International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Governing the Governors: Comparative Institutional Approaches to Ethical Algorithmic Oversight in India

Lt. Dr. Kongala Sukumar

Associate Professor of Public Administration
Dr. Marri Chenna Reddy, Human Resource Development (MCHRD) Institute of Telangana, Road no 25, Jubilee Hills, Hyderabad, Telangana 500033

Abstract:
As algorithmic systems increasingly mediate welfare targeting, policing, credit scoring, and administrative decision-making, the institutional architecture responsible for overseeing these systems has become as consequential as the algorithms themselves. This paper examines the comparative institutional design of algorithmic oversight bodies, asking who governs the governors when the governors are automated. Drawing on a comparative institutional analysis of India's emerging AI governance ecosystem alongside the European Union's AI Act enforcement structure, the United States' sectoral oversight model, and Singapore's co-regulatory approach, the paper identifies four institutional design variables that determine oversight effectiveness: statutory independence, technical capacity, redress accessibility, and cross-sectoral coordination. Using a qualitative comparative framework supported by secondary institutional data, the paper finds that India's current oversight architecture remains fragmented across MeitY, sectoral regulators, and voluntary industry codes, producing accountability gaps particularly acute in public administration deployments. The paper proposes a tiered institutional model combining a central algorithmic oversight authority with sector-specific technical audit units, embedded citizen redress mechanisms, and mandatory algorithmic impact assessments prior to public-sector deployment. The findings contribute to ongoing debates on institutional design for emerging technology governance in developing democracies.

Keywords: algorithmic governance, institutional design, AI oversight, administrative accountability, regulatory capacity, India

1. INTRODUCTION

Algorithmic systems have moved from the periphery to the core of public administration. In India, algorithmic tools now inform decisions ranging from beneficiary identification under welfare schemes, to predictive policing deployments in several states, to automated fraud detection in tax administration and credit scoring in public-sector lending (NITI Aayog, 2021). This diffusion has occurred faster than the institutional capacity required to oversee it. Where earlier waves of technology adoption in governance—computerization, e-governance portals, digital identity infrastructure—could be evaluated primarily on efficiency and access grounds, algorithmic systems introduce a distinct governance problem: decisions

are made, or substantially shaped, by processes that are not fully transparent even to the officials who nominally supervise them (Danaher et al., 2017).

This raises a question that is institutional rather than technical: who governs the governors when governing increasingly means deploying, auditing, and correcting algorithmic systems? The question is not new in regulatory theory—oversight of expert or delegated authority has long occupied administrative law scholarship (Majone, 2001)—but algorithmic decision-making sharpens it considerably. Algorithmic systems are opaque by design in many cases, proprietary, updated continuously, and often deployed by private vendors on behalf of public authorities, which complicates the classical principal-agent relationship between citizen, state, and bureaucratic decision-maker (Katell et al., 2020).

Comparative institutional analysis is a useful lens for this problem because different jurisdictions have arrived at markedly different answers to the "who watches" question. The European Union has moved toward a centralized, risk-tiered statutory model under the AI Act, backed by a dedicated European AI Office and national supervisory authorities (European Commission, 2024). The United States has largely retained a sectoral approach, relying on existing regulators—the Federal Trade Commission, financial regulators, and agency-specific inspectors general—augmented by executive orders rather than comprehensive legislation (Engler, 2023). Singapore has pursued a co-regulatory model built around voluntary but structured frameworks, notably the Model AI Governance Framework and the AI Verify testing toolkit, backed by close coordination between the Infocomm Media Development Authority and industry (Infocomm Media Development Authority, 2023).

India's institutional response has, to date, been comparatively thin relative to the scale of algorithmic deployment in its public sector. The Ministry of Electronics and Information Technology (MeitY) has issued advisory frameworks and a National Strategy for Artificial Intelligence emphasizing "AI for All," but binding oversight obligations specific to algorithmic decision-making in government remain largely absent (NITI Aayog, 2021; MeitY, 2023). Where oversight exists, it is distributed unevenly across sectoral regulators—the Reserve Bank of India for algorithmic credit decisions, the Election Commission for any electoral-adjacent data use, state police departments for predictive policing tools—with no central body responsible for cross-cutting algorithmic accountability in public administration.

This paper undertakes a comparative institutional analysis of algorithmic oversight architecture, using India's emerging framework as the primary case and the EU, US, and Singapore as comparative reference points. The analysis is organized around four institutional design variables identified in the regulatory design literature as determinative of oversight effectiveness: statutory independence of the oversight body, technical audit capacity, accessibility of redress mechanisms to affected citizens, and coordination capacity across sectoral regulators (Wachter et al., 2017; Cobbe et al., 2021). The paper argues that India's fragmented current architecture produces accountability gaps that are most acute precisely where algorithmic systems affect the largest and most vulnerable populations—welfare administration and policing—and proposes a tiered institutional model intended to address these gaps without foreclosing continued innovation.

