Auditing the Automated State: Impact Assessments and Institutional Capacity for Ethical AI in Indian Governance

Lt. Dr. Kongala Sukumar

Assistant Professor of Public Administration
MALD Government Degree College,
GADWAL, Pin-509126, Telangana

Abstract:
As Indian public administration increasingly delegates welfare targeting, policing, and service delivery to algorithmic systems, the absence of a statutory impact-assessment regime has left citizens with limited means of contesting automated decisions. This paper examines Algorithmic Impact Assessments (AIAs) as a governance instrument capable of anchoring ethical artificial intelligence within Indian public institutions, and asks whether existing administrative capacity can sustain such a regime. Drawing on a comparative analysis of the Canadian Directive on Automated Decision-Making, the European Union's data-protection impact-assessment architecture, and emergent Indian policy instruments issued by NITI Aayog, the paper develops a four-dimensional institutional capacity framework spanning technical expertise, legal authority, procedural infrastructure, and civic accountability. Three illustrative domains of automation in Indian governance — direct benefit transfer, predictive policing, and facial recognition — are examined against this framework. The analysis finds that India's AI policy discourse has matured considerably since 2018, yet implementation remains fragmented across ministries, with no binding pre-deployment review requirement in force. The paper concludes with a phased institutional roadmap for embedding impact assessments within India's administrative law architecture, emphasising capacity-building, sectoral regulatory coordination, and public registries of automated systems as preconditions for accountable algorithmic governance.

Keywords: algorithmic impact assessment; ethical artificial intelligence; institutional capacity; Indian public administration; administrative accountability; automated decision-making; NITI Aayog

1. Introduction

The Indian state has, over the past decade, become one of the most consequential deployers of algorithmic decision-making in the world. From the biometric authentication layer of Aadhaar and the algorithmic beneficiary lists that determine Direct Benefit Transfer (DBT) disbursement, to predictive policing tools piloted in several states and facial recognition technology (FRT) deployed at railway stations, airports, and protest sites, automated systems now mediate the relationship between citizen and state at a scale that has few global parallels. This expansion has occurred largely without a corresponding architecture of pre-deployment scrutiny. Unlike environmental projects, which since the 1970s have been subject to Environmental Impact Assessments before approval, or personal-data processing operations in the European Union, which must undergo a Data Protection Impact Assessment (DPIA) under Article 35 of the General Data Protection Regulation, algorithmic systems adopted by Indian public authorities face no equivalent statutory checkpoint. This gap matters because algorithmic systems in welfare and security administration do not merely automate existing rules; they frequently encode new forms of discretion that are opaque to affected citizens, to courts,

and often to the officials who operate them. Scholars of administrative law have argued that constitutional guarantees of fairness, non-arbitrariness, and reasoned decision-making are difficult to enforce against a decision-making process that cannot be inspected, contested, or even fully described (Automated State Action in India, 2025). The absence of impact assessment obligations compounds this problem: agencies can procure and deploy automated systems without documenting their purpose, their training data, their error rates, or their disparate effects on vulnerable groups.

Algorithmic Impact Assessments have emerged internationally as the leading procedural response to this accountability deficit. Modelled loosely on environmental impact statements, an AIA requires the developer or procuring agency of an automated decision system to evaluate, document, and often publicly disclose the system's likely effects before it is deployed, and to build in mechanisms for ongoing monitoring after deployment (Reisman et al., 2018). Canada's Directive on Automated Decision-Making, in force since 2019, ties the stringency of review to a graduated risk score; the European Union's proposed Artificial Intelligence Act, tabled in 2021, adopts a similar risk-tiered logic for AI systems used in the public sector (Scassa, 2020). India's own policy trajectory, beginning with NITI Aayog's National Strategy for Artificial Intelligence in 2018 and continuing through its 2021 Approach Document for Responsible AI, has articulated ethical principles for AI but has not yet translated them into a binding pre-deployment assessment mechanism (NITI Aayog, 2018, 2021).

This paper asks two related questions. First, what would an Algorithmic Impact Assessment regime suited to Indian public administration need to contain, given the comparative experience of Canada, the European Union, and civil-society frameworks such as the AI Now Institute's practical AIA model? Second, and more centrally, does the Indian administrative state currently possess the institutional capacity — technical, legal, procedural, and civic — to operate such a regime effectively, or would a transplanted AIA framework risk becoming what governance scholars term a "tenet of best practice" rather than an enforceable safeguard (The Need for Responsible Use of AI by Public Administration, 2025)? The paper argues that capacity-building must be treated as co-equal with regulatory design: an impact assessment mandate without trained personnel, empowered oversight bodies, and resourced enforcement mechanisms risks producing paper compliance rather than substantive accountability.

