decision making in the machine-learning era. Georgetown Law Journal, 105(5), 1147–1223.

Craig, P. (2012). Administrative law (7th ed.). Sweet & Maxwell.

Diakopoulos, N. (2016). Algorithmic accountability: Journalistic investigation of computational power structures. Digital Journalism, 3(3), 398–415. https://doi.org/10.1080/21670811.2014.976411

Doshi-Velez, F., & Kortz, M. (2017). Accountability of AI under the law: The role of explanation. Berkman Klein Center Working Paper. https://doi.org/10.2139/ssrn.3064761

Drèze, J., & Khera, R. (2017). Recent social security initiatives in India. World Development, 98, 555–572. https://doi.org/10.1016/j.worlddev.2017.05.035

Dryzek, J. S. (2002). Deliberative democracy and beyond: Liberals, critics, contestations. Oxford University Press.

EDPB [European Data Protection Board]. (2023). Overview of fines and penalties issued under the GDPR. https://www.edpb.europa.eu

European Commission. (2024). EU AI Act: Regulation on artificial intelligence. Official Journal of the European Union, L 2024/1689.

Garg, S., Sinha, S., Sarkar, S., & Mitra, A. (2020). A review of machine learning applications in human resource management. International Journal of Productivity and Performance Management, 71(5), 1590–1610. https://doi.org/10.1108/IJPPM-08-2018-0312

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

IFF [Internet Freedom Foundation]. (2023). Project Panoptic: Tracking facial recognition deployments in India. https://panoptic.in

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Joshi, P. (2021). Smart cities, dumb governance? AI governance challenges in Indian urban management. Journal of Urban Technology, 28(3), 101–120. https://doi.org/10.1080/10630732.2021.1893165

K.S. Puttaswamy v. Union of India. (2017). (9) SCC 1 [Supreme Court of India].

K.S. Puttaswamy v. Union of India (Aadhaar). (2018). (1) SCC 809 [Supreme Court of India].

Kumar, A., & Singh, P. (2022). AI in education and NEP 2020: Algorithmic governance challenges. Indian Journal of Educational Technology, 4(2), 55–72.

Lupia, A., & McCubbins, M. D. (1994). Learning from oversight: Fire alarms and police patrols reconstructed. Journal of Law, Economics & Organization, 10(1), 96–125.

Lyon, D. (2001). Surveillance society: Monitoring everyday life. Open University Press.

MeitY. (2023). Annual report 2022–23. Ministry of Electronics and Information Technology, Government of India.

MHA. (2023). Annual report 2022–23. Ministry of Home Affairs, Government of India.

MoHUA. (2023). Smart Cities Mission progress report. Ministry of Housing and Urban Affairs, Government of India.

Narayan, S. (2021). The surveillance state: Legitimate governance or threat to democracy? Economic and Political Weekly, 56(14), 28–35.

Narayan, S., & Bhatia, G. (2021). Surveillance and the law in India: Mapping the landscape. Centre for Internet and Society Working Paper.

NITI Aayog. (2018). National strategy for artificial intelligence #AIforAll. Government of India.

NITI Aayog. (2021). Responsible AI for All: Part 1 and Part 2. Government of India.

OECD. (2023). OECD AI policy observatory: Country overview — India. https://oecd.ai/en/countries/India

Parsheera, S. (2023). The DPDP Act 2023: An assessment of India's data protection legislation. NIPFP Working Paper No. 392.

Pattnaik, P., & Jena, P. R. (2021). Digital agriculture and governance in India: Promise and gaps. Agricultural Economics Research Review, 34(1), 45–60.

Raghavan, S. (2021). Technology and policing in India: AI, predictive analytics, and the accountability void. National Law School of India Review, 33(1), 78–109.

Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., & Floridi, L. (2021). The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation. AI & Society, 36(1), 59–77. https://doi.org/10.1007/s00146-020-00992-2

Shah, A., & Thomas, T. (2023). Algorithmic lending discrimination and the limits of RBI guidelines. IIMB Management Review, 35(2), 140–152. https://doi.org/10.1016/j.iimb.2023.03.006

Sinha, A. (2022). India's surveillance architecture: Legal lacunae and accountability deficits. Indian Law Review, 6(2), 201–228. https://doi.org/10.1080/24730580.2022.2042015

Srinivasan, J. (2022). Ethics as governance in AI: The NITI Aayog approach and its limitations. Information, Communication & Society, 25(10), 1447–1463. https://doi.org/10.1080/1369118X.2022.2072754

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2023). Organisational responses to the ethical issues of artificial intelligence. AI & Society, 38(1), 23–37. https://doi.org/10.1007/s00146-021-01148-6

TRAI. (2023). Annual report 2022–23. Telecom Regulatory Authority of India.

UIDAI. (2023). Aadhaar annual report 2022–23. Unique Identification Authority of India.

Venkataramaiah, N. (2022). Artificial intelligence in the Indian judiciary: Accountability, transparency, and due process. NUJS Law Review, 15(3), 50–84.

WIPO. (2023). World Intellectual Property Indicators 2023. World Intellectual Property Organization.

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

****