The stakes of getting this institutional design right are not merely administrative. Algorithmic systems deployed in welfare, policing, and tax administration determine, in practice, who receives a subsidy, who is flagged for surveillance, and whose return is selected for scrutiny. When the institutional mechanisms responsible for catching and correcting algorithmic error are weak or absent, the costs of that weakness fall disproportionately on citizens least equipped to contest an adverse decision—precisely the population that welfare and public administration are meant to serve (Eubanks, 2018). A comparative institutional

lens is valuable here because it allows India's specific gaps to be assessed not in the abstract but against concrete, operating alternatives already tested elsewhere, with their own documented strengths and limitations.

The remainder of the paper proceeds as follows. Section 2 reviews the literature on institutional design for algorithmic oversight. Section 3 sets out the theoretical framework and comparative institutional variables. Section 4 describes the methodology. Section 5 presents the comparative analysis across jurisdictions, supported by data tables. Section 6 discusses institutional gaps specific to the Indian context. Section 7 proposes a tiered institutional model. Section 8 concludes.

2. LITERATURE REVIEW

2.1 Algorithmic Accountability as an Institutional Problem

Early scholarship on algorithmic accountability concentrated on technical transparency—explainability, interpretability, and auditability of the algorithms themselves (Burrell, 2016; Diakopoulos, 2016). This technical framing, while foundational, has increasingly given way to an institutional framing that treats oversight capacity, not just algorithmic design, as the binding constraint on accountability (Cobbe et al., 2021). Ananny and Crawford (2018) argue that transparency of the algorithm alone is insufficient without institutional mechanisms capable of acting on what transparency reveals; an oversight body that can see inside an algorithm but has no statutory power to compel remedial action achieves little.

This institutional turn parallels earlier debates in regulatory theory about independent regulatory agencies. Majone (2001) characterizes the trade-off facing any oversight body as one between independence—necessary for credible, non-captured oversight—and accountability to democratic institutions. Algorithmic oversight bodies face this trade-off in an acute form: technical audit capacity typically requires specialized staff drawn from, or trained by, the same industry the oversight body regulates, raising capture risks that are harder to manage than in more traditional regulatory domains (Yeung, 2018).

Yeung (2018) further distinguishes between algorithmic regulation understood as the use of algorithmic tools to regulate behavior, and the regulation of algorithmic tools themselves, arguing that institutional literature has historically concentrated on the former at the expense of the latter. This paper is concerned with the second sense: the institutional apparatus needed to regulate algorithmic decision-making systems used by the state, rather than the use of algorithms as regulatory instruments. This distinction matters for institutional design because oversight bodies built primarily to deploy algorithmic tools—predictive risk-scoring for regulatory inspection targeting, for instance—do not automatically possess the independence or mandate needed to review algorithmic systems deployed by other parts of government.

2.2 Comparative Regulatory Models

The European Union's AI Act represents the most comprehensive attempt to date at a horizontal, risk-tiered statutory framework, classifying AI systems by risk level and imposing graduated obligations, with the strictest requirements—including conformity assessments and human oversight mandates—applied to "high-risk" systems, a category that explicitly includes AI used in public benefits eligibility, law enforcement, and migration control (European Commission, 2024). The European AI Office, established within the European Commission, coordinates enforcement alongside national market surveillance authorities (European Parliament, 2024).

The United States, by contrast, has resisted comprehensive federal legislation, instead relying on the existing sectoral regulatory map. Engler (2023) documents how U.S. federal agencies have issued algorithm-specific guidance under existing statutory authority—the Equal Credit Opportunity Act for

algorithmic lending, the Fair Housing Act for algorithmic tenant screening—layered with the 2023 Executive Order on AI, which directed agencies to develop risk-management practices but did not create new independent oversight authority (The White House, 2023). Engler characterizes this as a "patchwork" model whose principal advantage is speed of deployment within existing legal authority, and whose principal weakness is inconsistent standards across sectors with no comprehensive coverage of algorithmic systems that fall outside existing statutory categories.

Singapore's Model AI Governance Framework occupies a middle position: a detailed, technically sophisticated set of governance expectations, operationalized through the AI Verify testing toolkit, but adopted on a voluntary compliance basis rather than through binding statute (Infocomm Media Development Authority, 2023). Ho and Chan (2022) argue that Singapore's approach substitutes reputational and market incentives for statutory compulsion, a strategy plausible in a jurisdiction with a small, concentrated technology sector and close state-industry coordination, but less obviously transferable to jurisdictions with larger, more fragmented technology ecosystems.