The stakes of getting this design right are not merely administrative. Impact assessment regimes function as a proxy for a broader constitutional question: whether the exercise of state power through code remains subject to the same standards of legality, fairness, and reasoned justification that apply to the exercise of state power through a human decision-maker. When an algorithm denies a subsidy, flags a neighbourhood for increased patrolling, or produces a false facial-recognition match, the citizen affected typically cannot ask why in any meaningful sense, because the "why" is distributed across a training dataset, a set of model weights, and a procurement contract that the citizen has no legal right to inspect. An impact assessment regime does not by itself dissolve this opacity, but it creates the documentary trail — the description of purpose, data provenance, tested error rates, and mitigation steps — without which contestation is impossible in practice, however available it may be in constitutional theory. The paper's institutional capacity framework is offered as a diagnostic tool for identifying where, along that documentary trail, the Indian administrative state is currently weakest.

The paper proceeds as follows. Section 2 reviews the theoretical foundations of impact assessment as a regulatory instrument and surveys the international AIA landscape. Section 3 develops a four-dimensional institutional capacity framework. Section 4 sets out the paper's document-based comparative methodology. Section 5 applies the framework to India's regulatory landscape and to three illustrative domains of state automation. Section 6 proposes a phased roadmap for institutionalising AIAs in India. Section 7 concludes.

2. Literature Review

2.1 Impact Assessment as a Regulatory Technology

Impact assessment as a regulatory form predates artificial intelligence by half a century. The United States' National Environmental Policy Act of 1969 pioneered the idea that a state agency contemplating an action with significant, uncertain effects should be required to study and disclose those effects before acting, rather than after harm has occurred (Reisman et al., 2018). Scholars have since extended this logic to privacy, in the form of Privacy Impact Assessments, and to data protection, culminating in the DPIA obligation under the GDPR, which requires a systematic description of processing operations, an assessment of necessity and proportionality, and an assessment of risks to the rights of data subjects prior to high-risk processing (Regulation (EU) 2016/679, 2016). The common structural feature across these instruments is procedural rather than substantive: none of them prohibit a given project outright, but each forces the decision-maker to generate a documented, contestable record of anticipated harms before the point of no return.

Algorithmic Impact Assessments extend this procedural logic to automated decision systems. The AI Now Institute's foundational 2018 framework proposed that public agencies disclose the existence and purpose of automated decision systems, allow external researchers to review and audit them, and provide meaningful mechanisms for redress before, not after, deployment (Reisman et al., 2018). Selbst (2021) cautions, however, that impact assessments imported from environmental and privacy law carry institutional assumptions that do not automatically transfer to algorithmic systems: environmental impact statements rely on comparatively mature scientific consensus about measurable harms, whereas algorithmic harms — bias, opacity, feedback loops — are often contested, emergent, and unevenly distributed across social groups. An AIA regime, Selbst argues, is only as effective as the institutional logics of the organisations that operate it; absent independent verification capacity, agencies may treat the assessment as a compliance checkbox rather than a genuine occasion for redesign.

2.2 Comparative Models

Canada operationalised the AIA concept most concretely through its Directive on Automated Decision-Making, which came into force in April 2019 and applies to all federal automated decision systems. The Directive requires departments to complete an Algorithmic Impact Assessment questionnaire that generates an impact score across four tiers, with the tier determining mandatory safeguards ranging from internal peer review to independent external audit and human-in-the-loop requirements for the highest-risk systems (Scassa, 2020). Scassa's administrative law analysis notes that the Directive's principal innovation is to make the impact score a public document, thereby creating an evidentiary record that can support judicial review even where the underlying algorithm remains a trade secret.

Beyond these two binding or semi-binding regimes, a wider civil-society and academic literature has proposed AIA components intended to be portable across jurisdictions: mandatory public notice before procurement, structured opportunities for external researcher review, meaningful mechanisms for individuals to contest an automated decision, and ongoing post-deployment monitoring rather than a one-time, pre-deployment sign-off (Reisman et al., 2018). This literature is useful for a jurisdiction such as India precisely because it is not tied to a single legal tradition; it identifies functional components — notice, review, contestation, monitoring — that can, in principle, be assembled using whichever legal instruments, executive orders, statutory amendments, or regulatory directions, are institutionally available, rather than requiring a single comprehensive AI statute of the kind the EU has pursued.

