SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

Algorithmic Accountability and the Rise of AI-Driven Governance in Public Administration

Yachika, Department of Political Science, Faculty of Humanities, Baba Mastnath University, Rohtak, Haryana, India

Abstract

The diffusion of artificial intelligence (AI) technologies into public administration has fundamentally altered how governments make, justify, and review decisions that affect citizens' lives. From welfare eligibility screening to tax scrutiny selection, from judicial case management to law-enforcement risk assessment, algorithmic systems now mediate a growing share of state power. This shift raises urgent questions about algorithmic accountability — the capacity of citizens, courts, legislatures, and oversight bodies to understand, contest, and correct decisions made or substantially shaped by automated systems. This paper undertakes an analytical review of the theoretical foundations, global regulatory trajectories, and empirical experience of AI-driven governance, with particular attention to the Indian administrative context. Using a qualitative, document-based analytical methodology, the paper examines landmark international scholarship on algorithmic accountability alongside Indian case studies including the Aadhaar biometric identification system, the Faceless Assessment Scheme of the Income Tax Department, the Supreme Court's SUPACE and SUVAS initiatives, and NITI Aayog's Responsible AI policy architecture. The analysis finds that while algorithmic governance offers demonstrable gains in efficiency, consistency, and corruption reduction, it simultaneously produces new forms of opacity, exclusion, and diffused responsibility that existing administrative law doctrines are ill-equipped to address. The paper argues for a hybrid accountability model that combines technical auditability, human-in-the-loop safeguards, participatory design, and statutory rights to explanation, and it situates India's emerging regulatory architecture — particularly the Digital Personal Data Protection Act, 2023 and NITI Aayog's Responsible AI framework — within this global conversation. The paper concludes with policy recommendations for strengthening accountability without foreclosing the developmental promise of AI in governance.

Keywords: *algorithmic accountability, artificial intelligence, public administration, digital governance, India, Aadhaar, administrative law, transparency, NITI Aayog, algorithmic bias*

1. Introduction

Public administration has always relied on rules, discretion, and hierarchy to convert political decisions into administrative outcomes. For most of the twentieth century, that conversion process was substantially human: a clerk verified eligibility, an officer exercised judgment, and a supervisor reviewed the outcome. The last decade has witnessed a structural transformation of this process. Artificial intelligence and machine learning systems are now embedded at

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

multiple points in the administrative decision chain — screening applicants, flagging anomalies, ranking priorities, drafting orders, and in some jurisdictions, issuing decisions with minimal human intervention (Berryhill et al., 2019). Governments describe this transformation using the language of efficiency, objectivity, and modernization; scholars increasingly describe it using the language of opacity, discretion migration, and accountability deficit (Busuioc, 2021; Wirtz et al., 2020).

Algorithmic accountability, as a field of inquiry, asks a deceptively simple question: when a machine-assisted or machine-made decision harms a citizen, who is answerable, through what mechanism, and with what remedy available? The question is deceptively simple because the answer implicates multiple, overlapping domains — administrative law, data protection law, computer science, public administration theory, and constitutional rights. A welfare applicant denied benefits because a biometric authentication system failed to match a worn fingerprint does not experience an abstract "governance challenge"; she experiences hunger, humiliation, and an opaque bureaucratic wall that offers no clear point of appeal (Nickled and Dimed, 2025; Policy Circle, 2026). It is this gap between the technical elegance of AI systems and the lived experience of citizens subject to them that motivates the present inquiry.

This paper is animated by three interlocking objectives. First, it seeks to map the theoretical evolution of algorithmic accountability as a concept, tracing its roots in administrative law's due process tradition and its more recent articulation in computer science and information ethics literature. Second, it seeks to document, with empirical specificity, how AI-driven governance has actually unfolded in the Indian administrative state — a context that is analytically valuable precisely because India combines an ambitious digital public infrastructure agenda with deep socio-economic inequality, making the stakes of algorithmic error unusually visible. Third, it seeks to evaluate the adequacy of India's emerging legal and policy architecture, principally the Digital Personal Data Protection Act, 2023 (DPDP Act) and NITI Aayog's Responsible AI framework, against the accountability gaps identified in the global literature.

The urgency of this inquiry is not speculative. India operates what is arguably the world's largest biometric identification system, Aadhaar, covering over 1.3 billion residents and used as an authentication gateway for subsidised food rations, employment guarantee wages, pension disbursement, and banking access (Nickled and Dimed, 2025). India's Income Tax Department has, since 2019, used algorithmic case-allocation to conduct "faceless" scrutiny assessments affecting hundreds of thousands of taxpayers annually (ISAS, 2020). The Supreme Court of India has institutionalised AI-assisted legal research through its SUPACE portal and machine-translation through SUVAS (IndiaAI, 2021). Each of these systems illustrates a different facet of algorithmic governance — identity verification, risk-based selection, and decision-support — and each raises distinct accountability questions that this paper examines in turn.

The remainder of the paper is structured as follows. Section 2 reviews the international literature on algorithmic accountability in public administration. Section 3 develops a theoretical framework synthesising administrative law and information ethics perspectives. Section 4 outlines the research methodology. Section 5 surveys the global regulatory landscape.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

Section 6 presents an in-depth analysis of AI-driven governance in India through four case studies. Section 7 examines India's legal and policy framework. Section 8 offers a critical discussion of accountability gaps and cross-national comparison. Section 9 sets out policy recommendations, and Section 10 concludes.

2. Literature Review

2.1 The Emergence of Algorithmic Accountability as a Field

The term "algorithmic accountability" gained currency in the mid-2010s, initially within computational journalism and information science, before migrating into public administration scholarship. Diakopoulos (2015) framed algorithmic accountability as the practice of reverse-engineering opaque computational systems to understand and expose the power structures they encode, drawing an analogy between investigative journalism and algorithmic auditing. This early framing emphasised transparency as an instrumental good — the means by which hidden decision logics could be rendered visible to journalists, researchers, and ultimately the public.

Ananny and Crawford (2018) subsequently offered an influential critique of transparency-centred approaches, arguing that the "transparency ideal" — the assumption that seeing a system's inner workings is sufficient to hold it accountable — is analytically limited. They contended that accountability requires attention not merely to disclosure of code or data, but to the networked, socio-technical relationships in which algorithmic systems are embedded, including the institutions that deploy them and the humans who interpret their outputs. This shift from a transparency paradigm to a socio-technical accountability paradigm has been foundational for subsequent public administration scholarship.