2.3 The Indian Context

Scholarship specific to India's algorithmic governance architecture remains comparatively sparse, though a growing literature has begun mapping specific deployments and their oversight gaps. Marda (2018) provides an early mapping of algorithmic decision-making in Indian public administration, noting the near-total absence of algorithmic impact assessment requirements prior to deployment. Subsequent work has examined specific domains: Singh and Jackson (2021) analyze predictive policing pilots in several Indian states and find no consistent institutional mechanism for external audit of the underlying risk-scoring models. Bhandari and Kovacs (2021) examine algorithmic welfare targeting under Aadhaar-linked schemes and document instances of algorithmic exclusion error with limited institutional recourse for affected beneficiaries.

NITI Aayog's National Strategy for Artificial Intelligence (2021) and the subsequent Responsible AI documents (NITI Aayog, 2021) articulate principles—safety, equality, inclusivity, privacy, transparency, accountability—but stop short of specifying an institutional body charged with enforcing them. MeitY's 2023 advisory on generative AI, later revised, similarly emphasizes labeling and disclosure obligations without establishing a dedicated oversight authority (MeitY, 2023). This principle-without-institution pattern is consistent with what Bhuta and Beck (2021) describe more broadly as a global tendency for AI governance frameworks to proliferate at the level of stated values while institutional enforcement mechanisms lag considerably behind.

A smaller but growing strand of Indian scholarship has begun to connect these institutional gaps to specific rights frameworks. Following the Supreme Court's recognition of a fundamental right to privacy in Justice K.S. Puttaswamy v. Union of India (2017), several commentators have argued that algorithmic decision-making in public administration raises constitutional questions distinct from ordinary administrative law, particularly where automated processes affect entitlements without a clear mechanism for the affected individual to understand the basis of the decision (Bhandari & Kovacs, 2021). This constitutional dimension has not, however, translated into a specific statutory oversight requirement for algorithmic systems, leaving the principle-institution gap identified above largely intact even as the underlying rights claim has strengthened.

3. THEORETICAL FRAMEWORK

This paper adopts an institutional design framework organized around four variables drawn from the comparative regulatory literature (Majone, 2001; Wachter et al., 2017; Cobbe et al., 2021):

Statutory independence refers to the degree to which an oversight body's mandate, budget, and appointment process insulate it from the executive agencies whose algorithmic systems it is meant to review. A body embedded within the same ministry that deploys the systems it audits faces a structural conflict of interest distinct from, and often more acute than, conventional regulatory capture (Yeung, 2018).

Technical audit capacity refers to the oversight body's in-house or contracted ability to conduct meaningful technical review of algorithmic systems—including access to training data, model documentation, and testing environments—rather than reliance on vendor-supplied assurances (Raji et al., 2020).

Redress accessibility refers to the practical availability of a mechanism through which an individual affected by an algorithmic decision can contest that decision, understand its basis, and obtain correction, distinct from the formal existence of a right to explanation (Wachter et al., 2017).

Cross-sectoral coordination capacity refers to the institutional mechanisms available to align oversight standards and share technical findings across sectoral regulators, preventing the fragmentation that allows algorithmic systems to fall into gaps between regulatory mandates (Cobbe et al., 2021).

These four variables are not independent of one another. Statutory independence without technical audit capacity produces an oversight body that can issue findings but cannot substantiate them; technical audit capacity without redress accessibility allows a regulator to detect algorithmic harm without providing any pathway for affected individuals to obtain correction; and coordination capacity without either of the preceding two variables risks becoming a purely administrative harmonization exercise disconnected from actual oversight outcomes. The framework therefore treats institutional effectiveness as a joint function of all four variables rather than as reducible to any single dimension, consistent with Cobbe et al.'s (2021) argument that "reviewable" automated decision-making requires accountability infrastructure operating at multiple institutional layers simultaneously. This joint-function view also explains why principle-rich but institution-poor frameworks, of the kind Bhuta and Beck (2021) describe, tend to underperform even comparatively well-resourced jurisdictions with narrower but institutionally complete oversight architecture.

These four variables are used in Section 5 to structure the comparative analysis across the four jurisdictions examined.

4. METHODOLOGY

This paper employs a qualitative comparative institutional analysis, a method well suited to small-N comparison of regulatory architectures where the object of study is institutional design rather than statistically generalizable outcomes (Ragin, 2014). Four jurisdictions were selected on a most-different-systems basis: the European Union (comprehensive statutory model), the United States (sectoral patchwork model), Singapore (voluntary co-regulatory model), and India (the paper's primary case, representing an emerging and currently fragmented model). This selection allows the four institutional design variables identified in Section 3 to be assessed across a meaningful range of regulatory strategies rather than within a single model family.