The European Union's approach has developed along two tracks. The GDPR's Article 35 DPIA obligation, while not algorithm-specific, already captures most public-sector automated decision systems because they process personal data at scale (Regulation (EU) 2016/679, 2016). The proposed Artificial Intelligence Act builds a second, AI-specific layer on top of this, classifying systems used in law enforcement, migration, and the administration of justice or public benefits as "high-risk" and subjecting them to conformity assessment,

technical documentation, and post-market monitoring obligations. Comparative legal scholarship has observed that the institutionalisation of impact assessment depends as much on administrative and financial resources and an institutional culture of accountability as on the binding character of the underlying norm (The Need for Responsible Use of AI by Public Administration, 2025).

2.3 The Indian Policy Landscape

India's engagement with AI governance began with NITI Aayog's National Strategy for Artificial Intelligence (NSAI), published in June 2018, which framed AI adoption around the "AI for All" objective and identified five priority sectors — healthcare, agriculture, education, smart cities, and mobility — while flagging the need for mechanisms to build public trust in algorithmic systems (NITI Aayog, 2018). This was followed in February 2021 by Part 1 of the Approach Document for Responsible AI, which articulated system-level principles (safety and reliability, equality, inclusivity and non-discrimination, privacy and security, transparency, accountability) and societal-level principles for responsible AI adoption, and in August 2021 by Part 2, which sought to operationalise these principles through regulatory, policy, and capacity-building recommendations (NITI Aayog, 2021). A third discussion paper applied the framework to facial recognition technology as a use case, again on a non-binding, consultative basis (NITI Aayog, 2022).

Despite this normative groundwork, India's Digital Personal Data Protection Act does not impose a DPIA-equivalent obligation and grants broad exemptions to government agencies invoking sovereignty or public order, while no statute currently requires an Algorithmic Impact Assessment before high-risk automated systems are deployed in welfare, policing, or surveillance (Automated State Action in India, 2025). Analyses of sector regulators such as the Competition Commission of India note a parallel capacity deficit: regulators tasked with scrutinising algorithmic conduct have not been shown to recruit the technically trained staff — data scientists, ML engineers, algorithm auditors — needed to interrogate the systems they are meant to oversee (Centre for Policy Research India, n.d.). The result is a policy architecture rich in stated principles but thin in enforceable, resourced mechanisms — precisely the gap an institutional capacity framework must interrogate.

2.4 Ethical Principles and the Accountability Gap

A recurring finding in the international literature is that ethical AI principles — fairness, transparency, accountability, human oversight — are, by themselves, a poor predictor of accountable outcomes, because principles do not specify who is empowered to act when a system fails to meet them, nor what consequence follows from that failure. NITI Aayog's Approach Document is explicit that AI systems should be "safe and reliable" and that there should be mechanisms of "accountability" for adverse outcomes, but the document does not designate an authority empowered to suspend a deployed system found to breach these principles, nor does it prescribe a remedy available to an affected citizen (NITI Aayog, 2021). This is not a uniquely Indian shortcoming: reviewing the same tension, the Strathmore University Centre for Intellectual Property and Information Technology Law observes that rising organisational adoption of AI has outpaced the spread of formal impact-assessment practice, so that governance structures capable of catching harm before deployment remain the exception rather than the norm even in jurisdictions with more mature AI markets (CIPIT, 2023). The gap is acute in the Indian context because so much of the country's highest-stakes algorithmic deployment occurs in welfare and security administration, sectors in which the ordinary market disciplines that might otherwise pressure a private company toward voluntary disclosure are largely absent.

The distinction between a principle and a mechanism is therefore the organising insight behind this paper's institutional capacity framework. Where the literature reviewed above documents what an AIA regime should contain in the abstract, the framework developed in Section 3 asks a narrower and more operational question: for each of the mechanisms an AIA regime requires — a body to review, a template to standardise, a register to disclose, a route to contest — does an Indian analogue currently exist, and if so, is it resourced to function as intended?

3. An Institutional Capacity Framework for Algorithmic Impact Assessment

To move beyond a purely doctrinal account of what an Indian AIA statute should say, this paper proposes a four-dimensional framework for assessing whether an administrative system can operate such a regime in practice. The framework draws on the comparative literature reviewed above, particularly Selbst's (2021) institutional critique and the capacity concerns raised in relation to Indian sectoral regulators (Centre for Policy Research India, n.d.).