Wieringa (2020), in a widely cited systematic literature review synthesising over two hundred articles, proposed a definitional anchor for the field: algorithmic accountability concerns a networked account of a socio-technical algorithmic system across the stages of its lifecycle — from design and data collection through deployment and outcome review. This lifecycle framing has proven influential because it resists the temptation to treat accountability as a single moment of disclosure and instead treats it as a continuous administrative obligation.

2.2 Algorithmic Accountability in Public Administration Scholarship

Within public administration specifically, Wirtz, Weyerer and Geyer (2019) mapped the applications and challenges of AI adoption across government functions, identifying efficiency gains in service delivery alongside emergent risks in transparency and legal compliance. Their subsequent work, Wirtz, Weyerer and Sturm (2020), proposed an integrated AI governance framework for public administration explicitly organised around what they termed the "dark sides" of AI — bias, loss of human agency, and accountability diffusion — arguing that governance frameworks must be designed proactively rather than retrofitted after deployment failures.

Busuioc (2021), writing in *Public Administration Review*, provided one of the most rigorous accounts of why algorithmic systems strain conventional accountability mechanisms. She

argued that algorithms introduce accountability gaps along three dimensions: an *information gap*, because overseers often lack the technical capacity to interrogate model logic; an *explanation gap*, because even system designers may struggle to articulate why a machine-learning model produced a particular output; and a *justification gap*, because the normative standards against which algorithmic decisions should be judged remain contested. Busuioc's tripartite framework has become a standard reference point for subsequent empirical work, including comparative case analyses spanning Estonia, Singapore, the United States, and the United Kingdom, which found that algorithmic governance strengthens administrative effectiveness only when paired with robust oversight, auditability protocols, and public engagement mechanisms.

Taddeo and Floridi (2018), writing in *Nature*, situated algorithmic accountability within a broader call for anticipatory AI regulation, warning that the absence of coordinated governance frameworks risked both harmful deployment and a defensive "cyber arms race" among states competing to adopt AI without adequate safeguards — an argument with direct relevance to public administration's tendency to prioritise efficiency metrics over rights-protective design.

The Organisation for Economic Co-operation and Development's early empirical mapping (Berryhill et al., 2019) catalogued dozens of AI use cases across member and partner state bureaucracies, from chatbots to predictive maintenance to fraud detection, and observed that most governments were adopting AI faster than they were developing corresponding oversight capacity — a mismatch that recurs throughout the literature examined in this paper.

2.3 Administrative Law Foundations: Due Process and the Automated State

A distinct but complementary literature emerges from American administrative law, where scholars examined automated decision-making in welfare and benefits administration well before the current AI wave. Citron (2008), in her foundational article on "technological due process," documented how automated eligibility systems in United States welfare programs systematically violated procedural due process norms — denying claimants adequate notice, meaningful opportunity to be heard, and reasoned explanation — because system design decisions were made by engineers rather than reviewed through the deliberative processes traditionally required for rulemaking. Citron and Pasquale (2014) extended this analysis to algorithmic credit scoring, coining the phrase "the scored society" to describe a governance mode in which opaque predictive scores substitute for individualised assessment, with limited avenues for correction.

Pasquale's (2015) book-length treatment, widely cited across disciplines, characterised contemporary algorithmic governance — spanning both public and private sectors — as a "black box society," in which the reputational, financial, and security systems that govern modern life increasingly operate through proprietary logics shielded from public scrutiny by claims of trade secrecy and technical complexity. This account resonates strongly with the Indian Aadhaar experience discussed later in this paper, where the Unique Identification Authority of India's own submissions to the Supreme Court characterised its biometric-matching algorithms as proprietary and therefore not independently verifiable (The Wire, 2024).

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

2.4 Bias, Exclusion, and the Distributive Consequences of Algorithmic Governance

A parallel literature examines the distributive consequences of algorithmic decision-making, with particular attention to how automated systems can encode and amplify existing social inequalities. Eubanks (2018) documented how automated eligibility and risk-scoring systems in United States social services disproportionately burdened poor and working-class communities, terming this phenomenon the "digital poorhouse." O'Neil (2016) offered a broader indictment of what she termed "weapons of math destruction" — large-scale, opaque, and difficult-to-contest algorithmic systems that reinforce disadvantage at scale. Barocas and Selbst (2016) provided the doctrinal bridge between this empirical literature and legal remedy, analysing how "big data's disparate impact" can arise even from facially neutral algorithmic design, complicating conventional anti-discrimination frameworks that presume intentional bias.

Bovens (2007), writing more generally on public accountability theory prior to the AI wave, offered an analytical framework still widely used to evaluate algorithmic accountability: accountability as a relationship between an actor and a forum, in which the actor is obliged to explain and justify conduct, the forum can pose questions and pass judgment, and the actor may face consequences. Contemporary scholars, including Novelli, Taddeo and Floridi (2024), have extended this relational framework specifically to artificial intelligence, arguing that accountability in AI systems requires clearly identifying which actor — designer, deployer, or operator — is answerable at each stage of a system's lifecycle, echoing Wieringa's (2020) lifecycle approach.

2.5 Recent Systematic Reviews

More recent systematic reviews confirm that these accountability themes remain unresolved rather than settled. A 2026 systematic review guided by the PRISMA framework, analysing forty-five peer-reviewed publications on algorithmic accountability in the public sector, identified five recurring governance dimensions in the literature: transparency in algorithmic decision-making, explainability of AI systems, human oversight and administrative responsibility, ethical governance of AI, and — notably — the continued absence of consensus on enforceable legal standards (OSF, 2026). This finding underscores that despite a decade of scholarly attention, algorithmic accountability in public administration remains a field characterised more by diagnostic clarity than by settled remedial architecture — a gap this paper seeks to address with specific reference to the Indian administrative experience.

3. Theoretical Framework

This paper adopts a synthesis of two theoretical traditions: the relational accountability framework drawn from public administration theory, and the socio-technical systems perspective drawn from science and technology studies.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

3.1 Relational Accountability

Following Bovens (2007) and its extension by Busuioc (2021), accountability is understood here not as a static property of a system but as a relationship comprising three elements: (a) an obligation to explain and justify conduct, (b) a forum empowered to interrogate that conduct, and (c) the possibility of consequence — correction, sanction, or remedy — where conduct is found wanting. Applied to algorithmic governance, this framework generates three corresponding accountability questions for any AI-driven administrative system: Can the system's logic and outcomes be explained in terms intelligible to the affected citizen and to oversight bodies? Is there an institutional forum — judicial, legislative, or administrative — with the technical and legal capacity to interrogate that explanation? And does a meaningful remedy exist when the system errs?