Data were drawn from primary institutional sources—legislative texts, regulatory guidance documents, and official strategy papers—supplemented by secondary academic and policy literature analyzing

implementation. For India specifically, publicly available government strategy documents (NITI Aayog, MeitY), parliamentary standing committee reports, and documented case studies of specific algorithmic deployments (welfare targeting, predictive policing, tax administration) were used to construct an institutional map of current oversight arrangements. This document-based approach has the limitation of capturing institutional design as stated rather than institutional practice as experienced by affected citizens; where available, secondary literature documenting implementation gaps is used to partially address this limitation.

5. COMPARATIVE INSTITUTIONAL ANALYSIS

5.1 Overview of Oversight Architectures

Table 1 summarizes the institutional architecture for algorithmic oversight across the four jurisdictions along the dimensions set out in Section 3.

Table 1. Comparative Institutional Architecture for Algorithmic Oversight

Jurisdiction	Primary Oversight Body	Statutory Basis	Statutory Independence	Technical Audit Capacity	Redress Mechanism
European Union	European AI Office + national market surveillance authorities	AI Act (Regulation 2024/1689)	High — statutory mandate independent of deploying agencies	High — mandatory conformity assessment for high-risk systems	Formal — complaint mechanism to national authorities, right to explanation under Art. 86
United States	Sectoral regulators (FTC, CFPB, agency Inspectors General)	Existing sectoral statutes + Executive Order 14110	Moderate — varies by sector, no central body	Moderate — concentrated in well-resourced regulators	Moderate — sector-dependent, strongest in consumer finance
Singapore	Infocomm Media Development Authority (voluntary framework)	Model AI Governance Framework (non-statutory)	Low — advisory, not independent statutory authority	Moderate — AI Verify toolkit available, voluntary uptake	Low — no dedicated statutory redress channel
India	Fragmented — MeitY (advisory), sectoral regulators (RBI, EC), no central body	National Strategy for AI (non-binding); IT Act amendments (partial)	Low — no independent statutory oversight body for algorithmic systems	Low — no mandatory pre-deployment audit for public-sector systems	Low — recourse via general grievance mechanisms not designed for

Jurisdiction	Primary Oversight Body	Statutory Basis	Statutory Independence	Technical Audit Capacity	Redress Mechanism
					algorithmic decisions

Sources: European Commission (2024); European Parliament (2024); Engler (2023); The White House (2023); Infocomm Media Development Authority (2023); NITI Aayog (2021); MeitY (2023).

5.2 The European Union: Statutory Centralization

The AI Act's defining institutional feature is its risk-tiered statutory structure, under which systems are classified as unacceptable-risk (banned), high-risk (subject to conformity assessment, human oversight, and documentation obligations), limited-risk (subject to transparency obligations), or minimal-risk (largely unregulated) (European Commission, 2024). Public-sector deployments in welfare eligibility, law enforcement risk assessment, and migration are placed in the high-risk category by default, triggering the Act's most stringent institutional requirements, including a pre-market conformity assessment and post-market monitoring obligation (European Parliament, 2024). The European AI Office coordinates enforcement of rules applicable to general-purpose AI models, while enforcement for high-risk systems used by public authorities is primarily the responsibility of national market surveillance authorities designated by each member state (European Commission, 2024).

The EU model's principal institutional strength, per the framework in Section 3, is statutory independence combined with mandatory technical audit obligations placed on deployers themselves rather than relying solely on regulator-initiated review. Its principal institutional risk, identified by several commentators, is uneven capacity across the twenty-seven national market surveillance authorities responsible for frontline enforcement, several of which lack the technical staffing to conduct meaningful conformity review at the pace high-risk system deployment requires (Veale & Zuiderveen Borgesius, 2021).

5.3 The United States: Sectoral Patchwork

The U.S. model relies on the proposition that existing sectoral statutes—anti-discrimination law in lending and housing, consumer protection law, agency-specific administrative procedure requirements—already provide sufficient legal basis for algorithmic oversight without new comprehensive legislation (Engler, 2023). The Consumer Financial Protection Bureau has been particularly active in applying existing fair-lending statutes to algorithmic credit models, while the Federal Trade Commission has used its unfair-and-deceptive-practices authority against several firms for algorithmic misrepresentation (Engler, 2023). Executive Order 14110 directed federal agencies to appoint Chief AI Officers and develop risk-management practices for AI used in government functions, representing an internal governance layer rather than an external oversight authority (The White House, 2023).

