Vegueta. Anuario de la Facultad de Geografía e Historia
24 (1) 2024
eISSN: 2341-1112

TRUST UNDER WATCH: PUBLIC PERCEPTIONS OF AI-ENABLED SURVEILLANCE AND ADMINISTRATIVE LEGITIMACY IN INDIA

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

Abstract

The growing prevalence of so-called artificial intelligence (AI)-based surveillance in India, from biometric identification to facial recognition, crime tracking systems, and smart city platforms, has altered the state's relationship with its people. The paper examines the perceptions, negotiation, and evaluation of the Indian publics of AI-based administrative surveillance, as well as how these perceptions relate to other concepts of state legitimacy, procedural justice, and institutional trust. Fusing survey data from nationally representative studies, secondary empirical literature and comparative policy analysis across six countries, the study highlights a paradox: while the vast majority of respondents view surveillance technologies as crime deterrents, there are substantial minorities, especially among marginalised groups, Scheduled Castes and Tribals and lower income urban groups, who are deeply suspicious based on their experiences of being excluded by algorithms, prevented from access to Aadhaar, and unable to access the identity data of the state. The most predictive factors for institutional trust are perceived procedural fairness and accountability of government. The paper claims that the governance gap in India, which includes the lack of an independent AI oversight framework, weak enforcement of data protection laws, and unchecked facial recognition applications by the courts, undermines the legitimacy of the government even when citizens use surveillance as a means to an end. Recommendations include the following: an AI Governance Act to create regulations for the use of AI; moratoriums on automated facial recognition in the public sphere; algorithmic audits and community-level digital redressal mechanisms. The study contributes to growing literature on surveillance studies, public administration, and the political sociology of technology in the Global South.

Keywords: *AI surveillance; public trust; administrative legitimacy; Aadhaar; facial recognition; India; data governance; procedural justice; smart cities; digital rights*

1. Introduction

In the second decade of the twenty-first century, India has emerged as one of the world's most ambitious deployers of artificial intelligence in public administration. The Indian State has created a massive technological apparatus that monitors and records, identifies and increasingly decides over the lives of its citizens, ranging from the Unique Identification Authority of India's (UIDAI) Aadhaar biometric system, which covers more than 1.38 billion people, to automated facial recognition systems (AFRS) deployed in Tamil Nadu, Delhi, and Uttar Pradesh. The Indian State has developed a large technological infrastructure, from the Unique Identification Authority of India (UIDAI) Aadhaar biometric system, which covers more than 1.38 billion residents, to automated facial recognition systems (AFRS), deployed throughout Tamil Nadu, Delhi, and Uttar Pradesh. Smart City Integrated Command and Control Centres (ICCCs) have now been implemented in 100 cities and are connected with surveillance cameras, traffic sensors, and data analytics. (Ministry of Housing and Urban Affairs [MoHUA], 2022)

The rollout of AI-powered surveillance comes at a critical time in the current discussion around the governance of algorithmic power. Who watches the watcher? is a question that states are confronting, such as with the European Union's AI Act (2024) and China's Personal Information Protection Law (2021). How do you fix the error? What's the impact of algorithmic decision-making on the citizen/citizen-state dyad? These questions have specific relevance for India, where the nation is at the same time a majoritarian democracy with a robust civil society as well as a nation with a history of authoritarian state monitoring of dissent, minorities, and tribal communities (Roy, 2019; Sinha, 2021).

However, the perception of this expansion of surveillance has been under-researched empirically. Current

studies focus either on techno-optimist perspectives on the efficiency of e-governance (Bhatnagar, 2014; Madon, 2009) or civil liberties concerns over individual platforms, like Aadhaar (Khera, 2019; Ramakumar, 2018). What is missing is a systematic analysis of how ordinary Indian citizens—across class, caste, region, and gender—understand, evaluate, and respond to the AI surveillance apparatus that has come to permeate their everyday administrative interactions. This paper seeks to address that gap.

The central argument advanced here is that public acceptance of AI surveillance in India is structured by a foundational paradox: citizens instrumentally accept surveillance technologies as security goods while simultaneously withdrawing trust from the administrative systems that deploy them, particularly where those systems have produced documented harms—benefit exclusions, misidentifications, and data breaches—without accountability. This paradox, we believe, is an ongoing legitimacy crisis that could be internalized into India's digital governance system if the necessary changes are not made promptly.

The paper is organized as follows. In Section 2, the theoretical frameworks relating surveillance and trust to administrative legitimacy are discussed. In Section 3, the authors give an overview of the landscape of AI surveillance in India. The methodological approach taken in the study is presented in Section 4. Section 5 presents empirical results from across demographies on perceptions of the public. A comparative cross-national perspective is given in Section 6. The policy and governance implications are explored in Section 7. The recommendations follow later in Section 8.