3.2 The Socio-Technical Systems Perspective

Following Ananny and Crawford (2018), this paper treats algorithmic systems not as discrete technical artefacts but as socio-technical assemblages comprising code, data, institutional procedure, and human judgment. This perspective resists two common analytical errors: first, the tendency to treat algorithmic failure purely as a technical bug correctable through better engineering, and second, the tendency to treat algorithmic accountability purely as a transparency problem solvable by disclosing source code. Instead, the socio-technical lens directs attention to the full deployment context — who designed the eligibility criteria embedded in a welfare algorithm, who decided that biometric mismatch should trigger service denial rather than a manual fallback, and who bears institutional responsibility when that decision produces harm.

3.3 Integrating the Two Traditions: A Lifecycle Accountability Model

Building on Wieringa's (2020) lifecycle conception and Novelli et al.'s (2024) actor-specific accountability mapping, this paper's analytical framework identifies four accountability-relevant stages in any AI-driven governance system: **design** (what values and assumptions are embedded in system architecture and training data), **deployment** (what institutional procedures govern the system's introduction into administrative practice), **operation** (how the system functions in practice, including its interaction with frontline officials and citizens), and **review** (what mechanisms exist to detect, contest, and correct erroneous or harmful outcomes). This four-stage model is used throughout the remainder of the paper to structure the analysis of Indian case studies in Section 6.

4. Research Methodology

This study adopts a qualitative, analytical, and doctrinal research design appropriate to a socio-legal and public administration inquiry of this kind. The methodology combines three complementary approaches.

First, a **systematic conceptual literature review** was undertaken, drawing on peer-reviewed scholarship in public administration, administrative law, and information ethics journals,

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

supplemented by working papers and institutional reports from bodies such as the OECD. This review followed the accountability-dimension categorisation developed in prior systematic reviews of the field (e.g., OSF, 2026; Wieringa, 2020) to ensure comprehensive coverage of the theoretical landscape.

Second, the paper employs a **case study method**, examining four purposively selected instances of AI-driven governance in India — Aadhaar-linked welfare authentication, the Faceless Assessment Scheme in income tax administration, AI-assisted judicial infrastructure (SUPACE and SUVAS), and the NITI Aayog Responsible AI policy process. These cases were selected because, collectively, they span the principal administrative functions in which AI is deployed globally — identity verification, regulatory enforcement, adjudicative support, and policy governance — and because each has generated sufficient documented experience, including parliamentary scrutiny, judicial review, and independent research, to support analytical evaluation rather than speculative assessment.

Third, the paper undertakes a **doctrinal and policy analysis** of India's principal legal instruments bearing on algorithmic governance, namely the Digital Personal Data Protection Act, 2023, and the *Justice K.S. Puttaswamy* line of Supreme Court jurisprudence, assessed against the theoretical accountability criteria developed in Section 3.

This is not an empirical study involving primary data collection, surveys, or statistical modelling; rather, it is an analytical synthesis of secondary sources — academic literature, government policy documents, parliamentary records, judicial pronouncements, and credible investigative and research reporting — evaluated through the lifecycle accountability framework outlined above. This method is consistent with established practice in comparative public administration and socio-legal research, where the object of inquiry is institutional design and normative adequacy rather than a quantifiable causal effect.

5. The Global Regulatory Landscape of AI-Driven Governance

Before turning to the Indian context, it is analytically necessary to situate India's experience within the broader global trajectory of AI governance regulation, since India's own policy instruments are explicitly designed with reference to comparable international frameworks.

5.1 The European Union: Rights-Based Regulation

The European Union has pursued the most codified regulatory response to algorithmic decision-making, beginning with Articles 13, 15, and 22 of the General Data Protection Regulation (GDPR), which establish a qualified right to be informed of automated decision-making, a right of access to the logic involved, and safeguards against decisions based solely on automated processing that produce legal or similarly significant effects. Scholars have noted that these provisions, while significant, were designed with private-sector profiling primarily in mind, and their application to public administration remains contested, particularly because the GDPR grants member states considerable procedural autonomy to authorise algorithmic enforcement tools for public authorities, a derogation that must be balanced against due process, judicial review, and equal treatment guarantees. The subsequent EU Artificial

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

Intelligence Act extends this rights-based approach through a risk-tiered classification system, treating many public administration use cases — including eligibility determination for public benefits and law enforcement risk assessment — as "high-risk," triggering mandatory conformity assessments, human oversight requirements, and documentation obligations.

5.2 The United States: Administrative Law by Adaptation

The United States has taken a comparatively fragmented approach, relying on existing administrative law doctrine adapted to algorithmic contexts rather than comprehensive AI-specific legislation. A major study commissioned by the Administrative Conference of the United States examined the use of AI tools by the Social Security Administration in adjudicating disability benefits claims and by the Securities and Exchange Commission in targeting enforcement, finding that litigation challenging these systems has tended to centre not on the abstract opacity concerns emphasised in academic literature, but on concrete procedural due process failures — inadequate notice, unrealistically short response windows, and deficient explanation — echoing Citron's (2008) earlier "technological due process" critique. This finding is significant for the present paper because it suggests that, regardless of jurisdiction, the most legally cognisable harms from algorithmic governance tend to arise not from the sophistication of the algorithm itself but from the erosion of basic procedural fairness that automation enables.

5.3 The OECD Framework and Global Coordination

The OECD's Recommendation of the Council on Artificial Intelligence, first adopted in 2019 and periodically updated, has functioned as an influential soft-law reference point, articulating principles of inclusive growth, human-centred values, transparency, robustness, and accountability that numerous national strategies — including India's — explicitly draw upon. India became a founding member of the Global Partnership on Artificial Intelligence in 2020, built around shared commitment to the OECD Recommendation, and subsequently chaired the Global Partnership, hosting the Global India AI Summit in 2024, reflecting India's aspiration to be a norm-shaping rather than merely norm-receiving participant in global AI governance discourse.