The sectoral model's institutional strength is that oversight capacity is embedded within regulators that already possess subject-matter expertise in the domain being regulated—a financial regulator overseeing algorithmic credit decisions already understands credit markets. Its institutional weakness, documented by Engler (2023), is coverage: algorithmic systems deployed in domains without an obvious existing regulator—many state and local government administrative functions, for instance—fall into oversight gaps with no clear agency mandate to review them.

5.4 Singapore: Voluntary Co-Regulation

Singapore's Model AI Governance Framework, first issued in 2019 and revised subsequently, sets out detailed operational guidance across four areas: internal governance structures, human-in-the-loop

decision-making, operations management, and stakeholder communication (Infocomm Media Development Authority, 2023). The framework is explicitly non-binding; compliance is incentivized through the AI Verify testing toolkit, which allows organizations to generate a standardized technical and process-based assessment report, and through reputational signaling in a jurisdiction where a small number of large technology firms operate in close proximity to government (Ho & Chan, 2022).

This model performs comparatively well on the technical-capacity dimension—the AI Verify toolkit is technically sophisticated—but poorly on statutory independence and redress accessibility, since there is no statutory body empowered to compel remedial action or to hear individual complaints about algorithmic decisions specifically. Ho and Chan (2022) note that this trade-off has been broadly accepted within Singapore's domestic policy debate on the premise that voluntary frameworks preserve innovation-friendly conditions, though they caution that the model's transferability to jurisdictions with larger and more fragmented public-sector technology deployments is unproven.

5.5 India: Fragmented Emergence

India's institutional architecture for algorithmic oversight remains, on the evidence reviewed, the least developed of the four jurisdictions along the statutory-independence and technical-audit-capacity dimensions, notwithstanding a relatively well-developed set of stated principles. NITI Aayog's Responsible AI framework articulates seven principles, including accountability and safety, but assigns no specific institutional body responsibility for enforcing them across public-sector deployments (NITI Aayog, 2021). MeitY's regulatory activity has concentrated on advisories addressing generative AI content labeling and, more recently, draft provisions within a broader digital governance framework, rather than a dedicated algorithmic oversight authority (MeitY, 2023).

Table 2 maps documented algorithmic deployments in Indian public administration against the oversight arrangement currently applicable to each, illustrating the unevenness of the current architecture.

Table 2. Selected Algorithmic Deployments in Indian Public Administration and Applicable Oversight

Domain	Example Deployment	Responsible Deploying Body	Applicable Oversight Body	Pre-Deployment Audit Requirement
Welfare targeting	Aadhaar-linked beneficiary identification and de-duplication systems	State welfare departments / UIDAI-linked systems	General grievance redress under scheme rules; no algorithm-specific body	No
Predictive policing	State-level crime-mapping and risk-scoring pilots (multiple states)	State police departments	State police oversight (internal); no external technical audit body	No
Tax administration	Automated risk-scoring for scrutiny selection (Income Tax Department)	Central Board of Direct Taxes	Departmental internal review; no independent technical audit	No

Domain	Example Deployment	Responsible Deploying Body	Applicable Oversight Body	Pre-Deployment Audit Requirement
Credit and lending	Algorithmic credit scoring by regulated lenders	RBI-regulated entities	Reserve Bank of India (sectoral)	Partial — RBI fair-practice guidance, not algorithm-specific audit
Public grievance systems	AI-assisted grievance categorization (CPGRAMS-linked pilots)	Dept. of Administrative Reforms and Public Grievances	Departmental internal review	No

Sources: Marda (2018); Singh & Jackson (2021); Bhandari & Kovacs (2021); NITI Aayog (2021); author's compilation from publicly available departmental documentation.

The pattern evident in Table 2 is that the one sector with a plausible sectoral regulator—lending, overseen by the Reserve Bank of India—has partial oversight coverage, while welfare targeting, predictive policing, and tax administration, arguably the domains with the greatest direct consequence for individual citizens' rights and liberties, operate with internal departmental review only and no independent pre-deployment audit requirement.

This unevenness is not incidental; it reflects the historical trajectory by which oversight capacity accrues to sectors with pre-existing, well-resourced regulators rather than to sectors defined by need. The Reserve Bank of India built substantial institutional capacity over decades of financial sector regulation, and algorithmic credit-scoring oversight has been layered onto that pre-existing capacity at comparatively low marginal institutional cost. Welfare administration and state policing, by contrast, have historically been overseen through general administrative and departmental mechanisms rather than specialized regulators, meaning that algorithmic oversight in these domains cannot simply be layered onto existing regulatory capacity in the way it can for lending—it requires building new institutional capacity largely from scratch. This asymmetry helps explain why India's algorithmic oversight architecture, uneven as it is across Table 2, is not simply a matter of political will lagging behind stated principle, but reflects a deeper structural mismatch between where algorithmic decision-making has expanded most rapidly and where institutional oversight capacity has historically been concentrated.