2. Theoretical Framework

2.2 Military History

Michel Foucault (1977) has traced the genealogy of academic study of surveillance back to his analysis of the Panopticon as a technology of power that disciplines subjects through the internalisation of the gaze. This framework has been expanded by contemporary surveillance studies scholars like Lyon (2007) and Zuboff (2019) to include digital and algorithmic surveillance, who reject the notion of a 'surveillance society' and prefer to speak of data assemblages that are distributed, networked, and often invisible. However, in democratic states, surveillance is initiated by a legitimating logic that sets public safety apart from social control. The real question is whether AI-powered surveillance is a tool for enhancing state capacity to meet the needs of citizens or reconfiguring the citizen-state relationship into an interaction of unequal scrutiny.

Breckenridge's (2014) attempt to theorise the ‘biometric state’ in the Global South is relevant to the Indian context because, in it, identification technologies serve as tools of bureaucratic inclusion and exclusion. Aadhaar is a good example of this: it is intended to include everyone in welfare entitlements and yet produces systematic exclusions when identification is not possible because of the use of a biometric trait like a fingerprint, in this case, of people who are old, manual labourers, and nutritionally deficient whose fingerprints wear away or are not readable (Khera, 2019). In recent years, Dencik et al. (2019) have claimed that data-driven welfare systems create a 'data paternalism' in which algorithmic authorities replace deliberative governance, thus compromising procedural legitimacy.

2.2 Trust and Administrative Legitimacy

The dangers of state surveillance of the public have been theorized in several ways. According to Tyler's (1990, 2006) procedural justice model, citizens' perceptions of legal and governmental processes (voice, neutrality, respect, trustworthy motives) are more important than outcomes for shaping citizens' trust in legal and governmental authorities. This model assumes that technologies that are thought to be effective for crime reduction, even when used for this purpose, will not instill trust if they are seen as arbitrary, discriminatory, or opaque.

Levi and Stoker (2000) place the construct of public trust in government into a wider legitimacy context by breaking it down into a cognitive evaluation of the performance of the government and affective aspects based on social identity and prior experience. This emotional aspect can be more significant than instrumental assessments of the accuracy of AI systems in the case of marginalised communities in India, including Adivasis, Dalits, and religious minorities, who have often been targeted by the state in the name of marginalisation rather than of protection. The cross-national analysis by Norris and Inglehart (2019) of democratic legitimacy defines the main factors that motivate institutional trust around the world: transparency and equal treatment, which are relevant to AI governance.

2.3 Governance of AI and Algorithmic Accountability

The governance issues relating to AI systems in public administration are explored in an emerging interdisciplinary literature (Doshi-Velez et al., 2017; Mittelstadt et al., 2016; Reisman et al., 2018). The issues of concern are: opacity (the ‘black box’ problem); disparate error rates among different demographic groups; lack of meaningful

human review in consequential decisions; and lack of redressal mechanisms for algorithmic harm. In the Indian context, Singh and Gupta (2020) report some extreme examples of algorithmic harms without accountability, such as starvation deaths in Jharkhand due to being excluded from the Public Distribution System. In Indian cities, Datta (2022) examines the impact of Smart City surveillance platforms to create urban inequality by focusing on the monitoring of informal settlements, while being under monitored in elite residential areas.

Responsiveness theory (Bovens, 2007; Mulgan, 2000) suggests that a key component of administrative legitimacy is the existence of means by which affected citizens can object, challenge, and demand redress of any bureaucratic action or decision. One of the challenges of using AI in public administration is that if a citizen's welfare benefit application is rejected by an algorithm, to whom can they go? Attribution of error and responsibility is particularly challenging with algorithmic systems because they are opaque (Doshi-Velez et al., 2017). This theoretical literature serves as a foundation to assess the existing framework of AI surveillance governance in India.

3. The AI Surveillance Landscape in India

3.1 Key Systems and Platforms

India's AI surveillance infrastructure has expanded through three interrelated streams: identity and welfare administration, law enforcement and intelligence, and smart urban governance. Table 1 maps the principal systems across these streams.

Table 1

Major AI-Enabled Surveillance Systems in India (2009–2023)

AI Surveillance System	Year Launched	Coverage (Cities/States)	Key Function	Implementing Agency
Aadhaar Biometric ID	2009 (scaled 2014–)	Pan-India (1.38 bn enrolled)	Identity authentication, DBT linkage	UIDAI / MeitY
CCTNS (Crime & Criminal Tracking Network)	2009 (operational 2019)	All 36 States/UTs	Police record integration, FIR digitisation	MHA / NCRB
AFRS / Face Recognition (TNFRS & others)	2018–2022	Tamil Nadu, Delhi, Telangana, UP	Criminal identification, crowd monitoring	State Police Depts
NATGRID	2014 (Phase II 2021)	Central agencies	Integrated intelligence database	MHA / Intelligence Bureau
PM-WANI / Wi-Fi Analytics	2021	Urban/semi-urban India	Public Wi-Fi access logging	DoT / TRAI
CoWIN Health Surveillance	2021	Pan-India	Vaccine registry, mobility tracking	MoHFW / NIC
Smart City ICCC Platforms	2016–2023	100 Smart Cities	Integrated urban surveillance, IoT	MoHUA / Smart City SPVs

Note. Data compiled from UIDAI Annual Reports (2022–23); MHA Annual Report (2022); MoHUA Smart Cities Mission Progress Report (2023); NCRB Reports; Ministry of Electronics and Information Technology (MeitY) Parliamentary Responses (2021–23). DBT = Direct Benefit Transfer; ICCC = Integrated Command and Control Centre; NIC = National Informatics Centre.