5.4 The Common Thread: Efficiency-Accountability Tension

Across these divergent regulatory approaches, a common structural tension recurs: the same features that make algorithmic systems administratively attractive — speed, consistency, scale, and reduced discretion — are the features that most complicate conventional accountability mechanisms, which were designed around individualised, human, and reviewable decision-making. This tension is not resolved by any of the frameworks reviewed above; it is, at best, managed through layered safeguards. It is against this backdrop that India's own regulatory and administrative experience must be evaluated.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

6. AI-Driven Governance in India: Case Studies

India presents an unusually rich empirical setting for studying algorithmic accountability because it combines an ambitious "Digital India" agenda, a large and diverse population, significant socio-economic inequality, and a constitutional framework with a strong fundamental rights tradition. This section examines four case studies using the four-stage lifecycle accountability model developed in Section 3.

6.1 Case Study I: Aadhaar and Algorithmic Exclusion in Welfare Delivery

6.1.1 Background and Design

Aadhaar, administered by the Unique Identification Authority of India (UIDAI), is a twelve-digit biometric identity number linked to fingerprint, iris, and demographic data, covering over 1.3 billion residents. Since the mid-2010s, Aadhaar-based biometric authentication has been progressively mandated as a gateway to accessing the Public Distribution System (subsidised food rations), the Mahatma Gandhi National Rural Employment Guarantee Scheme (rural employment wages), pension disbursement, and an expanding range of banking and welfare services (Nickled and Dimed, 2025). The underlying design rationale, consistent with the efficiency logic identified in Section 5.4, was to eliminate "ghost" beneficiaries and leakage in welfare distribution through algorithmic, biometric verification rather than paper-based, discretion-laden local administration.

6.1.2 Deployment and Operation: Documented Exclusion

The deployment stage of Aadhaar-linked authentication has generated substantial documented harm, particularly for the very populations the welfare schemes were designed to serve. India's Parliamentary Public Accounts Committee has heard evidence of persistently high rates of biometric verification failure, with committee members across party lines warning that faulty fingerprint and iris scans are excluding eligible beneficiaries from subsidised rations and employment guarantee wages (Biometric Update, 2025). Independent analysis indicates that of roughly 312 million Aadhaar-based biometric authentication attempts conducted monthly for welfare, banking, and public services, approximately 20.3 million fail — a failure rate of roughly 6.5 percent that has remained largely unchanged for over a decade, despite controlled laboratory accuracy rates of 98 to 99 percent for fingerprint and iris matching (Policy Circle, 2026). This gap between controlled testing accuracy and real-world operational accuracy is analytically significant: it demonstrates that algorithmic systems validated under idealised conditions can nonetheless produce systemic exclusion when deployed at population scale in real administrative environments.

Crucially, these failures are not randomly distributed. Manual labourers — agricultural workers, construction workers, and domestic workers — experience disproportionately higher failure rates because repetitive physical labour degrades fingerprint ridges over time, while elderly citizens experience iris pattern changes with age (Policy Circle, 2026). This produces what can be termed a *structural contradiction*: the demographic groups most dependent on welfare entitlements are precisely those most likely to be algorithmically misidentified and

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

therefore excluded. This finding directly illustrates the disparate-impact concern articulated by Barocas and Selbst (2016) in the abstract: a facially neutral biometric-matching algorithm produces systematically unequal outcomes correlated with poverty and manual labour, without any need for discriminatory intent in system design.

6.1.3 The Accountability Gap: Proprietary Algorithms and Opaque Failure

The review stage — where citizens or oversight bodies might contest and correct erroneous exclusions — reveals the most acute accountability deficit. During Supreme Court litigation challenging the Aadhaar Act, the UIDAI itself stated that its biometric-matching algorithms were proprietary, meaning that neither courts, researchers, nor affected citizens could independently verify how well the matching mechanism actually performed (The Wire, 2024). As one doctoral researcher studying digital governance observed, when an administrative authority attributes a service denial to a generic "technical failure," it remains genuinely unclear "what does it mean" for accountability purposes, since the testing methodology and error mechanism are neither published nor independently auditable (The Wire, 2024). This is a textbook illustration of Busuioc's (2021) *information gap* — an oversight forum (in this instance, the judiciary itself) lacking the technical access necessary to interrogate the system it is asked to evaluate.

The constitutional dimension of this gap was addressed, though not fully resolved, in *Justice K.S. Puttaswamy (Retd.) v. Union of India* (2019) 1 SCC 1, in which the Supreme Court upheld the constitutionality of the Aadhaar Act's core architecture while reading down several provisions and severing mandatory Aadhaar-linkage for private telecom and banking services. Notably, the dissenting opinion in that case emphasised that biometric identification systems, being inherently probabilistic, will inevitably misclassify some genuine beneficiaries, and warned that "dignity and the rights of individuals cannot depend on algorithms or probabilities" — a formulation frequently cited in subsequent scholarship as capturing the core normative objection to algorithmic gatekeeping of subsistence entitlements (Nickled and Dimed, 2025). This concern was foreshadowed as early as 2009, when the UIDAI's own committee report acknowledged that the system could not reliably capture damaged or worn fingerprints common among manual labourers and the elderly — a limitation known prior to the system's mandatory rollout for essential welfare access (Nickled and Dimed, 2025).

Empirical economic research on biometric authentication in Indian welfare delivery presents a more mixed picture than the exclusion narrative alone suggests. Randomised evaluations of biometric smartcard-based payment systems found that biometric authentication reduced leakage and transaction costs for most beneficiaries, while acknowledging that authentication failures generate real, if unevenly distributed, additional burdens, including repeated trips to obtain entitlements (Muralidharan, Niehaus, & Sukhtankar, 2016). This body of evidence usefully complicates a purely condemnatory account: Aadhaar-linked authentication has generated genuine administrative gains in reducing corruption and diversion, even as it has generated genuine, unevenly distributed harms of exclusion — a tension that exemplifies the efficiency-accountability trade-off identified in Section 5.4, rather than a simple story of technological failure.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

6.1.4 Lifecycle Assessment

Applying the four-stage framework: at the *design* stage, exclusion risk was foreseeable and documented but not adequately mitigated before mandatory rollout. At the *deployment* stage, exception-handling mechanisms (manual override procedures) exist on paper but are inconsistently implemented and poorly publicised. At the *operation* stage, failure rates remain stubbornly high and are not demographically disaggregated in published data, meaning the state cannot even fully measure the exclusion it is causing (Policy Circle, 2026). At the *review* stage, algorithmic opacity — reinforced by proprietary-technology claims — substantially forecloses meaningful judicial or administrative interrogation of individual failures, leaving affected citizens with limited recourse beyond informal, localised problem-solving.