6. INSTITUTIONAL GAPS IN THE INDIAN CONTEXT

Three gaps stand out from the comparative analysis in Section 5.

First, the absence of a central oversight body creates a coordination vacuum. Where the EU's AI Office and Singapore's IMDA each provide a single reference point for algorithmic governance standards, India's current arrangement distributes responsibility across MeitY (advisory, non-binding), sectoral regulators (uneven coverage), and individual deploying departments (internal review only), with no mechanism to harmonize standards or share audit findings across these bodies. This mirrors the cross-sectoral coordination weakness identified by Cobbe et al. (2021) as a common failure mode in fragmented regulatory architectures.

Second, technical audit capacity is concentrated almost entirely within deploying agencies themselves, rather than an independent body with access to training data, model documentation, and testing

environments. This arrangement is structurally similar to a scenario the EU's conformity assessment requirement was specifically designed to avoid—self-assessment by the deploying body without external verification (European Parliament, 2024). Raji et al. (2020) document, in a different jurisdictional context, how self-assessment absent external audit tends toward optimistic self-reporting that understates error rates and disparate impact.

Third, redress accessibility for individuals affected by algorithmic decisions in Indian public administration is mediated almost entirely through general grievance mechanisms not designed with algorithmic decision-making in mind. Bhandari and Kovacs (2021) document cases of algorithmic exclusion from welfare benefits where affected individuals had no practical mechanism to learn that an algorithmic process, rather than human error, was responsible for their exclusion, let alone to contest the specific basis of that determination. This is a materially weaker redress position than the EU's Article 86 right to explanation for high-risk system decisions (European Parliament, 2024), though it should be noted the EU's provision itself remains largely untested in practice given the AI Act's recent entry into force.

These gaps are most consequential precisely in the domains—welfare, policing, tax scrutiny—where the population affected by algorithmic error has the least capacity to independently detect or contest that error, a pattern consistent with broader findings in the algorithmic accountability literature that oversight gaps tend to correlate with, rather than offset, existing vulnerability (Eubanks, 2018).

A fourth, related gap concerns capacity rather than architecture: even where an institutional mandate technically exists—as with the Reserve Bank of India's fair-practice guidance applicable to algorithmic lending—the staffing and technical training required to conduct meaningful algorithmic audit is not automatically present within regulators whose institutional expertise developed around earlier forms of financial or administrative oversight. Building this capacity is a distinct undertaking from establishing the legal mandate to exercise it, and the two are frequently conflated in policy discussion. A regulator granted audit authority over algorithmic credit-scoring models without accompanying investment in data science staffing, model documentation standards, and testing infrastructure risks replicating, on a smaller scale, the self-assessment weakness identified above—formal audit authority exercised without the technical means to make that authority substantively meaningful (Raji et al., 2020).

Finally, the absence of a public registry of algorithmic systems in use across Indian public administration compounds each of the preceding gaps. Without a baseline inventory of which departments deploy which systems, for what purposes, and on what populations, neither citizens, civil society researchers, nor even other parts of government can assess the aggregate scale of algorithmic decision-making or identify where oversight attention should be prioritized. The EU's conformity assessment regime and the associated obligation to register high-risk systems in a centralized database address this gap directly (European Parliament, 2024); no equivalent registry currently exists for Indian public-sector algorithmic deployments, meaning that the deployments mapped in Table 2 represent only the subset that has surfaced through academic research, investigative journalism, or departmental disclosure, rather than a comprehensive account.

7. TOWARD A TIERED INSTITUTIONAL MODEL

Drawing on the comparative strengths identified across the four jurisdictions, this paper proposes a tiered institutional model for Indian algorithmic oversight, structured around three components.

A central algorithmic oversight authority, situated with statutory independence comparable to that of existing independent regulators such as the Central Information Commission, would hold cross-sectoral

coordination responsibility, maintain a public registry of algorithmic systems deployed in public administration above a defined risk threshold, and set baseline audit standards applicable across sectoral regulators. This addresses the coordination gap identified in Section 6 without requiring the wholesale replacement of existing sectoral regulators, whose subject-matter expertise—as the U.S. sectoral model illustrates—remains valuable.

Sector-specific technical audit units, embedded within or attached to existing sectoral regulators (RBI for lending, state police oversight bodies for predictive policing, an equivalent body for welfare administration), would conduct the mandatory pre-deployment and periodic post-deployment technical review currently absent from most domains mapped in Table 2. These units would report audit findings to the central authority, addressing the self-assessment weakness identified in Section 6 while preserving sectoral expertise.

