

Who Governs the Machine?

Authority, accountability and institutional control when artificial intelligence begins to act.

A bounded inquiry into institutional control.

This dossier examines laws, treaties, official policies, administrative directives and risk-management frameworks. It distinguishes documented mechanisms, Foundation inference and open questions. It does not evaluate a named deployment or determine legal compliance.

INSTITUTIONAL AUTHOR

Nataraja Labs Research Foundation

PUBLICATION STATE

Released public dossier

GEOGRAPHY

India in comparative institutional context

EVIDENCE CUT-OFF

2 September 2026

EDITION

Version 1.0 · first public release

CITATION

Nataraja Labs Research Foundation (2026), Who Governs the Machine?, Dossier 002, version 1.0.

When a machine enters a chain of command, who still possesses the power to say no?

Artificial intelligence becomes an institutional question when its output enters a chain of authority: deciding what receives attention, what is classified as risk, which action is recommended, what is executed and how an affected person may challenge the result.

Governance begins before deployment—where purpose, permission, evidence and recourse are designed.

F/01

STRONG INFERENCE

Authority may be delegated.
Responsibility cannot be automated away.

Across the examined frameworks, responsibility remains attributable to people or legal entities even when AI supplies the recommendation, ranking, prediction or action pathway.

F/02

DOCUMENTED

A human in the loop is not, by itself, meaningful oversight.

Effective oversight requires competence, information, time and authority to disregard, reverse or stop the system.

F/03

STRONG INFERENCE

The consequential object is the decision system, not the model alone.

Risk emerges from mandate, data, model, interface, workflow, operator, affected population, vendor and remedy.

F/04

DOCUMENTED

Inventories and public records are becoming governance infrastructure.

Governments now require or maintain use-case inventories, transparency records or published impact assessments so systems can be located before scrutiny.

F/05

STRONG INFERENCE

Recourse is part of system design.

Notice, explanation, complaint, appeal and correction determine what records must exist and whether an adverse outcome can be reversed.

F/06

DOCUMENTED

Procurement determines institutional sovereignty.

Portability, data rights, documentation, testing, subcontractor visibility, licensing and continuity are material AI controls.

F/07

STRONG INFERENCE

Autonomy changes the containment architecture.

When a system can use tools, communicate, alter records or initiate transactions, governance must constrain permissions and preserve actions.

F/08

STRONG INFERENCE

India's emerging architecture is layered and institution-building.

The examined record combines data protection, national guidelines, proposed institutions and sector-specific supervisory frameworks.

Ten layers between purpose and termination

01

Mandate

What lawful or organisational purpose permits the system to exist?

Named owner and bounded purpose.

02

Classification

What changes if the system is wrong, manipulated or unavailable?

Context-specific impact tier and prohibited uses.

03

Acquisition

What control is surrendered to providers and subcontractors?

Portability, testing access and exit terms.

04

Data

Which records shape the system and who may use them?

Provenance, lawful basis, quality, retention and access.

05

Model & tools

What may the system infer, retrieve, communicate or execute?

Approved models, connected tools and permission boundaries.

Ten layers between purpose and termination

06

Human authority

Who may accept, reject, suspend or reverse an output?

Competent officials with real authority.

07

Evidence

Can the institution reconstruct what happened?

Versioned inputs, outputs, actions, approvals and exceptions.

08

Monitoring

How will drift, bias, misuse and control failure become visible?

Metrics, testing, incident thresholds and review cadence.

09

Recourse

How can an affected person understand and challenge the result?

Notice, explanation, complaint, appeal and correction.