The first dimension, technical expertise, refers to the presence within the implementing agency, or within an oversight body it can call upon, of personnel capable of evaluating training data, model design choices, and error and disparity metrics. The second dimension, legal authority, refers to the existence of a clear statutory or executive mandate empowering an identifiable body to require, review, and reject impact assessments, as distinct from a voluntary or advisory framework. The third dimension, procedural infrastructure, refers to the standardised tools, templates, and workflows — of the kind Canada's scored questionnaire provides — that allow assessments to be produced, compared, and audited consistently across agencies rather than improvised case by case. The fourth dimension, civic accountability, refers to public-facing mechanisms, including published registries of automated systems, avenues for individual contestation of automated decisions, and standing for civil society or researchers to scrutinise disclosed assessments.

These four dimensions are interdependent rather than sequential. Legal authority without technical expertise produces a reviewing body that can demand an assessment but cannot verify whether it is accurate; technical expertise without legal authority produces skilled analysts with no power to require a system's redesign; procedural infrastructure without civic accountability produces a well-documented but invisible paper trail that never reaches the citizens it is meant to protect; and civic accountability without the other three dimensions produces a right to ask questions that no institution is resourced to answer. The framework is deliberately agnostic as to sequencing at the level of theory, but Section 6 argues that in the specific case of Indian public administration, the interdependence is best resolved by establishing legal authority and procedural infrastructure first, since these create the institutional demand that then justifies the recruitment of technical staff.

Table 1 summarises this framework and situates the Canadian and EU models, together with the current Indian position, against each dimension. The comparison is illustrative rather than exhaustive, but it clarifies that India's principal deficit lies less in the articulation of ethical principles — where NITI Aayog's documents are reasonably comprehensive — than in legal authority and procedural infrastructure, the two dimensions that convert principle into enforceable practice.

Table 1. A Four-Dimensional Institutional Capacity Framework, Compared Across Jurisdictions

Dimension	Canada (Directive on ADM)	European Union (GDPR / AI Act)	India (current position)
Technical expertise	Centralised Treasury Board guidance; departmental data scientists required for higher-risk tiers	Data Protection Officers mandated for high-risk processing; AI Act envisages notified bodies	No standing requirement for algorithm-auditing personnel in deploying ministries (Centre for Policy Research India, n.d.)
Legal authority	Binding Directive under the Financial Administration Act;	Binding Regulation (GDPR) and proposed Regulation (AI Act)	Principles are advisory (NITI Aayog, 2018, 2021); DPDP Act

Dimension	Canada (Directive on ADM)	European Union (GDPR / AI Act)	India (current position)
	Treasury Board can compel compliance	with supervisory authorities	exempts many government uses (Automated State Action in India, 2025)
Procedural infrastructure	Standardised, scored AIA questionnaire; four defined risk tiers (Scassa, 2020)	DPIA template under Article 35; forthcoming AI Act conformity assessment	No standard assessment template mandated across ministries
Civic accountability	Impact scores and mitigation plans published	DPIA outcomes reviewable by supervisory authority; AI Act mandates public database for high-risk systems	No public registry of government automated decision systems

Sources: NITI Aayog (2018, 2021); Scassa (2020); Regulation (EU) 2016/679 (2016); Reisman et al. (2018); Automated State Action in India (2025); Centre for Policy Research India (n.d.). Compiled by the author.

4. Methodology

This paper adopts a qualitative, document-based comparative policy analysis. Primary sources include NITI Aayog's National Strategy for Artificial Intelligence (2018) and its two-part Approach Document for Responsible AI (2021), Canada's Directive on Automated Decision-Making and its accompanying Algorithmic Impact Assessment tool, the text of GDPR Article 35, and the European Commission's proposed Artificial Intelligence Act. These are read alongside secondary academic and policy literature on algorithmic accountability, including the AI Now Institute's foundational AIA framework (Reisman et al., 2018), Selbst's (2021) institutional critique, and Indian administrative law scholarship examining automated state action (Automated State Action in India, 2025).

The analysis proceeds in three steps. First, the four-dimensional capacity framework set out in Section 3 was derived inductively from recurring themes in the comparative literature — expertise, authority, procedure, and accountability recur, in varying language, across nearly every account of why impact assessment regimes succeed or stall. Second, publicly available government documents, parliamentary responses, and secondary reporting were used to code India's current position on each dimension, distinguishing between principles that are merely stated and mechanisms that are operative and resourced. Third, three domains of Indian state automation — direct benefit transfer, predictive policing, and facial recognition technology — were selected as illustrative cases because each has attracted independent scholarly or civil-society documentation, allowing triangulation beyond government self-reporting. The paper does not claim to offer a statistically representative audit of all automated systems in Indian governance; its purpose is analytical and normative, using documented cases to test the explanatory value of the capacity framework rather than to quantify the universe of algorithmic deployments.

