AI GOVERNANCE GUIDELINES

B. Ravindran

Wadhwani School of Data Science and AI (WSAI)
Centre for Responsible AI (CeRAI)

Indian Institute of Technology Madras

BACKGROUND

- **Sep 2023 – Advisory Group** constituted under **Principal Scientific Advisor**
- **Nov 2023 – PSA Subcommittee** tasked with drafting actionable recommendations for AI governance in India
- **Jan 2025 – Subcommittee report** underscored need for a coordinated, whole-of-government approach
- **Feb 2025 –** Over 2,500 submissions received from **public consultation**
- **July 2025 – Draft Committee** convened to develop revised AI governance guidelines based on feedback, existing literature, laws and international practice
- **Members:** Balaraman Ravindran, Abhishek Singh, Debjani Ghosh, Kalika Bali, Rahul Matthan, Amlan Mohanty, Sharad Sharma, Kavita Bhatia, Abhishek Aggarwal, Avinash Agarwal, Shreeppriya Gopalakrishnan

VISION AND OBJECTIVES

Balanced, inclusive AI ecosystem

Foster **equity** and **readiness** for future technological advancements

Unlock AI for inclusive growth

Drive **economic** and **social development** across strata

Mitigate societal risks from AI

Manage risks effectively, ensure **safety** and **ethical use**

Align AI with global standards

Ensure policy alignment with **international norms** and **best practices**

Create an enabling ecosystem

Build a regulatory environment that encourages **innovation** and **adaptability** for responsible AI innovation