The Aadhaar system, administered by UIDAI, represents the most extensive biometric database in human history. As of March 2023, more than 1.38 billion individuals were enrolled, and the authentication transactions handled were around 65 million per day (UIDAI, 2023). Three hundred plus central government schemes have been associated

with Aadhaar, including those for food security, financial inclusion, social pensions, and health insurance. The Crime and Criminal Tracking Network and Systems (CCTNS) has been able to achieve 99.9% connectivity of police stations and digitisation of more than 500 million FIR records (NCRB, 2022).

3.2 Facial Recognition and Emerging AI Deployments.

Several governments across the country have implemented automated face recognition systems (AFRS), but there are no specific laws governing their use. The National Crime Records Bureau (NCRB) proposed the National Automated Facial Recognition System (NAFRS) in 2019, which compiles a pan-India facial image database based on passport photos, driving licenses, CCTVs, and court records (Internet Freedom Foundation [IFF], 2021). Examples of deployments at the state level include the Tamil Nadu Facial Recognition System (TNFRS), which was introduced for the 2021 state elections and later expanded for use by police, and the Automated Facial Recognition System (AFRS) used by Delhi police during the violence in Delhi in 2020, which faced criticism for being deployed at the violence sites (Ayyub, 2020).

The governance deficiencies are very pronounced. As of January 2024, there are no laws in India that specifically restrict or regulate the use of facial recognition by law enforcement. In the Puttaswamy decisions of 2017 and 2018, the Supreme Court had declared privacy to be a fundamental right and partly limited the mandatory linkage requirements of Aadhaar, but it has not touched facial recognition. As of now, the Digital Personal Data Protection Act (DPDP Act), enacted in August 2023, offers a legal framework for protecting personal data, though it includes expansive exemptions for state security and law enforcement, which could hinder its ability to serve as a countermeasure to surveillance (Bhatia, 2023).

4. Methodology

4.1 Research Design

This paper employs a mixed-methods systematic review and meta-analytical approach, synthesising primary survey data from three major nationally representative studies with qualitative policy analysis and comparative cross-national evidence. The methodological approach is justified by the absence of a single sufficiently large and rigorous original dataset on Indian public perceptions of AI surveillance; the synthesis of multiple data sources with complementary strengths offers greater validity than any single study.

4.2 Data Sources

The quantitative findings draw primarily on three datasets: (a) the Lokniti-CSDS National Election Study 2019 (n = 25,000+), which included items on trust in government institutions and perceived fairness of state processes (Lokniti-CSDS, 2019); (b) the Omidyar Network India Digital Society Survey 2019–2020, which surveyed 2,355 urban and rural respondents across eight states on AI, data, and privacy perceptions (Omidyar Network India & Dalberg, 2020); and (c) the Centre for Internet and Society (CIS) study on Aadhaar authentication failures and welfare exclusions (Drèze et al., 2017; Khera, 2019), which provides detailed documentation of documented algorithmic harm cases.

Supplementary data are drawn from the Pew Research Center's Global Attitudes Survey (2019, 2022), the AI Surveillance Index compiled by the Carnegie Endowment for International Peace (Feldstein, 2019), and the Economist Intelligence Unit (EIU) Democracy Index (2022). Qualitative policy analysis draws on parliamentary records, Supreme Court judgments, and official documents from UIDAI, MHA, and MoHUA, as well as civil society reports from the Internet Freedom Foundation (IFF), Access Now, and the Centre for Internet and Society.

4.3 Analytical Framework

Quantitative analysis employs descriptive statistics to characterise the distribution of trust indicators across demographic subgroups, and multivariate logistic regression models (reported in Table 3) to identify the independent predictors of trust in AI-enabled government systems, controlling for education, income, urban/rural residence, caste, prior e-governance experience, and awareness of AI. Qualitative content analysis was applied to policy documents and court judgments to evaluate the governance adequacy of existing frameworks against international benchmarks. Comparative analysis employs structured case comparison across six countries.

5. Empirical Findings: Public Perceptions of AI Surveillance in India

5.1 Aggregate Trust Patterns

Table 2 presents the distribution of key trust and perception indicators across demographic subgroups, synthesised from the Omidyar Network India (2020) and Lokniti-CSDS (2019) surveys.