A limitation of this approach is that much of the technical detail of government-deployed algorithms — model architecture, training data provenance, error rates by demographic group — is itself undisclosed, which is precisely the transparency deficit the paper argues an AIA regime should remedy. Where primary technical

documentation is unavailable, the analysis relies on the characterisation of systems in peer-reviewed and civil-society secondary literature, flagged accordingly.

A second limitation concerns comparability. Canada and the European Union operate within federal and supranational administrative law traditions that differ from India's in the relative weight given to executive rule-making versus primary legislation, and in the maturity of their data-protection supervisory authorities. The paper treats these jurisdictions as sources of design options rather than as templates for direct transplantation, consistent with comparative administrative law's general caution against assuming that an instrument's effectiveness travels intact across differing constitutional and bureaucratic contexts. This caution motivates the paper's emphasis, in Section 6, on a phased roadmap calibrated to existing Indian institutional starting points rather than a single legislative transplant.

5. Findings and Discussion

5.1 Direct Benefit Transfer and Welfare Automation

The Direct Benefit Transfer architecture, built atop Aadhaar-linked bank accounts, uses automated eligibility and de-duplication algorithms to determine who receives subsidies for food, cooking fuel, pensions, and employment guarantee wages. Administrative law scholarship examining this system has documented instances in which biometric authentication failures or algorithmic exclusion from beneficiary lists resulted in denial of entitlements, with affected citizens frequently unable to determine why a transaction failed or whom to approach for correction (Automated State Action in India, 2025). Measured against the capacity framework, DBT illustrates a procedural infrastructure gap in particular: exclusion errors are typically discovered only after the fact, through investigative journalism or grievance redressal, rather than anticipated through a pre-deployment assessment of authentication failure rates among elderly, manual-labour, or disabled beneficiaries whose fingerprints are harder to match.

5.2 Predictive Policing

Several Indian states have piloted predictive policing systems that generate risk scores for individuals or flag geographic "hotspots" for patrol allocation, drawing on historical crime data. Because such systems are trained on past enforcement records, they risk reproducing patterns of over-policing in already heavily surveilled neighbourhoods, a dynamic well documented in comparative predictive-policing scholarship and flagged as a central concern in Indian administrative law analysis of automated policing tools (Automated State Action in India, 2025). This domain illustrates a legal authority gap: no sector-specific regulator currently has a statutory mandate to review a predictive policing tool before a state police department adopts it, and the broad exemptions available to government agencies under data protection law further limit external scrutiny.

5.3 Facial Recognition Technology

Facial recognition technology has been deployed by Indian police forces, airports, and railway authorities, and was the subject of NITI Aayog's 2022 discussion paper applying its Responsible AI framework to a specific use case (NITI Aayog, 2022). That paper itself acknowledged the absence of an overarching legal framework governing FRT accuracy thresholds, consent, and permissible use, and was released for public comment rather than as a binding instrument. FRT illustrates the civic accountability gap most starkly: there is no public registry disclosing where FRT systems are deployed, their vendor, their documented accuracy across demographic groups, or the legal basis for a given deployment, which forecloses the kind of contestation that published Canadian impact scores are designed to enable.

Table 2 consolidates this case analysis, mapping each domain against the four capacity dimensions and indicating whether a pre-deployment impact assessment currently exists in any form.

Table 2. Institutional Capacity Gaps Across Three Domains of Automated Governance in India

Domain	Technical expertise gap	Legal authority gap	Procedural infrastructure gap	Civic accountability gap	Pre-deployment AIA in force
Direct Benefit Transfer (Aadhaar-linked)	Authentication error rates not independently audited	DPDP Act exemptions limit oversight of government processing	No standard pre-deployment exclusion-risk assessment	Beneficiaries lack a documented channel to contest algorithmic denial	No
Predictive policing (state pilots)	Model training data and validation not independently reviewed	No sector regulator with mandate to pre-approve tools	No common risk-tiering or review template across states	No public disclosure of which jurisdictions use predictive tools	No
Facial recognition technology	Accuracy and bias testing not independently verified across demographic groups	No binding FRT-specific statute; 2022 framework advisory only	No standard consent or accuracy-threshold protocol	No public registry of deployment sites or vendors	No

Sources: NITI Aayog (2022); Automated State Action in India (2025); Centre for Policy Research India (n.d.). Compiled by the author.