10

Termination

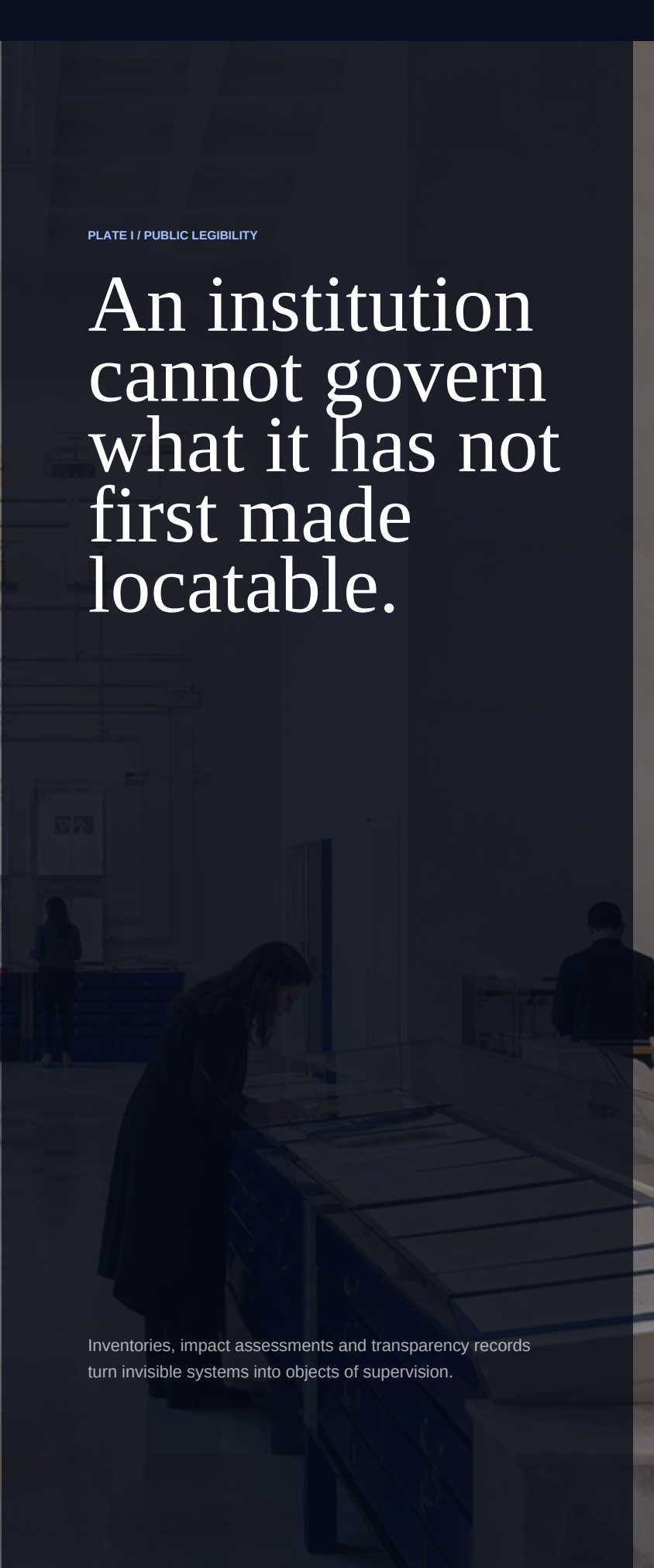
Can the institution stop the system and continue its mission?

Kill switch, fall-back process and vendor exit.



PLATE I / PUBLIC LEGIBILITY

An institution cannot govern what it has not first made locatable.



Inventories, impact assessments and transparency records turn invisible systems into objects of supervision.