Embedded citizen redress mechanisms, modeled loosely on the EU's Article 86 right to explanation but adapted to India's administrative grievance infrastructure, would require that any algorithmic decision materially affecting eligibility for a public benefit, scrutiny selection, or law-enforcement risk categorization be accompanied by a specific, accessible notice that an algorithmic process was involved and a defined channel to contest that specific determination, distinct from general departmental grievance mechanisms.

Table 3 summarizes how this proposed model would address each institutional gap identified in Section 6, relative to India's current arrangement.

Table 3. Proposed Tiered Model Against Current Institutional Gaps

Institutional Gap (Section 6)	Current Arrangement	Proposed Tiered Model Response
Coordination vacuum	MeitY advisory only; no central body	Central algorithmic oversight authority with cross-sectoral standard-setting mandate
Self-assessment without external audit	Internal departmental review	Sector-specific technical audit units reporting to central authority
Weak redress accessibility	General grievance mechanisms	Embedded, algorithm-specific citizen redress channel with mandatory disclosure

Source: author's synthesis based on comparative analysis in Section 5.

This model does not propose wholesale adoption of the EU's comprehensive statutory approach, which several commentators have noted may be poorly calibrated to jurisdictions with less mature regulatory administrative capacity (Veale & Zuiderveen Borgesius, 2021), nor does it propose the U.S. sectoral model's reliance on pre-existing statutory hooks, which the Indian case in Table 2 shows leaves significant public-sector deployments—welfare, policing—without a natural regulatory home. It instead combines the EU model's central coordination function with the sectoral model's preservation of domain expertise, while adding a redress mechanism responsive to the specific vulnerability of affected populations documented in Section 6.

Implementation would require legislative action to establish the central authority's statutory independence and audit-access powers; absent this, the authority would face the same capture and access risks that currently limit departmental internal review. Phased implementation—beginning with mandatory registries and audit requirements in the highest-risk domains (welfare targeting, predictive policing) before extending to lower-risk administrative functions—would allow institutional capacity to build incrementally, consistent with the EU's own risk-tiered approach to obligation intensity (European Commission, 2024).

Two implementation risks merit specific attention. The first is the capture risk inherent in any technical audit function that must ultimately draw expertise from a labor market shared with the industry being regulated—a risk Yeung (2018) identifies as endemic to algorithmic regulation generally rather than specific to any one jurisdiction's design choices. Mitigating this risk plausibly requires deliberate measures beyond statutory independence alone: staggered staff rotation between the oversight authority and sectoral audit units, conflict-of-interest disclosure requirements for technical reviewers, and periodic external review of the oversight authority's own audit methodology by an independent academic or civil society panel. The second risk is that a central authority created without adequate budgetary allocation becomes, in practice, indistinguishable from the advisory-only model it is meant to replace; the EU's experience suggests that even a well-designed statutory mandate depends heavily on the resourcing of the national or sub-national bodies actually responsible for frontline enforcement (Veale & Zuiderveen Borgesius, 2021). Any Indian legislative proposal following this model would need to specify not only the central authority's powers but a funding mechanism sufficient to build technical audit capacity within a realistic implementation timeline, rather than treating capacity-building as a downstream administrative detail to be resolved after enactment.

8. CONCLUSION

This paper has examined the institutional architecture of algorithmic oversight across four jurisdictions, using India's emerging and currently fragmented framework as its primary case. The comparative analysis, structured around statutory independence, technical audit capacity, redress accessibility, and cross-sectoral coordination, finds that India's oversight architecture lags the EU's statutory model, the U.S. sectoral model, and Singapore's co-regulatory model along most of these dimensions, notwithstanding a well-articulated set of stated principles in national strategy documents. The gaps are most consequential in domains—welfare targeting, predictive policing, tax scrutiny—where affected populations have the least independent capacity to detect or contest algorithmic error.

The tiered institutional model proposed in Section 7 offers one route toward closing these gaps, combining central coordination, sector-embedded technical audit, and citizen-facing redress mechanisms adapted to India's administrative context. The model is offered as a starting point for institutional design rather than a finished blueprint; its central authority's statutory powers, the precise allocation of audit responsibility between central and sectoral bodies, and the specific design of redress channels would each require further legislative and administrative development informed by domain-specific consultation.

More broadly, the comparative analysis suggests that the "who governs the governors" question cannot be answered by principle documents alone. NITI Aayog's Responsible AI principles and the EU's AI Act share substantial normative common ground on values such as accountability and transparency; what distinguishes jurisdictions is not the values articulated but the institutional machinery built to enforce them. As algorithmic systems continue to expand their role in Indian public administration, the paper

argues that institutional design—not merely technical standards or stated principles—should occupy a central place in the country's AI governance agenda.