6.2 Case Study II: The Faceless Assessment Scheme and Algorithmic Tax Administration

6.2.1 Background and Design

India's Income Tax Department introduced the e-Assessment Scheme in September 2019, renamed the Faceless Assessment Scheme in August 2020 as part of the government's "Transparent Taxation — Honouring the Honest" platform. The scheme uses data analytics and artificial intelligence to randomly allocate tax scrutiny cases to assessing officers located outside the taxpayer's home jurisdiction, explicitly designed to eliminate territorial jurisdiction and direct physical interaction between taxpayers and tax officials (ISAS, 2020). The declared policy rationale was to curb corruption and rent-seeking associated with face-to-face scrutiny assessment, where personal interaction between taxpayer and jurisdictional officer had historically created opportunities for undesirable practices (ISAS, 2020).

6.2.2 Deployment and Operation

Under the scheme's architecture, a case scrutiny selected for assessment is randomly assigned via algorithm to an officer in a different city than the taxpayer's location; any subsequent appeal is likewise randomly assigned to a reviewing officer in a third location, and a final review may involve officers in yet another jurisdiction (ISAS, 2020). In its first phase, over 58,000 cases were allocated through this automated, randomised process, with the finance ministry describing the reform as having reduced the proportion of tax returns subject to scrutiny substantially over the preceding decade (Deccan Herald, n.d.). Taxpayers receive notices electronically and respond through a centralised e-filing portal, with a dedicated National e-Assessment Centre coordinating case management (ISAS, 2020).

6.2.3 Accountability Assessment

The Faceless Assessment Scheme offers an instructive counterpoint to the Aadhaar case because its stated objective was explicitly accountability-enhancing: reducing officer discretion was framed as a corruption-reduction and transparency measure rather than merely an efficiency measure. In this respect, it aligns with a strand of literature suggesting that algorithmic decision rules can, in specific institutional contexts characterised by endemic low-

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

level corruption, *increase* rather than decrease certain forms of accountability by removing opportunities for individualised, unrecorded discretion.

However, this case also illustrates important limits to that logic. First, the scheme has not eliminated corruption risk but relocated it: in one documented instance, the Central Bureau of Investigation arrested a deputy commissioner of income tax and a chartered accountant for allegedly leaking confidential case-allocation information — including the identity of the assigned assessing officer — in exchange for bribes, demonstrating that algorithmic randomisation can be circumvented through insider information leakage rather than eliminated as a governance risk (Deccan Herald, n.d.-b). This finding is analytically important because it demonstrates that algorithmic accountability solutions targeting one failure mode (face-to-face discretion) can generate displaced or novel failure modes (information-security breaches around the algorithmic allocation process itself) that require distinct oversight mechanisms.

Second, taxpayer grievance mechanisms for the scheme — including dedicated grievance email addresses established for faceless assessment, penalty, and appeal complaints — indicate persistent implementation friction, including complaints regarding inadequate opportunity to explain complex facts through a purely written, portal-based process, echoing the due-process concerns regarding notice and meaningful hearing opportunity identified by Citron (2008) in the American welfare-automation context. Applying the lifecycle framework, the *design* stage of the Faceless Assessment Scheme reflects an accountability-conscious rationale (corruption reduction through discretion removal), but the *operation* and *review* stages reveal that algorithmic case allocation alone does not resolve — and may partially obscure — questions of substantive fairness in how allocated officers actually evaluate complex, fact-intensive tax positions without in-person clarification opportunities.

6.3 Case Study III: Artificial Intelligence in the Judiciary — SUPACE and SUVAS

6.3.1 Background and Design

The Supreme Court of India constituted an Artificial Intelligence Committee in 2019, chaired by Justice L. Nageswara Rao, to explore the integration of machine learning and AI tools into judicial administration with the explicit aim of enhancing efficiency and reducing case pendency (IndiaAI, n.d.-a). This initiative produced two flagship tools. SUPACE (Supreme Court Portal for Assistance in Courts Efficiency), launched in April 2021, is an AI-enabled research-assistance tool that extracts facts, issues, and points of law from case documents, supports legal researchers and judges through a chatbot interface, and assists in drafting case documents (IndiaAI, n.d.-b). SUVAS (Supreme Court Vidhik Anuvaad Software) is a machine-translation tool designed to render judgments into regional languages, addressing India's significant linguistic diversity (IndiaAI, n.d.-c).

6.3.2 Operation and Distinguishing Features

A critical accountability-relevant design feature distinguishes SUPACE from many algorithmic governance systems examined elsewhere in this paper: SUPACE is explicitly positioned as a decision-*support* tool rather than a decision-*making* tool. Its architecture

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

comprises four functional components — file preview, an interactive chatbot, a fact-extraction "logic gate," and an integrated notebook for drafting — all of which feed information to a human judge who retains ultimate decisional authority (IndiaAI, n.d.-b). This design choice is significant because it retains the "human-in-the-loop" safeguard that Wirtz et al. (2020) identify as essential to mitigating the "dark side" risks of AI governance, and it substantially narrows the accountability gap relative to systems like Aadhaar authentication, where the algorithmic output (a biometric match or non-match) can directly and automatically trigger service denial without meaningful human intermediation.

Empirical assessment of judicial efficiency gains associated with AI-supported case management has been documented at the state level: the Madras High Court, for instance, recorded a case clearance rate of 114 percent in 2022 — meaning it disposed of more cases than were newly filed — a period coinciding with expanded technology adoption in judicial administration under the broader e-Courts initiative (IndiaAI, n.d.-d). Scholars examining the SUPACE and SUVAS initiatives against the backdrop of the proposed EU AI Act have observed that India's approach has thus far proceeded through incremental, procurement-driven pilot projects — including a 2023 bid for AI-based transcription of Supreme Court constitutional bench proceedings — rather than through a comprehensive, ex-ante statutory framework of the kind the EU has pursued, raising questions about whether accountability safeguards are being designed contemporaneously with system deployment or retrofitted afterward (Verma, 2023).