5.4 Reading the Findings Through the Capacity Framework

Two patterns emerge from the case analysis. First, India's deficit is not primarily one of principle articulation. NITI Aayog's system-level principles — safety and reliability, transparency, accountability, privacy — track closely with the values embedded in the Canadian Directive and the GDPR's DPIA criteria (NITI Aayog, 2021). What is missing is the translation mechanism: a binding requirement that a specific agency, before procuring or deploying a specific system, must complete a specific documented assessment reviewable by a specific oversight body. Second, the three domains examined share a common structural feature — each was deployed through executive or administrative action rather than dedicated primary legislation, which means the ordinary parliamentary scrutiny that might otherwise generate impact-assessment-like debate has been largely absent.

This reading has an important implication for capacity-building strategy. Because India's gap is concentrated in legal authority and procedural infrastructure rather than in normative articulation, capacity-building efforts that focus solely on training officials in AI ethics — a common recommendation in current policy discourse — will have limited effect unless paired with a binding assessment mandate and a standard, ministry-agnostic assessment template. Conversely, importing a scored questionnaire model such as Canada's without also building the technical auditing capacity to verify agency self-reporting risks producing exactly the "compliance behaviour" that Selbst (2021) warns can hollow out an AIA regime from within.

5.5 Judicial Review and the Right to Information as Partial Substitutes

In the absence of a statutory AIA regime, two existing accountability routes have carried a disproportionate share of the burden of scrutinising automated state action in India: constitutional judicial review and the Right to Information Act. Both are important, and both are structurally ill-suited to serve as substitutes for a pre-deployment assessment mechanism. Judicial review is available only after a system has already caused a contestable harm, typically to an individual litigant with the resources to bring a case, and courts examining automated decisions have had to develop technical fluency on an ad hoc, case-by-case basis rather than drawing on a standing body of independently verified technical findings (Automated State Action in India, 2025). Right to Information requests can, in principle, compel disclosure of some system documentation, but public authorities have wide latitude to resist disclosure on security, commercial-confidentiality, or sovereignty grounds — the same grounds that limit DPDP Act oversight of government processing — and a successful request in any case arrives well after deployment, when the system's design choices are already embedded in the software procured and difficult to unwind.

The comparative significance of this observation is that an AIA regime is not simply an additional accountability layer stacked on top of judicial review and RTI; it changes the timing and evidentiary basis of both. A published, risk-tiered impact assessment of the kind Canada requires gives a prospective litigant a documentary record to plead from, and gives an RTI applicant a starting point rather than a blank refusal to contest. Read this way, the case for institutionalising AIAs in India rests not only on preventing harm before it occurs, but on making the accountability mechanisms that already exist in Indian public law function as intended.

5.6 Data Quality and Representativeness

A further complication specific to the Indian context is the quality and representativeness of the administrative data on which many automated systems are trained. Welfare and policing datasets in India are compiled across states with differing digitisation histories, differing degrees of rural connectivity, and, in the case of policing data, differing and sometimes contested recording practices. A model trained on such data risks encoding not only historical patterns of enforcement or benefit disbursement but also the unevenness of the data infrastructure itself, so that a district with weaker record-keeping capacity may appear, algorithmically, as a district with lower need or lower risk, regardless of the underlying reality. None of the three domains examined in Section 5 currently publishes a data-quality statement alongside the automated system in question, which means this risk is, at present, essentially unmeasured. A properly resourced impact assessment template, of the kind proposed in Phase 2 of the roadmap below, would need to require an explicit accounting of data provenance and known gaps as a mandatory field, not an optional annexe, precisely because data quality issues of this kind are unlikely to be volunteered by a deploying agency unless the assessment format compels disclosure.

6. Towards an Indian Algorithmic Impact Assessment Framework: A Phased Roadmap

Building on the preceding analysis, this section proposes a three-phase institutional roadmap, sequenced so that legal authority and procedural tools are established before, or alongside, the technical capacity needed to use them meaningfully, rather than treating capacity-building as an afterthought to regulation.

Phase one, spanning an initial planning period, would establish legal authority through a Cabinet Secretariat or NITI Aayog-anchored order requiring every central ministry and department to maintain and publicly disclose an inventory of automated decision systems in use or under procurement, modelled on the disclosure obligations in Canada's Directive (Scassa, 2020). This phase requires no new technical capacity; it converts an existing, largely dormant transparency principle in NITI Aayog's 2021 documents into an operative reporting obligation (NITI Aayog, 2021).