Different jurisdictions regulate different points in the chain

R/01	INDIA	Risk-based governance; proposed coordination, safety, incident and accountability mechanisms.
AI Governance Guidelines		
R/02	INDIA · FINANCE	Board policy, lifecycle controls, audits, consumer communication and third-party controls.
RBI FREE-AI Committee Report		
R/03	EUROPEAN UNION	Risk categories, provider and deployer duties, human oversight, logging and impact assessment.
AI Act		
R/04	CANADA	Published assessment, notice, explanation, peer review, human intervention and recourse.
Directive on Automated Decision-Making		
R/05	UNITED KINGDOM	Public records explaining purpose, ownership, suppliers and decision support.
Algorithmic Transparency Recording Standard		
R/06	UNITED STATES	Inventories, high-impact controls, acquisition testing, portability and data rights.
OMB M-25-21 and M-25-22		
R/07	INTERNATIONAL	Human-rights duties, impact assessment, procedural safeguards, complaint and remedy.
Council of Europe Framework Convention		



A/01 · INDUSTRIAL ENTERPRISE

The agent inside the operating core.

An AI agent coordinates procurement, maintenance and inventory across plants, suppliers and enterprise systems.

FAILURE FIELD

- Unsafe maintenance deferral
- Inventory distortion
- Unauthorised commitment
- Supplier manipulation
- Cascading outage

CONTROL FIELD

- Segregated permissions
- Financial and safety thresholds
- Action and override logs
- Manual operating continuity

THE GOVERNANCE TEST

Can the enterprise reconstruct, interrupt and continue the process when the agent or provider fails?



A/02 · PUBLIC ADMINISTRATION

The queue before the right.

An automated system triages applications for benefits or public services and directs cases toward approval, review or investigation.

FAILURE FIELD

- Invisible exclusion
- Proxy discrimination
- Automation bias
- Unintelligible denial
- No effective correction

CONTROL FIELD

- Published impact assessment
- Notice of automated assistance
- Meaningful explanation
- Human reversal authority
- Accessible appeal

THE GOVERNANCE TEST

Can the affected person discover the system's role and obtain timely human reconsideration?



A/03 · REGULATED AUTHORITY

The machine near the coercive edge.

AI assists investigation, supervision, enforcement triage or risk prioritisation within an authority able to impose legal or economic consequences.

FAILURE FIELD

- Biased attention
- Opaque evidence chain
- Automation escalation
- Vendor influence
- Inability to contest

CONTROL FIELD

- Lawful-purpose boundary
- Evidentiary provenance
- Independent validation
- Senior review
- External supervision

THE GOVERNANCE TEST

Can the authority defend the path from signal to state action before an independent forum?

Eight domains of institutional control

C/01 Purpose Name the decision or action the system may influence—and those it may not.	C/02 Ownership Assign a mission owner, technical owner, data owner and independent assurance function.	C/03 Permission Treat tool access, communications and transaction rights as institutional authority.	C/04 Evidence Preserve enough of the operating record to reconstruct consequential outputs and actions.
C/05 Intervention Specify who can pause, override, reverse and decommission the system.	C/06 Recourse Design notice, explanation, complaint and correction before the first affected decision.	C/07 Continuity Maintain portable records, fall-back procedures and rights to survive provider failure.	C/08 Supervision Create review beyond the delivery team as the context requires.

The right to transfer and continue a system is part of the right to govern it.

Procurement determines what the institution may understand, test, move, preserve and replace after deployment.

P/01

System boundary

Exact models, data sources, tools, environments, subprocessors and permitted changes.

P/02

Evaluation rights

Access needed to test performance, security, bias and context-specific limitations.

P/03

Operational evidence

Logs, version records, incident information, known limitations and change notification.

P/04

Data and training

Ownership, lawful use, retention, training restrictions, deletion and derivative data.

P/05

Portability

Export formats, documentation, interfaces, knowledge transfer and transition assistance.

P/06

Continuity

Service failure, provider exit, subcontractor change, replacement and manual fallback.

P/07

Liability

Responsibility for unauthorised action, security failure, infringement and remediation.

P/08

Termination

Power to suspend unsafe use without losing mission records or operating control.

I/01

DOCUMENTED

A national governance architecture is now explicit.

The 2025 guidelines describe a risk-based, light-touch and techno-legal approach, including proposed coordination, safety, incident and accountability mechanisms.

I/02

STRONG INFERENCE

Data protection is necessary but not sufficient.