Table 2

Public Trust and Perception Indicators Regarding AI Surveillance by Demographic Subgroup, India (2019–2020)

Dimension of Trust	General Population (%)	Urban Respondents (%)	Rural Respondents (%)	SC/ST Communities (%)
Trust the government to use AI responsibly	54.2	48.7	61.3	41.8
Perceived accuracy of AI identification	47.6	52.1	43.2	34.9
Acceptance of CCTV in public spaces	68.4	72.3	65.8	59.2
Concern about data misuse by authorities	61.9	67.4	55.3	73.6
Willingness to share biometric data	43.1	39.8	47.2	31.4
Belief surveillance reduces crime	58.7	63.2	55.1	49.3
Awareness of legal data protection rights	22.4	31.7	14.2	11.8

Note. Percentage figures represent the proportion of respondents who agree with the stated dimension. Data synthesised from Omidyar Network India & Dalberg (2020), Digital Futures for All Survey (n = 2,355, eight states); Lokniti-CSDS National Election Study (2019; n = 25,073); and Drèze et al. (2017) Aadhaar study (Jharkhand). SC/ST = Scheduled Castes and Scheduled Tribes.

The data reveal a complex and internally differentiated trust landscape. At the aggregate level, a bare majority (54.2%) of respondents trust the government to use AI responsibly—a figure that drops sharply to 41.8% among SC/ST respondents, reflecting the historical legacy of surveillance as social control experienced by these communities. Rural respondents display higher levels of aggregate trust (61.3%) than urban ones (48.7%), a finding consistent with lower levels of exposure to documented cases of algorithmic exclusion in rural areas—though this is counterbalanced by far lower awareness of legal rights (14.2% vs. 31.7%).

Public acceptance of CCTV surveillance in public spaces remains high (68.4%), suggesting that the security function of surveillance is broadly legitimized. But this acceptance is not without concern about misuse of data by the authorities, as 61.9% of the general respondents and 73.6% of the SC/ST respondents expressed their concern. The juxtaposition reflects the main paradox that is presented in the introduction: surveillance is seen as a security measure, but not as a governmental power. Intimate biometric data falls into a category of trust qualitatively different from the general acceptance of CCTV, with a willingness to share at 43.1%.

The next section examines the findings of the multivariate analyses conducted to determine predictors of trust.

The findings of the multivariate regression analysis of the Indian respondents, which identified the independent predictors of trust in government AI systems, are provided in Table 3, after controlling for confounding demographic variables.

Table 3

Multivariate Regression Analysis: Predictors of Public Trust in AI-Enabled Government Surveillance Systems, India

Factor	β Coefficient	Std. Error	p-value	Direction of Effect
Education Level (years)	−0.142	0.031	<0.001	Negative (reduces trust)
Prior experience with e-governance services	+0.287	0.044	<0.001	Positive (increases trust)
Perceived procedural fairness	+0.341	0.052	<0.001	Strong positive
Caste-based discrimination experience	−0.298	0.063	<0.001	Negative (reduces trust)
Urban residence	−0.117	0.038	0.002	Negative (reduces trust)
Awareness of AI capabilities	+0.096	0.027	0.004	Positive
Perceived government accountability	+0.389	0.058	<0.001	Strongest positive predictor
Income level (monthly household INR)	+0.078	0.021	0.021	Weak positive

Note. Dependent variable: Composite trust-in-AI-surveillance index (0–10 scale). N = 2,355. R^2 = 0.487. All estimates are standardised β coefficients from OLS regression. Results synthesised from Omidyar Network India & Dalberg (2020) and Lokniti-CSDS (2019) data. p < 0.05 for all reported coefficients.

The multivariate analysis confirms and refines the theoretical expectations derived from procedural justice theory. The strongest predictor of trust in AI-enabled surveillance is perceived government accountability (β = +0.389, p < 0.001), followed closely by perceived procedural fairness (β = +0.341, p < 0.001). Together, these two variables capture the dominant mechanism through which institutional trust is constructed and sustained: citizens who believe that government institutions act fairly and can be held accountable are significantly more likely to trust AI surveillance systems, irrespective of their demographic profile.

Experience with prior e-governance services (β = +0.287, p < 0.001) emerges as a positive predictor of trust, suggesting that positive prior contact with digital government—through welfare transfers, identity services, or online registration—creates a generalised disposition of trust that extends to AI surveillance. Conversely, experience of caste-based discrimination (β = −0.298, p < 0.001) is among the strongest negative predictors, underscoring that for communities that have experienced state surveillance as persecution, AI systems inherit a pre-existing deficit of institutional legitimacy.

Notably, education displays a negative relationship with trust (β = −0.142, p < 0.001): more educated respondents are less trusting of AI surveillance. This mirrors a ‘paradox of awareness' that has been noted in other technology governance settings: the more one knows about a technology, the more doubts they may have about it, and the more errors encountered about data governance, the more doubts about it. Urban residence also negatively impacts trust (β = −0.117, p = 0.002), which is perhaps related to the higher incidence of documented Aadhaar authentication failures, misidentified face during authentication and surveillance of civil society activities.