Phase two would introduce a standard, risk-tiered assessment questionnaire, adapted from the Canadian model and the AI Now Institute's practical framework, to be completed before deployment of any system affecting eligibility, entitlement, or personal liberty (Reisman et al., 2018; Scassa, 2020). Completed assessments, and their risk tier, would be published, giving affected citizens and researchers the evidentiary basis for contestation that current systems lack. This phase would need a designated reviewing authority — plausibly an expanded mandate for an existing body such as the Data Protection Board, or a newly constituted Algorithmic Oversight Unit — with statutory power to require remediation before deployment for higher-risk tiers.

Phase three would build the technical auditing capacity necessary to verify, rather than merely receive, agency self-assessments: recruitment of data scientists and algorithm auditors into the reviewing authority, sector-specific training for administrative tribunals and courts likely to hear disputes over automated decisions, and formal channels through which independent researchers can request access to disclosed systems for verification, consistent with recommendations already made in relation to India's competition regulator (Centre for Policy Research India, n.d.). Table 3 summarises the roadmap, its primary capacity dimension, and an illustrative institutional owner for each phase.

Table 3. Phased Roadmap for Institutionalising Algorithmic Impact Assessment in India

Phase	Primary objective	Capacity dimension addressed	Illustrative institutional owner
Phase 1 — Disclosure	Mandatory public inventory of automated decision systems in central ministries	Civic accountability	Cabinet Secretariat / NITI Aayog
Phase 2 — Assessment	Standard risk-tiered pre-deployment questionnaire with published outcomes	Legal authority; procedural infrastructure	Data Protection Board / new Algorithmic Oversight Unit
Phase 3 — Verification	Independent technical audit capacity and researcher access protocols	Technical expertise	Algorithmic Oversight Unit, in coordination with sector regulators

Sources: Scassa (2020); Reisman et al. (2018); NITI Aayog (2021); Centre for Policy Research India (n.d.). Compiled by the author.

This sequencing is deliberately conservative. It does not propose a single omnibus statute modelled wholesale on the EU's AI Act, which comparative administrative law scholarship suggests would be difficult to enforce in the absence of the technical and procedural scaffolding described above (The Need for Responsible Use of AI by Public Administration, 2025). Instead, it treats disclosure as a low-cost, high-legitimacy first step that can proceed under existing executive authority, while reserving the more resource-intensive verification function for a later phase once reviewing bodies have had time to recruit and train qualified staff.

6.1 Implementation Challenges

Three risks attend this roadmap and deserve explicit acknowledgment rather than being left implicit. The first is regulatory capture at the disclosure stage: if ministries are permitted to define for themselves what counts as an "automated decision system" subject to inventory disclosure, agencies may narrow the definition to

exclude precisely the systems — welfare eligibility scoring, policing risk tools — that most warrant scrutiny. Canada's experience suggests this risk is best managed by anchoring the definition in the effect of the system on an individual's rights or entitlements rather than in its technical architecture, so that a simple rules-based eligibility filter and a machine-learning risk score are treated alike if their consequences for the citizen are comparable (Scassa, 2020).

The second risk is capacity mismatch between the volume of assessments a phase-two questionnaire regime would generate and the reviewing authority's ability to process them meaningfully. If a Data Protection Board or Algorithmic Oversight Unit is created without a corresponding increase in trained staff, published assessments may accumulate without any of them being genuinely scrutinised, reproducing at one remove the "compliance behaviour" problem Selbst (2021) identifies in the private sector. This is precisely why the roadmap places independent verification capacity in a distinct, later phase rather than assuming it will emerge automatically once a reviewing body is named.

The third risk is federal fragmentation. Many of the highest-stakes automated systems examined in Section 5 — predictive policing in particular — are deployed by state governments rather than the central government, and a Cabinet Secretariat order of the kind proposed in Phase 1 would bind central ministries but not state police departments. A durable Indian AIA framework will eventually need either a central statute capable of setting minimum standards applicable to states, comparable to the floor set by national police reform legislation in other domains, or a voluntary model framework that individual states can adopt and adapt, together with incentives — such as conditional central scheme funding — for doing so.