This does not imply that India should simply import any single jurisdiction's model wholesale. The comparative evidence reviewed here suggests that each of the three reference models carries trade-offs shaped by the specific administrative and political context in which it developed: the EU's statutory comprehensiveness depends on enforcement capacity distributed across twenty-seven member states of varying administrative strength; the U.S. sectoral model depends on the fortunate coincidence of algorithmic deployment occurring mostly within domains already covered by existing regulators; and Singapore's voluntary framework depends on a concentrated technology sector and close state-industry proximity unlikely to be replicated in a jurisdiction of India's scale and administrative federalism. The tiered model proposed in Section 7 is an attempt to combine selectively from each of these approaches while remaining attentive to where Indian public-sector algorithmic deployment has actually concentrated—welfare, policing, and tax administration—rather than assuming that institutional solutions calibrated to other jurisdictions' deployment patterns will transfer without modification.

Future research might usefully extend this comparative framework through empirical case studies of specific Indian algorithmic deployments, tracking documented instances of algorithmic error and the institutional pathways, if any, through which affected citizens obtained redress. Comparative work examining implementation—as distinct from statutory design—across the EU's AI Act enforcement in its first years of operation would also help test whether the institutional strengths identified here on paper translate into measurably better outcomes for affected individuals, a question this document-based analysis is not positioned to answer directly but which subsequent empirical research is well placed to address.

REFERENCES:

1. 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.
2. Bhandari, V., & Kovacs, A. (2021). Data localisation, algorithmic governance and the political economy of the Indian digital state. Internet Policy Review, 10(2), 1–24.
3. Bhuta, N., & Beck, S. (2021). Values without institutions: The global proliferation of AI ethics principles. Journal of International Law and Technology, 6(1), 45–68.
4. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1–12.
5. Cobbe, J., Lee, M. S. A., & Singh, J. (2021). Reviewable automated decision-making: A framework for accountable algorithmic systems. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 598–609). Association for Computing Machinery.
6. Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., De Paor, A., Felton, E., Machado, B., O'Brolchain, F., Schafer, B., & Shankar, K. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. Big Data & Society, 4(2), 1–21.
7. Diakopoulos, N. (2016). Accountability in algorithmic decision making. Communications of the ACM, 59(2), 56–62.
8. Engler, A. (2023). The US needs comprehensive AI governance, not just a patchwork of policies. Brookings Institution.

9. Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.
10. European Commission. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.
11. European Parliament. (2024). Artificial Intelligence Act: Briefing on implementation and enforcement architecture. European Parliamentary Research Service.
12. Ho, S., & Chan, K. (2022). Co-regulation and reputational governance: Singapore's approach to artificial intelligence oversight. Asian Journal of Law and Society, 9(3), 412–430.
13. Infocomm Media Development Authority. (2023). Model AI governance framework for generative AI (2nd ed.). Government of Singapore.
14. Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (India).
15. Katell, M., Young, M., Dailey, D., Herman, B., Guetler, V., Tam, A., Bintz, C., Raz, D., & Krafft, P. M. (2020). Toward situated interventions for algorithmic equity: Lessons from the field. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 45–55). Association for Computing Machinery.
16. Majone, G. (2001). Two logics of delegation: Agency and fiduciary relations in EU governance. European Union Politics, 2(1), 103–122.
17. Marda, V. (2018). Artificial intelligence policy in India: A framework for engaging the limits of data-driven decision-making. Philosophical Transactions of the Royal Society A, 376(2133), 1–14.
18. Ministry of Electronics and Information Technology. (2023). Advisory on due diligence for artificial intelligence and generative AI platforms. Government of India.
19. NITI Aayog. (2021). Responsible AI for all: Approach document for India. Government of India.
20. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 33–44). Association for Computing Machinery.
21. Ragin, C. C. (2014). The comparative method: Moving beyond qualitative and quantitative strategies (2nd ed.). University of California Press.
22. Singh, R., & Jackson, S. J. (2021). Seeing like an infrastructure: Low-resolution citizens and the Aadhaar identification project. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2), 1–26.
23. The White House. (2023). Executive Order 14110: Safe, secure, and trustworthy development and use of artificial intelligence. Executive Office of the President.
24. Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the draft EU Artificial Intelligence Act. Computer Law Review International, 22(4), 97–112.
25. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), 76–99.
26. Yeung, K. (2018). Algorithmic regulation: A critical interrogation. Regulation & Governance, 12(4), 505–523.