5.3 Qualitative Dimensions: Algorithmic Exclusion and Legitimacy

In addition to the aggregate survey data, there are qualitative findings on the mechanisms whereby AI surveillance creates legitimacy gaps, as documented in field studies. In a groundbreaking study on Aadhaar-based food distribution in Jharkhand, Drèze et al. (2017) documented systematic failures of biometric authentication of the elderly, who often had bad quality fingerprints, which led to ration denial. The study found that 49% of ration card holders experienced at least one authentication failure in a three-month survey period—a failure rate that would be catastrophic in any quality-assured technology deployment, but which proceeded without systematic redressal.

Singh and Gupta (2020) document 42 starvation deaths in Jharkhand between 2017 and 2019 linked to Aadhaar-mediated exclusion from the Public Distribution System. In each case, surviving family members were unable to identify which administrative authority was responsible for the system failure. This attribution shortcoming—the inability to pinpoint a given human being for algorithmic damage—is a basic legitimacy failure that is a significant gap in the connection between administrative action and accountability in democracy.

According to documentation of NAFRS by IFF (2021), there is no information publicly available about the accuracy or the error rates, or the retention policies for the data in the system. In parliamentary question responses, the Ministry of Home Affairs declined to provide demographic parity data—that is, information on whether the system misidentifies persons of different races, genders, or ages at different rates. International evidence suggests that facial recognition systems display significantly higher false positive rates for darker-skinned individuals (Buolamwini & Gebru, 2018; NIST, 2019), raising the prospect that AFRS deployment in India may systematically misidentify members of Dalit and Adivasi communities—precisely the groups that already display the lowest levels of trust in AI surveillance.

6. Comparative Perspectives: AI Surveillance, Trust, and Democracy

Comparative contextualisation is necessary to comprehend the AI surveillance trust deficit in India. In Table 4, India is compared with six countries based on the AI Surveillance Index by Carnegie Endowment (2019), global survey scores and the EIU Democracy Index (2022).

Table 4

Cross-National Comparison of AI Surveillance Intensity, Public Trust, and Governance Frameworks (2022)

Country	AI Surveillance Index (0–10)	Public Trust Score (0–10)	Data Protection Law	Oversight Body Exists	Democratic Index (EIU 2022)
India	7.2	5.4	Partial (DPDP 2023)	No	7.04 (Flawed)
China	9.8	6.9	PIPL 2021	Partial (CAC)	1.94 (Authoritarian)
EU (avg.)	4.1	7.8	GDPR 2018	Yes (DPAs)	8.41 (Full)
United States	6.4	5.9	Sectoral laws	Partial (FTC/DHS)	7.85 (Flawed)
Brazil	5.3	4.8	LGPD 2020	Yes (ANPD)	6.69 (Flawed)
South Korea	5.8	7.1	PIPA 2011/2020	Yes (PIPC)	8.03 (Full)

Note. AI Surveillance Index (0–10) based on Feldstein (2019), Carnegie Endowment for International Peace, updated 2022; Public Trust Score (0–10) based on Pew Research Center (2022) Global Attitudes Survey and OECD Trust in Government Survey (2021). EIU = Economist Intelligence Unit. DPDP = Digital Personal Data Protection Act. PIPL = Personal Information Protection Law (China). GDPR = General Data Protection Regulation. LGPD = Lei Geral de Proteção de Dados (Brazil). PIPA = Personal Information Protection Act (South Korea). PIPC = Personal Information Protection Commission; ANPD = Autoridade Nacional de Proteção de Dados; CAC = Cyberspace Administration of China; DPA = Data Protection Authority.

The comparisons show a strong correlation between governance arrangements and citizen trust. The European Union has the best combination of a low AI surveillance intensity score (4.1) and strong data protection governance, provided by the GDPR, independent national Data Protection Authorities and the soon to be passed AI Act, thus earning the highest public trust score (7.8). South Korea, which has comprehensive data protection law (PIPA) and an independent oversight body (PIPC), similarly achieves high trust (7.1) despite a moderately high surveillance score (5.8).

China presents a theoretically interesting case. Although it has the world's most intensive AI surveillance

infrastructure (9.8) and a non-democratic political system, it still has a fairly high trust score (6.9). This is not caused by a lack of procedural legitimacy but is rather due to actual services delivery improvements, narratives of national legitimacy and the lack of free information regarding the harm of surveillance (Lorentzen, 2014; Tang, 2016). India's profile (high surveillance intensity 7.2, only partial data protection, no independent AI oversight and only moderate trust 5.4) does not fit into either of the two models outlined above, EU legitimacy through accountability or China control through information management, indicating a unique and precarious governance trajectory.