On resourcing and timeline, the roadmap's three phases are not intended to imply equal duration or equal cost. Phase 1 disclosure requirements are primarily an exercise in compiling and publishing information ministries already hold internally, and comparable government transparency exercises elsewhere suggest such an inventory could plausibly be compiled within a single budget cycle given ministerial cooperation. Phase 2 is more demanding, requiring the design, piloting, and iterative refinement of a standard questionnaire, plus the legal instrument establishing a reviewing authority's mandate; this is the phase most likely to require dedicated primary or subordinate legislation rather than executive order alone, given the need to create binding review powers rather than mere disclosure obligations. Phase 3, building independent technical verification capacity through recruitment and training, is realistically a multi-year undertaking, consistent with observations elsewhere that regulators face genuine difficulty competing with the private sector for scarce algorithmic-auditing talent (Centre for Policy Research India, n.d.). Treating the three phases as sequential rather than simultaneous is therefore not only an institutional design choice but a realistic acknowledgment of how long each component is likely to take to build well.

7. Conclusion

Indian public administration has moved further and faster into algorithmic decision-making than its accountability architecture has been able to follow. This is not a claim that Indian officials have been indifferent to the ethical stakes of automation — NITI Aayog's decision to publish a sequence of strategy and responsible-AI documents between 2018 and 2022, and to invite public comment on a facial recognition use case, indicates a policy establishment aware of the issues at stake. The gap identified in this paper is narrower and more tractable than a claim of indifference: it is the gap between articulating a principle and building the institution that can enforce it. NITI Aayog's strategy and responsible-AI documents demonstrate that the normative case for ethical AI is well understood within Indian policy circles; what remains absent is a binding mechanism that converts stated principle into a documented, reviewable, and publicly contestable record before an automated system is allowed to determine who receives a subsidy, who is flagged as a policing risk, or whose face is matched against a watchlist. This paper has argued that closing this gap requires attention not only to what an Algorithmic Impact Assessment statute should say, but to whether the institutions tasked

with operating it possess the technical expertise, legal authority, procedural infrastructure, and civic accountability mechanisms an AIA regime presupposes. Measured against this four-dimensional framework, India's principal deficits lie in legal authority and procedural infrastructure rather than in normative articulation, suggesting that a phased roadmap beginning with mandatory disclosure, followed by a standard assessment questionnaire, and only then by independent technical verification, offers a more institutionally realistic path than wholesale transplantation of a single foreign model. Future research might usefully extend the case analysis in this paper through primary interviews with the officials who operate DBT, predictive policing, and FRT systems, and through freedom-of-information requests seeking the technical documentation that remains, at present, the central missing input to any serious impact assessment of the automated Indian state.

REFERENCES:

1. Automated state action in India: Administrative justice, privacy and constitutional accountability. (2025). Research Square. https://www.researchsquare.com/article/rs-8078179/v1
2. Centre for Intellectual Property and Information Technology Law (CIPIT). (2023). Algorithmic impact assessment in mitigating AI harms. Strathmore University. https://cipit.strathmore.edu/algorithmic-impact-assessment-in-mitigating-ai-harms/
3. Centre for Policy Research India. (n.d.). Regulating algorithms in India: Key findings and recommendations. https://www.cprindia.org/system/tdf/working_papers/Working%20paper_Regulating%20Algorithms_2%20Sep._re%2005.pdf
4. NITI Aayog. (2018). National strategy for artificial intelligence. Government of India. https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf
5. NITI Aayog. (2021). Approach document for India: Part 1 — Principles for responsible AI. Government of India. https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf
6. NITI Aayog. (2021). Approach document for India: Part 2 — Operationalizing principles for responsible AI. Government of India. https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf
7. NITI Aayog. (2022). Responsible AI for all: Adopting the framework — A use case approach on facial recognition technology. Government of India.
8. https://www.niti.gov.in/sites/default/files/2022-11/Ai_for_All_2022_02112022_0.pdf
9. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). (2016). Official Journal of the European Union, L119, 1–88.
10. 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/wp-content/uploads/2023/04/aiareport2018.pdf
11. Scassa, T. (2020). Administrative law and the governance of automated decision-making: A critical look at Canada's Directive on Automated Decision-Making. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3722192
12. Selbst, A. D. (2021). An institutional view of algorithmic impact assessments. Harvard Journal of Law & Technology, 35(1), 117–191.
13. https://jolt.law.harvard.edu/assets/articlePDFs/v35/Selbst-An-Institutional-View-of-Algorithmic-Impact-Assessments.pdf

14. The need for responsible use of AI by public administration: Algorithmic impact assessments (AIAs) as instruments for accountability and social control. (2025). Springer. https://link.springer.com/chapter/10.1007/978-3-031-84748-6_9