6.3.3 Accountability Assessment

Applying the lifecycle framework, the *design* stage of India's judicial AI tools reflects a deliberate accountability-preserving choice to limit AI to an assistive rather than adjudicative function — an instructive contrast to the "predictive justice" ambitions found in some other jurisdictions' judicial analytics tools, which attempt to forecast case outcomes based on historical data and thereby risk encoding historical biases into prospective case management (Anbarasi & Sankar, 2025). At the *operation* stage, because SUPACE outputs are advisory and mediated by a human judge who remains the accountable decision-maker under existing judicial review doctrine, the explanation and justification obligations central to Bovens's (2007) accountability relationship remain anchored in a human, legally cognisable actor. This represents, within the Indian case studies examined in this paper, the most accountability-conscious design among the systems reviewed — though it is also the system with the narrowest direct impact on individual citizen entitlements, since its primary users are judges and legal researchers rather than the general public directly.

6.4 Case Study IV: NITI Aayog and the Institutional Architecture of Responsible AI

6.4.1 Background

NITI Aayog, India's apex public policy think tank, was mandated in the 2018–19 Union Budget to establish the National Program on AI, resulting in the June 2018 publication of the National Strategy for Artificial Intelligence (NSAI), organised around the tagline "AI for All" (NITI Aayog, 2018). The strategy identified five priority sectors expected to benefit most from AI-

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

driven public intervention: healthcare, agriculture, education, smart cities and infrastructure, and — significant for present purposes — smart mobility and transportation, while explicitly flagging barriers including limited technical expertise, absent data ecosystems, high resource costs, and unclear ethical and privacy frameworks as constraints to be addressed (NITI Aayog, 2018).

6.4.2 The Responsible AI Trilogy

Recognising the governance gap identified in its own foundational strategy, NITI Aayog subsequently commissioned a three-part Responsible AI policy series, developed in collaboration with the Robert Bosch Centre for Data Science and Artificial Intelligence at IIT Madras. Part 1, "Principles for Responsible AI," published in February 2021, articulated seven core principles for AI governance in the Indian context: safety and reliability, inclusivity and non-discrimination, equality, privacy and security, transparency, accountability, and the protection and reinforcement of positive human values (Access Partnership, 2025). These principles were explicitly divided into "system considerations" — addressing algorithmic decision logic, beneficiary inclusion, and accountability of AI-driven decisions — and "societal considerations" — addressing the broader impact of automation on employment (Access Partnership, 2025). Part 2, "Operationalizing Principles for Responsible AI," published in August 2021, translated these abstract principles into actionable regulatory, policy, and capacity-building recommendations for both government and private-sector AI deployers (Access Partnership, 2025). Part 3, released as a discussion paper in November 2022, applied the framework specifically to facial recognition technology (FRT) — a particularly sensitive AI application given its use in law enforcement and public surveillance contexts — outlining the challenges, risks, and design guidelines needed for responsible FRT deployment, and inviting public comment before finalisation (NITI Aayog, 2022).

6.4.3 Accountability Assessment

NITI Aayog's Responsible AI architecture represents India's most direct institutional engagement with the accountability concepts developed in the international scholarship reviewed in Section 2. Its explicit adoption of "accountability" and "transparency" as named principles reflects direct engagement with global normative frameworks, including the OECD Recommendation on Artificial Intelligence, and its system/societal bifurcation implicitly mirrors the distinction in academic literature between algorithm-specific accountability (Wieringa's lifecycle model) and broader socio-economic governance concerns (Ananny and Crawford's socio-technical framing).

However, a critical limitation must be noted: these NITI Aayog documents are policy guidance instruments — "approach papers" and "discussion papers" — rather than binding legal instruments. They articulate principles without creating enforceable rights, statutory obligations, or independent regulatory oversight bodies with power to audit, sanction, or order remediation. This positions India's Responsible AI framework closer to the OECD's soft-law recommendation model than to the EU's binding, risk-tiered regulatory model discussed in Section 5.1. Applying Bovens's (2007) accountability relationship, NITI Aayog's framework establishes normative *obligations to explain and justify* AI system design but does not yet

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

establish a clearly empowered *forum* with binding authority to interrogate specific deployed systems, nor codified *consequences* for non-compliance — an accountability gap that the discussion in Section 8 returns to directly.

7. Legal and Regulatory Framework in India

7.1 Constitutional Foundations: Privacy and Due Process

India's algorithmic governance landscape must be understood against the constitutional backdrop established by *Justice K.S. Puttaswamy (Retd.) v. Union of India* (2017) 10 SCC 1, in which a nine-judge bench of the Supreme Court unanimously recognised the right to privacy as a fundamental right protected under Article 21 of the Constitution. This judgment, decided prior to the more specific Aadhaar-focused ruling discussed in Section 6.1.3, established the doctrinal foundation for subsequent challenges to algorithmic and biometric governance systems, holding that any state intrusion on informational privacy must satisfy tests of legality, legitimate state aim, and proportionality. This proportionality framework has since become the primary judicial tool for evaluating algorithmic governance systems in India, and its application in the Aadhaar judgment — including the dissenting emphasis on the probabilistic, and therefore inherently fallible, nature of biometric matching — continues to shape how Indian courts approach algorithmic accountability claims.

7.2 The Digital Personal Data Protection Act, 2023

India's first comprehensive data protection statute, the Digital Personal Data Protection Act, 2023 (DPDP Act), received presidential assent on 11 August 2023, with substantive provisions commencing in phases beginning November 2025 (Wikipedia, 2026; Deloitte, n.d.). The Act establishes a framework built around the concepts of "Data Principal" (the individual to whom personal data relates), "Data Fiduciary" (the entity determining the purpose and means of data processing), and "Data Processor," and creates a Data Protection Board of India empowered to inquire into data breaches and impose financial penalties, with the Act enabling penalties of up to 250 crore rupees for breaches of data fiduciary obligations (PIB, 2023).