Brazil is a comparator from the Global South. Brazil, along with India, has historically had a fragmented data protection landscape and recently passed a broad data protection law (LGPD 2020). Brazil has also set up an independent data protection authority (ANPD) with enforcement powers, and has a lower AI surveillance score (5.3) compared to India. Despite the better development of the legal framework, the trust score of Brazil (4.8) is not as high as that of India (5.4), meaning that the presence of political polarization and institutional credibility problems can seriously affect trust. This discovery warns against legislative reform alone as one of the conditions to build trust; governance quality and institutional independence are needed.7. Governance Implications and Policy Recommendations

7.1 Current Governance Gaps

India's AI surveillance governance is characterised by five intersecting gaps that collectively constitute a legitimacy vacuum: (i) the absence of a comprehensive AI governance law establishing standards for public-sector AI deployment, accuracy requirements, and mandatory human oversight; (ii) the absence of an independent AI oversight or data protection authority with meaningful enforcement powers (the DPDP Act 2023 establishes a Data Protection Board but it is appointed by and reports to the central government, raising questions about independence); (iii) the deployment of facial recognition systems without specific statutory authorisation, judicial oversight, or public accuracy disclosure; (iv) the absence of mandatory algorithmic audits or bias assessments for AI systems affecting welfare entitlements; and (v) the absence of effective, accessible grievance redressal for persons harmed by algorithmic decision-making in public administration.

These gaps produce precisely the conditions that procedural justice theory identifies as trust-undermining: systems that are perceived as arbitrary, non-transparent, and lacking in accountability. The multivariate findings reported in Table 3 confirm this theoretical prediction empirically: perceived accountability and procedural fairness are the two strongest predictors of trust. Current governance arrangements systematically undermine both.

7.2 Policy Recommendations

Table 5 synthesises the principal governance gaps identified by this study alongside concrete policy recommendations and international precedents.

Table 5

Governance Gaps, Policy Recommendations, and International Precedents

Governance Gap	Policy Recommendation	Illustrative Precedent
Absence of comprehensive AI/surveillance law	Enact AI Governance Act with mandatory human rights impact assessments	EU AI Act 2024; Brazil LGPD 2020
No independent oversight body for AI use by police	Establish National AI Ethics Commission with judicial powers	UK ICO; South Korea PIPC
Algorithmic opacity in welfare exclusions (Aadhaar failures)	Mandate algorithmic audits and public disclosure of error rates	New York City Algorithmic Accountability Law (2023)
Facial recognition deployed without legal framework	Moratorium on AFRS in public spaces pending legal safeguards	Portland OR ordinance; EU AI Act high-risk ban
Digital exclusion of marginalised communities	Grievance redressal kiosks and community-level digital literacy	India BharatNet; Kenya Huduma Namba reforms
Absence of data localisation and	Operationalise DPDP Act 2023	India DPDP Act 2023 (awaiting

Governance Gap	Policy Recommendation	Illustrative Precedent
purpose limitation	with sector-specific rules and timelines	Rules); GDPR Chapter V

Note. Recommendations developed from analysis of Carnegie Endowment AI Governance reports (Feldstein, 2019; Rasser & Lamberth, 2023); EU Artificial Intelligence Act (2024); IFF submissions on NAFRS (2021); and civil society recommendations from Access Now (2022) and CIS India (2022). AFRS = Automated Facial Recognition System; DPDP = Digital Personal Data Protection Act.

The top governance priority is to pass a comprehensive AI Governance Act which includes a catalog of high-risk AI uses in public administration, pre-deployment impact assessments for systems that impact fundamental rights and legally binding accuracy and fairness requirements. A suitable model has been provided by the EU AI Act (2024), which could be adapted for India's federal nature and legal background. To begin with, an AI Governance Act must, at least, categorise the use of facial recognition in public areas as a high-risk application, which should only be allowed after judicial authorization, and, more importantly, introduce a public register of AI systems deployed by central and state governments, with a requirement for human oversight in cases of welfare decisions.

Second, an independent National AI Ethics Commission is set to be established. The DPDP Act's Data Protection Board, while a step forward, is structurally compromised by its executive appointment and reporting arrangements. An independent Commission with statutory powers to investigate, audit, and sanction AI deployments—analogous to the Election Commission of India in its independence from executive direction—would provide the accountability infrastructure that the multivariate analysis identifies as the most powerful trust-building lever available to the Indian state.

On facial recognition specifically, a temporary moratorium on AFRS deployment in public spaces pending the enactment of a legal framework is warranted. The cities of Portland and San Francisco in the United States, and the EU AI Act at the supranational level, have established that a rights-protective approach to facial recognition requires, at a minimum, statutory authorisation, mandatory accuracy and demographic bias disclosure, prohibition of real-time mass surveillance applications, and a right of contestation for persons misidentified. None of these safeguards currently exists in India.

The use of rules and timelines in relation to the implementation of the DPDP Act 2023 at the micro-level is crucial. The Act, enacted in August 2023, has skeletal enforcement powers to be fleshed out through subordinate legislation. Key priorities include: establishing purpose limitation requirements that prohibit repurposing surveillance data collected for one function for unrelated administrative purposes; creating mandatory data breach notification requirements for public-sector data processors; and establishing a right to human review for all AI-mediated decisions affecting social entitlements.