The DPDP framework does not by itself answer who may delegate a decision, what evidence must be preserved or how an AI-assisted outcome may be appealed.

I/03

DOCUMENTED

Sectoral institutions may move first.

RBI translates general principles into board policy, product approval, lifecycle controls, audit, incident reporting and third-party governance.

I/04

OPEN QUESTION

Public-sector operating evidence remains thin.

The examined record does not establish a single mandatory public inventory, cross-government impact assessment or uniform recourse standard.

The institution
remains
sovereign only
if it can stop
the machine
and continue
the mission.

A kill switch without a viable fall-back process stops the system—and may also stop the institution.

-
- Q/01 Which Indian public institutions currently use AI to rank, triage, investigate, allocate or deny—and where is that use publicly recorded?
-
- Q/02 What threshold should convert an ordinary automation project into a high-impact institutional system?
-
- Q/03 Who is accountable when provider, integrator, model developer and deploying institution share the chain?
-
- Q/04 What minimum operating record is necessary for audit, judicial review and affected-person recourse?
-
- Q/05 Which classes of autonomous action require dual control, prior approval or categorical prohibition?
-
- Q/06 How should institutions test systems whose behaviour changes through models, retrieval, tools or agents?
-
- Q/07 What continuity rights should be non-negotiable in public or critical operations?
-
- Q/08 Which supervisor can evaluate both technical performance and legality of institutional purpose?
-

The rule, the inference and the unknown remain separate.

M/01 This edition examines laws, treaties, official policy memoranda, regulatory reports, administrative directives, public transparency standards and technical risk-management frameworks current to the stated evidence cut-off.

M/02 It does not provide legal advice, determine compliance in a named jurisdiction or evaluate a deployed system. The three institutional anatomies are illustrative contexts used to test how authority, evidence, intervention and recourse change across domains.

M/03 Binding law, administrative policy, voluntary standards and Foundation analysis are not treated as equivalent. The source register states institutional provenance; the finding register distinguishes documented mechanisms from comparative inference.

The rule should remain reachable

R/01	Office of the Principal Scientific Adviser, Government of India	India AI Governance Guidelines	2025
R/02	Office of the Principal Scientific Adviser, Government of India	Strengthening AI Governance Through a Techno-Legal Framework	2026
R/03	Reserve Bank of India	FREE-AI Committee Report	2025
R/04	Ministry of Electronics and Information Technology	Digital Personal Data Protection Rules	2025
R/05	NITI Aayog	Responsible AI: Operationalizing Principles	2021
R/06	National Institute of Standards and Technology	AI Risk Management Framework 1.0	2023
R/07	Organisation for Economic Co-operation and Development	OECD AI Principles	2024

The rule should remain reachable

R/08	European Union	Regulation (EU) 2024/1689 — Artificial Intelligence Act	2024–26
R/09	Treasury Board of Canada Secretariat	Directive on Automated Decision-Making	2025
R/10	Government of the United Kingdom	Algorithmic Transparency Recording Standard	2025
R/11	Council of Europe	Framework Convention on Artificial Intelligence	2024
R/12	UNESCO	Guidelines for the Use of AI Systems in Courts and Tribunals	2025
R/13	Executive Office of the President of the United States	OMB Memorandum M-25-21	2025
R/14	Executive Office of the President of the United States	OMB Memorandum M-25-22	2025



DOSSIER / 002 · VERSION 1.0

NATARAJA LABS RESEARCH FOUNDATION

The machine is governed where authority can still be named.

Nataraja Labs Research Foundation (2026), Who Governs the Machine?, Dossier 002,
version 1.0. First released 2 September 2026.

CORRECTIONS & SUBSTANTIVE EVIDENCE

contact@natarajalabs.org · Subject: Dossier 002 correction

REGISTERED OFFICE

A310, Phase 1, Chhatarpur Enclave, New Delhi - 110074