From an algorithmic accountability perspective, the DPDP Act's most significant limitation is definitional and structural rather than merely procedural. Comparative analysis of the Act against the GDPR and similar frameworks notes that the DPDP Act removes purpose limitation obligations as far as state processing is concerned, and — critically for the present inquiry — does not contain provisions specifically addressing harm arising from automated decision-making, nor does it establish an explicit right to explanation for algorithmically-influenced government decisions comparable to Article 22 of the GDPR (Forrester, n.d.). This is a material gap when evaluated against the theoretical framework developed in Section 3: the DPDP Act addresses the *design* stage of the algorithmic lifecycle through consent and data-minimisation obligations, and creates a *review* forum through the Data Protection Board, but does not squarely address the *operation* stage question of whether and how an individual can demand an explanation of, or contest, a specific automated or algorithmically-assisted government decision affecting them — precisely the accountability gap that Citron's (2008) technological due process critique identified nearly two decades earlier in a different jurisdiction.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

7.3 Sectoral and Emerging Instruments

Beyond the DPDP Act, India's algorithmic governance landscape includes the Information Technology Rules, 2021, which regulate intermediaries including social media and digital platforms and have been interpreted as extending to AI-generated content concerns such as deepfakes, and sector-specific regulatory engagement, including the Telecom Regulatory Authority of India's 2023 recommendations on leveraging AI and big data in telecommunications, which called for a common cross-sectoral AI regulatory framework (Access Partnership, 2025). The absence, to date, of a horizontal, binding AI-specific statute — comparable to the EU AI Act — means that India's algorithmic governance currently operates through a patchwork of sectoral rules, constitutional proportionality doctrine, the DPDP Act's data-processing obligations, and non-binding NITI Aayog policy guidance, an architecture that mirrors, in structural terms, the fragmented "administrative law by adaptation" approach observed in the United States (Section 5.2) more closely than the EU's comprehensive risk-tiered model.

8. Discussion: Cross-Cutting Accountability Gaps and Comparative Reflection

8.1 The Persistence of the Information and Explanation Gaps

Across all four Indian case studies examined in Section 6, Busuioc's (2021) information gap recurs as the most consistent structural obstacle to algorithmic accountability. In the Aadhaar case, this gap is explicit and acknowledged: the UIDAI's own proprietary-technology claim in Supreme Court proceedings directly forecloses independent verification of matching-algorithm performance. In the Faceless Assessment Scheme, the gap is procedural: taxpayers subject to algorithmic case allocation have no visibility into the selection criteria or risk model underlying their scrutiny selection. In the judicial AI case, the gap is narrower precisely because human judges retain interpretive authority over SUPACE's outputs — reinforcing the paper's broader finding that accountability outcomes correlate strongly with the degree of human decisional intermediation retained in system design, consistent with Wirtz et al.'s (2020) emphasis on human-in-the-loop safeguards as a mitigating design choice.

8.2 Efficiency Gains Are Real, But Unevenly Distributed

A recurring pattern across the Indian case studies is that algorithmic governance genuinely delivers the efficiency and corruption-reduction benefits its designers intend — reduced welfare leakage, reduced face-to-face tax-scrutiny corruption opportunities, accelerated judicial case clearance — while simultaneously distributing the *costs* of algorithmic error disproportionately onto citizens with the least capacity to contest them. This asymmetry, most starkly visible in the Aadhaar case's disproportionate impact on manual labourers and the elderly, echoes Eubanks's (2018) "digital poorhouse" thesis and O'Neil's (2016) observation that algorithmic systems tend to be deployed with greatest coercive force precisely against populations with the least political and economic power to resist erroneous classification. This suggests that algorithmic accountability cannot be evaluated solely at the level of aggregate efficiency metrics — the standard by which government agencies typically justify AI adoption — but must incorporate distributive impact analysis as a core evaluative criterion.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

8.3 Soft-Law Principles Without Hard-Law Remedies

India's NITI Aayog Responsible AI framework and its DPDP Act together represent significant normative progress relative to a decade ago, when no comprehensive Indian policy instrument addressed algorithmic governance at all. Yet, evaluated against the relational accountability model (Section 3.1), a structural weakness persists: principles exist without a clearly empowered forum for individual grievance redress specific to algorithmic harm, and consequences for algorithmic malfunction remain addressed, if at all, through general data-protection penalty provisions rather than through remedies tailored to the specific harm of erroneous automated exclusion from an entitlement. This is consistent with Wieringa's (2020) broader finding that lifecycle accountability requires attention to *review*-stage mechanisms specifically, not merely design-stage principles — a stage that both the DPDP Act and the NITI Aayog framework, as presently constituted, address only partially.

8.4 Displacement Rather Than Elimination of Discretion and Risk

The Faceless Assessment Scheme's documented corruption case (Section 6.2.3) offers a valuable cautionary finding for the broader algorithmic accountability literature: automating a decision process does not eliminate the underlying governance risk it targets so much as it relocates that risk to a new point in the system — in this instance, from face-to-face bargaining to insider leakage of algorithmically-generated case-allocation data. This finding supports Ananny and Crawford's (2018) socio-technical caution against treating algorithmic reform as a self-contained technical fix; effective accountability design must anticipate that human actors within the surrounding institutional system will adapt their behaviour around new algorithmic constraints, sometimes in ways that reproduce the original governance failure in a new form.

8.5 Comparative Positioning

Positioned against the global regulatory landscape surveyed in Section 5, India's current approach most closely resembles a hybrid of the OECD soft-law model and the United States' adaptation-through-existing-doctrine model, rather than the EU's binding risk-tiered statutory approach. This positioning is neither inherently deficient nor inherently adequate — Busuioc's (2021) comparative case analysis across Estonia, Singapore, the United States, and the United Kingdom found that governance effectiveness depended less on the specific regulatory model chosen than on whether oversight, auditability, and public engagement mechanisms were substantively operationalised in practice, regardless of their formal legal source. Applied to India, this suggests that the central policy question is not primarily "does India need a binding AI statute," though there are strong reasons discussed below to think it does, but more fundamentally "what independent auditability and grievance-redress capacity currently exists for the algorithmic systems already deployed at scale" — and the case studies in Section 6 suggest that capacity remains thin.

9. Policy Recommendations

Drawing on the preceding analysis, this paper offers the following recommendations, organised according to the four-stage lifecycle accountability framework developed in Section 3.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

9.1 Design-Stage Recommendations

Government agencies deploying AI systems affecting individual entitlements should be statutorily required to conduct and publish pre-deployment algorithmic impact assessments, disaggregated by demographic characteristics known to correlate with system error — for instance, occupation-linked biometric degradation in the Aadhaar context. Such assessments should be modelled on, though need not be identical to, the "high-risk" conformity assessment obligations found in the EU AI Act, adapted to India's administrative capacity and developmental priorities. NITI Aayog's existing Responsible AI principles provide a sound normative starting point; the priority now is converting these principles into binding procurement and deployment conditions for government AI systems.

9.2 Deployment-Stage Recommendations

Mandatory human-in-the-loop fallback mechanisms should accompany any algorithmic system whose output can trigger denial of an essential entitlement — food subsidy, employment wages, or pension access. The existing exception-handling provisions nominally present in Aadhaar-linked welfare delivery should be strengthened, standardised, and made genuinely accessible at the point of service delivery, rather than requiring citizens to navigate a separate, often distant, grievance process after the point of exclusion. This recommendation directly reflects the accountability-preserving design choice already evident in the SUPACE judicial AI case study, where human decisional authority is retained precisely at the point where AI output could otherwise directly determine an outcome.