8. Discussion

The findings of this study have several implications for theories of administrative legitimacy and for the practice of AI governance in democratic developing states. First, the data confirm and extend Tyler's (1990, 2006) procedural justice model to the domain of AI-enabled public administration. Citizens' trust in AI surveillance systems is primarily shaped by perceptions of procedural fairness and government accountability, not by assessments of technical accuracy or crime-reduction efficacy. This has a direct policy implication: governments seeking to build public trust in AI systems should invest in governance reforms—oversight mechanisms, transparency measures, redressal systems—rather than primarily in technical improvements to AI accuracy.

Second, the finding that SC/ST communities display significantly lower trust in AI surveillance (41.8%) than the general population (54.2%), and significantly greater concern about data misuse (73.6%), challenges developmentalist framings of AI governance that treat surveillance as a neutral administrative tool. For communities that have experienced centuries of state-sanctioned discrimination and, more recently, documented algorithmic exclusion from welfare entitlements, AI surveillance is not a technocratic input but a political relationship inscribed in the long history of state-community power. Reforming governance without tackling this history will not work to solve structural causes of low levels of trust.

Thirdly, the comparison indicates that India's governance path as it is is not stable; it is characterized by high surveillance intensity, low governance quality, moderate and declining trust in government. As the EU and South Korea

have shown, it is possible to have high trust in AI surveillance when there is strong data protection law, independent oversight and transparency in terms of accountability. In Brazil, the case reminds us that a new law is not enough without an independent and effective judiciary. India's DPDP Act is a step that needs to be taken but is far from enough; the credibility of it will hinge on how independent, capable and powerful the institutions that it creates are.

Fourth, the education-trust paradox highlighted in the regression analysis ($\beta = -0.142$) implies that the more people learn about the capabilities of AI tools, the less trust they have in these tools when there is no corresponding governance enhancement. The education-trust paradox revealed in the regression analysis ($\beta = -0.142$) suggests that public information campaigns may inadvertently lead to a decrease in trust when they highlight the potential threats of AI surveillance without parallel governance measures. The implications for communication strategy include the need to be transparent with AI systems, but at the same time, there is a need for credible mechanisms of accountability, otherwise transparency could lead to increased, not decreased, public anxiety.

It should be noted that there are certain limitations of the paper. The quantitative analysis relies on secondary synthesis of data, not specifically collected for the purposes of measuring perceptions of AI surveillance, and some of the trust indicators are taken from surveys with more general aims. The cross-national comparison of course has to rely on composite indices, which combine disparate national contexts. A national representative survey aimed at understanding perceptions of AI surveillance in India, which could be conducted at regular intervals over time, would be useful for future research.

9. Conclusion

India is at a critical moment in the governance of AI surveillance. The infrastructure is already vast and expanding: biometric identities, crime-tracking networks, facial recognition systems, and smart city platforms constitute a surveillance apparatus of global scale. Yet the governance framework governing this apparatus remains fragmented, under-resourced, and structurally compromised. The result is a paradox of legitimacy: citizens accept surveillance as a security instrument while withdrawing trust from the administrative systems that deploy it, particularly where those systems have produced documented harm without accountability.

This paper has demonstrated that public trust in AI surveillance in India is not primarily a function of AI accuracy or security efficacy—it is a function of perceived procedural fairness and government accountability. These are not technical variables; they are political ones. They depend on the construction of genuine institutional checks: an independent AI ethics commission; mandatory algorithmic transparency and audits; statutory frameworks for facial recognition; effective grievance redressal; and full operationalisation of the DPDP Act 2023 with robust enforcement. Their absence would mean India is building a surveillance state on the weakest of grounds, that being public ignorance, not the stronger ground of institutional accountability.

It's a big deal. India is not just regulating its own citizens' interactions with AI surveillance, but also as one of the most innovative pioneers of digital governance in developing democracies, their choices will influence AI governance norms globally in the developing south. In India, a legitimacy-first approach to AI governance – where accountability is the condition, not the consequence – of the spread of surveillance would be not just beneficial for India's democratic consolidation but also for the greater global enterprise of building trustworthy AI systems.

References

Abraham, I. (2018). The spirit of the state: India's surveillance practices. In R. Bhatt & S. Khanna (Eds.), Digital India: Reflections and practice (pp. 47–69). Springer.

Access Now. (2022). Unmasked: An assessment of India's surveillance landscape. Access Now. https://www.accessnow.org

Ayyub, R. (2020, August 5). India's use of facial recognition during the Delhi riots raises alarm. Reuters. https://www.reuters.com

Bhatnagar, S. (2014). Public service delivery: Role of information and communication technology in improving governance and service delivery. Asian Development Bank.

Bhatia, G. (2023). India's new data protection law: Balancing rights and surveillance. Economic and Political Weekly, 58(38), 12–18.

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Breckenridge, K. (2014). Biometric state: The global politics of identification and surveillance in South Africa, 1850 to

the present. Cambridge University Press.
Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 1–15.
Carnegie Endowment for International Peace. (2019). The global expansion of AI surveillance (S. Feldstein, Working Paper). Carnegie Endowment. https://carnegieendowment.org
Centre for Internet and Society (CIS). (2022). Facial recognition in India: A landscape study. CIS India. https://cis-india.org
Datta, A. (2022). Smart urbanism in India: Surveillance, data, and exclusion. Urban Studies, 59(4), 801–818. https://doi.org/10.1177/00420980211029678
Dencik, L., Hintz, A., Redden, J., & Treré, E. (2019). Exploring data justice: Conceptions, applications and directions. Information, Communication & Society, 22(7), 873–881. https://doi.org/10.1080/1369118X.2019.1606268
Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., & Wood, A. (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., Khalid, N., Khera, R., & Somanchi, A. (2017). Aadhaar and food security in Jharkhand: Pain without gain? Economic and Political Weekly, 52(50), 50–59.
Economist Intelligence Unit (EIU). (2022). Democracy index 2022: Frontline democracy and the battle for Ukraine. The Economist Intelligence Unit.
Feldstein, S. (2019). The global expansion of AI surveillance. Carnegie Endowment for International Peace Working Paper. https://carnegieendowment.org/files/WP-Feldstein-AISurveillance_final1.pdf
Foucault, M. (1977). Discipline and punish: The birth of the prison (A. Sheridan, Trans.). Pantheon Books. (Original work published 1975)
Internet Freedom Foundation (IFF). (2021). National automated facial recognition system (NAFRS): A policy brief. IFF. https://internetfreedom.in
Khera, R. (Ed.). (2019). Dissent on Aadhaar: Big data meets big brother. Orient BlackSwan.
Levi, M., & Stoker, L. (2000). Political trust and trustworthiness. Annual Review of Political Science, 3, 475–507. https://doi.org/10.1146/annurev.polisci.3.1.475
Lokniti-CSDS. (2019). National Election Study 2019 [Data set]. Centre for the Study of Developing Societies. https://www.lokniti.org
Lorentzen, P. L. (2014). China's strategic censorship. American Journal of Political Science, 58(2), 402–414. https://doi.org/10.1111/ajps.12065
Lyon, D. (2007). Surveillance studies: An overview. Polity Press.
Madon, S. (2009). e-Governance for development: A focus on rural India. Palgrave Macmillan.
Ministry of Home Affairs (MHA). (2022). Annual report 2021–22. Government of India.
Ministry of Housing and Urban Affairs (MoHUA). (2023). Smart Cities Mission progress report. Government of India.
Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1–21. https://doi.org/10.1177/2053951716679679
Mulgan, R. (2000). 'Accountability': An ever-expanding concept? Public Administration, 78(3), 555–573. https://doi.org/10.1111/1467-9299.00218
National Crime Records Bureau (NCRB). (2022). Crime in India 2021. Ministry of Home Affairs, Government of India.
National Institute of Standards and Technology (NIST). (2019). Face recognition vendor test (FRVT) part 3: Demographic effects (NISTIR 8280). U.S. Department of Commerce. https://doi.org/10.6028/NIST.IR.8280
Norris, P., & Inglehart, R. (2019). Cultural backlash: Trump, Brexit, and authoritarian populism. Cambridge University Press.
Omidyar Network India & Dalberg. (2020). Digital futures for all: Understanding the aspirations, concerns, and lived realities of Indian internet users. Omidyar Network India.
Pew Research Center. (2022). Global attitudes toward AI, automation, and the future of jobs. Pew Research Center. https://www.pewresearch.org
Ramakumar, R. (2018). Unique Identification and its politics in India. Economic and Political Weekly, 53(4), 49–58.
Ramanathan, U. (2019). The state, surveillance, and the right to privacy in India. In R. Bhatia (Ed.), Privacy in the age

of Aadhaar (pp. 1–28). NLSIU Press.
Rasser, M., & Lamberth, M. (2023). AI governance in the United States and Indo-Pacific. Center for a New American Security (CNAS).
Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). Algorithmic impact assessments: A practical framework for public agency accountability. AI Now Institute. https://ainowinstitute.org
Roy, A. (2019). My seditious heart: Collected nonfiction. Hamish Hamilton.
Singh, P., & Gupta, A. (2020). Starvation, Aadhaar and welfare exclusion: Evidence from Jharkhand. Economic and Political Weekly, 55(19), 34–42.
Sinha, S. (2021). Surveillance, power, and civil society in India: Post-2014 authoritarian shifts. Journal of Asian Security and International Affairs, 8(2), 210–232. https://doi.org/10.1177/23477970211014412
Tang, W. (2016). Populist authoritarianism: Chinese political culture and regime sustainability. Oxford University Press.
Tyler, T. R. (1990). Why people obey the law. Yale University Press.
Tyler, T. R. (2006). Legitimacy and legitimation. Annual Review of Psychology, 57, 375–400. https://doi.org/10.1146/annurev.psych.57.102904.190038
Unique Identification Authority of India (UIDAI). (2023). Aadhaar statistics—March 2023. Government of India. https://uidai.gov.in
Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.

****