9.3 Operation-Stage Recommendations

Government agencies should be required to publish periodic, disaggregated error and exclusion statistics for algorithmic systems affecting the public, comparable to the parliamentary evidence already generated through the Public Accounts Committee's scrutiny of Aadhaar authentication failure rates. Institutionalising this reporting as a routine statutory obligation, rather than relying on ad hoc parliamentary questioning, would meaningfully narrow the information gap identified throughout this paper's case studies.

9.4 Review-Stage Recommendations

India's DPDP Act should be amended, or complemented by dedicated administrative law reform, to establish an explicit right to a reasoned explanation for government decisions substantially informed by automated or algorithmic processing, alongside a clearly designated, accessible administrative forum for contesting such decisions — addressing the gap identified in Section 7.2. Independent, adequately resourced algorithmic auditing capacity, potentially housed within or alongside the Data Protection Board of India, should be empowered to review proprietary claims of the kind made by UIDAI in Aadhaar litigation, ensuring that trade-secrecy claims do not categorically foreclose judicial or regulatory scrutiny of algorithmic decision logic affecting fundamental rights.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
Vol. 3, Issue 11, May 2026
Available Online : https://shodhprakashan.in/

9.5 Cross-Cutting Recommendation: Participatory Design

Consistent with the socio-technical systems perspective adopted in this paper's theoretical framework, affected communities — particularly manual labourers, elderly citizens, and other groups disproportionately burdened by biometric authentication failure — should be substantively consulted during the design and periodic revision of algorithmic welfare-delivery systems, rather than treated solely as post-deployment complainants. NITI Aayog's public consultation process for its facial recognition technology discussion paper offers a replicable procedural template that could be extended to other high-stakes algorithmic governance domains, including welfare authentication systems.

10. Conclusion

The rise of AI-driven governance in public administration represents neither an unambiguous triumph of administrative modernisation nor a wholesale erosion of citizen protection; the evidence examined in this paper, drawn substantially from the Indian administrative experience, supports a more differentiated conclusion. Algorithmic systems have delivered measurable gains in efficiency, consistency, and — in specific contexts such as the Faceless Assessment Scheme — reduced opportunities for individualised corruption. At the same time, these same systems have introduced accountability gaps that existing legal and institutional architecture, in India as elsewhere, has not yet adequately resolved: an information gap that shields algorithmic logic from independent scrutiny, an explanation gap that leaves affected citizens without meaningful understanding of adverse decisions, and a justification gap reflecting continuing normative disagreement about the standards against which algorithmic governance should be judged.

The Indian case studies examined in this paper — Aadhaar's biometric welfare gatekeeping, the Faceless Assessment Scheme's algorithmic tax administration, the Supreme Court's human-in-the-loop judicial AI tools, and NITI Aayog's Responsible AI policy architecture — demonstrate that accountability outcomes are not determined by the mere fact of algorithmic deployment, but by specific, identifiable design choices: whether human decisional authority is retained at critical junctures, whether error rates are transparently measured and disaggregated, whether affected citizens have accessible avenues for explanation and correction, and whether proprietary or trade-secrecy claims are permitted to foreclose independent audit. Where these design choices favour accountability — as in the judicial AI case — algorithmic governance appears compatible with, and even supportive of, administrative legitimacy. Where these choices favour opacity and unmediated automation — as in aspects of the Aadhaar authentication system — algorithmic governance risks reproducing and amplifying the very administrative failures, exclusion and unaccountable discretion, it was often introduced to remedy.

India's regulatory trajectory, anchored in the Digital Personal Data Protection Act, 2023 and NITI Aayog's Responsible AI framework, reflects genuine and increasing engagement with the global algorithmic accountability discourse reviewed in this paper. Yet a persistent gap remains between principle and enforceable remedy — a gap this paper has argued must be closed through binding algorithmic impact assessment requirements, statutorily mandated human

fallback mechanisms, transparent error reporting, an explicit right to explanation for algorithmically-influenced state decisions, and independent audit capacity empowered to pierce proprietary-technology claims where fundamental rights are at stake. As AI-driven governance continues its rapid expansion across public administration globally, the central lesson from the Indian experience examined here is that algorithmic accountability is not a technical afterthought to be addressed once systems are deployed, but a design obligation that must be built into public administration's algorithmic future from its inception.

References

- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973–989.
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review, 104*(3), 671–732.
- Berryhill, J., Heang, K. K., Clogher, R., & McBride, K. (2019). *Hello, World: Artificial intelligence and its use in the public sector* (OECD Working Papers on Public Governance No. 36). OECD Publishing.
- Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal, 13*(4), 447–468.
- Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review, 81*(5), 825–836.
- Citron, D. K. (2008). Technological due process. *Washington University Law Review, 85*(6), 1249–1313.
- Citron, D. K., & Pasquale, F. (2014). The scored society: Due process for automated predictions. *Washington Law Review, 89*(1), 1–33.
- Diakopoulos, N. (2015). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism, 3*(3), 398–415.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (Supreme Court of India).
- Justice K.S. Puttaswamy (Retd.) v. Union of India, (2019) 1 SCC 1 (Supreme Court of India).
- Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2016). Building state capacity: Evidence from biometric smartcards in India. *American Economic Review, 106*(10), 2895–2929.
- NITI Aayog. (2018). *National strategy for artificial intelligence: #AIforAll*. Government of India.
- NITI Aayog. (2021a). *Approach document for India, Part 1: Principles for responsible AI*. Government of India.
- NITI Aayog. (2021b). *Approach document for India, Part 2: Operationalizing principles for responsible AI*. Government of India.
- NITI Aayog. (2022). *Responsible AI for all: Adopting the framework — A use case approach on facial recognition technology*. Government of India.

- Novelli, C., Taddeo, M., & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society, 39*(4), 1871–1882.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Organisation for Economic Co-operation and Development. (2024). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449).
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Taddeo, M., & Floridi, L. (2018). Regulate artificial intelligence to avert cyber arms race. *Nature, 556*(7701), 296–298.
- Wieringa, M. (2020). What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 1–18). Association for Computing Machinery.
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration, 42*(7), 596–615.
- Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration, 43*(9), 818–829.