KEY PRINCIPLES: THE 7 SUTRAS

Trust is the Foundation

Innovation depends on trust across the value chain

People First

Human-centric design and oversight

Innovation over Restraint

Encourage responsible experimentation

Fairness and Equity

Prevent bias and promote inclusion

Accountability

Assign clear responsibility across actors

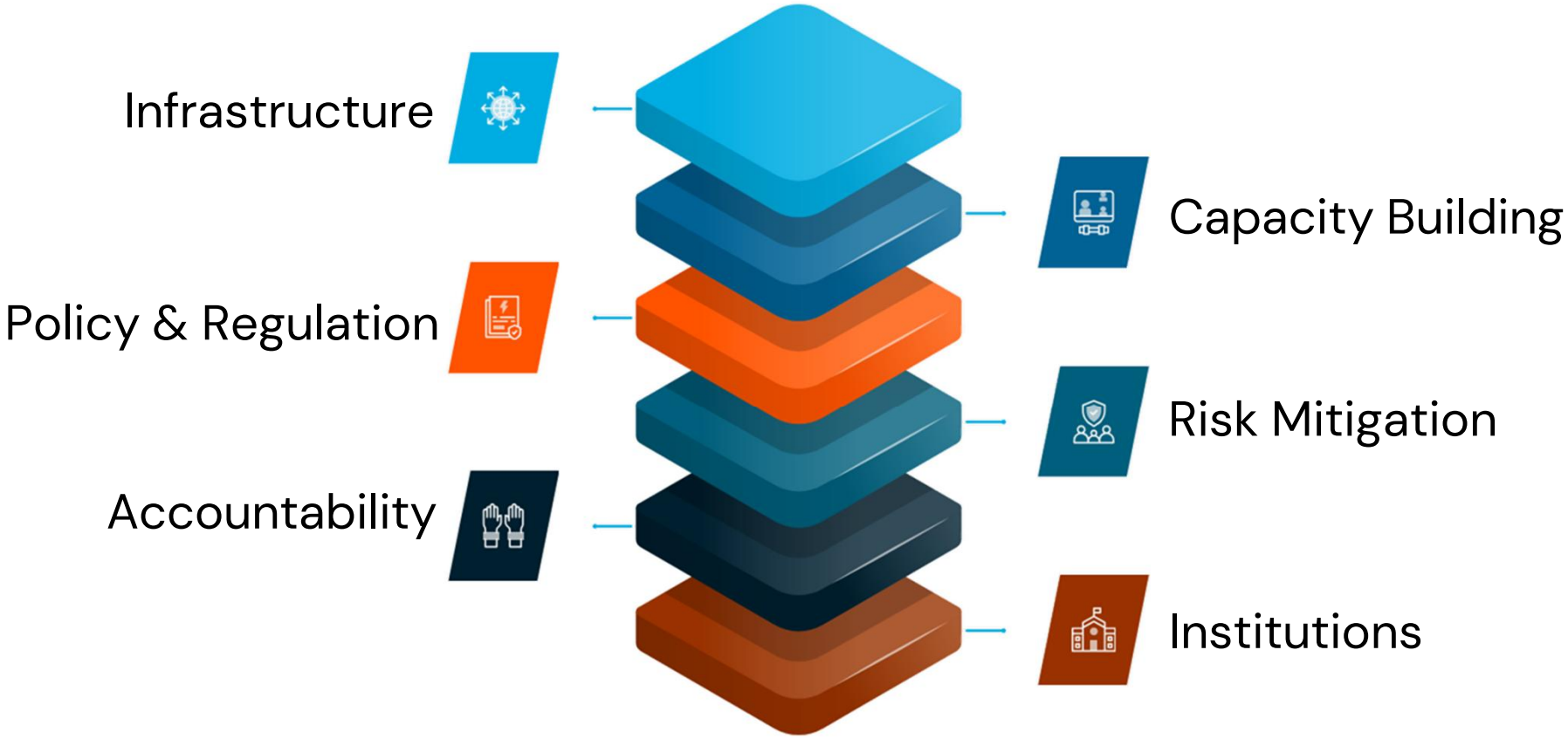
Understandable by Design

Ensure transparency and explainability

Safety, Resilience and Sustainability

Build robust, secure, and energy-efficient systems

GOVERNANCE ARCHITECTURE



INFRASTRUCTURE:
BUILDING THE FOUNDATION

- Expand access to compute and data resources for AI development
- Support indigenous foundation models to ensure security and relevance
- Integrate AI with Digital Public Infrastructure (DPI) for better governance
- Incentivize MSME adoption and innovation in AI technologies

RECOMMENDATIONS

1. Empower the India AI mission, line ministries, sectoral regulators and state governments to increase AI adoption through initiatives on infrastructure development and increasing access to data and computing resources.

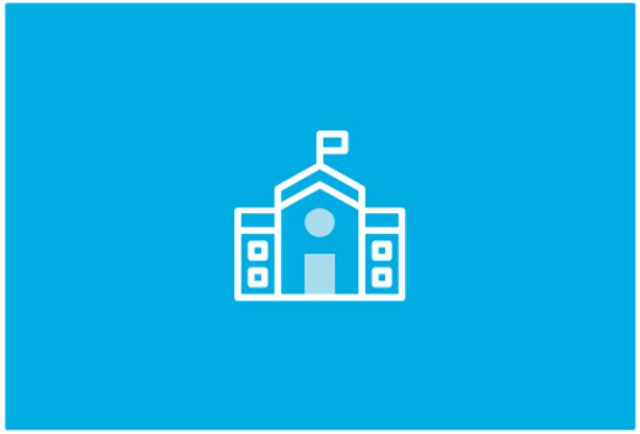
2. Increase data availability, sharing, and usability for AI development and adoption with robust data portability standards and data governance frameworks.

3. Encourage the use of locally relevant datasets to support the creation of culturally representative models and applications.

4. Promote access to reliable evaluation datasets and compute infrastructure for AI development and deployments, and to conduct safety testing and evaluations.

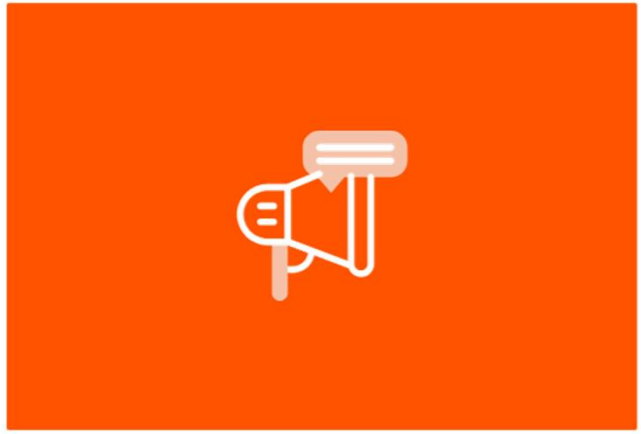
5. Integrate AI with Digital Public Infrastructure (DPI) to promote scalability, interoperability and inclusivity.

CAPACITY BUILDING:
EMPOWERING PEOPLE & INSTITUTIONS



Scale AI skilling and education programs

Train regulators and public officials



Promote public awareness and trust building

RECOMMENDATIONS

1. Increase societal trust and public awareness about the risks and capabilities of AI through regular training programs and publicity campaigns.

2. Conduct training programs for government officials, regulators and civil servants to understand AI technology developments, to manage public procurements effectively, and to encourage the responsible use of AI in the public sector.

3. Develop the capacity of law enforcement agencies (LEAs), police, cybercrime units, and prosecutors to detect, investigate and resolve AI-enabled crimes.

4. Expand capacity building initiatives to achieve deeper penetration of AI into tier-2 and tier-3 cities, and in vocational institutes.

POLICY AND REGULATION: AGILE, PRINCIPLE-BASED GOVERNANCE

Leverage existing laws on a technological continuum

Reinterpret and expand laws for copyright, data protection, liability

Pilot regulatory sandboxes for innovation

Promote technology-led approaches for compliance

RECOMMENDATIONS

1. Develop governance frameworks that are balanced, agile, flexible, and principle-based, and enable monitoring and recalibration based on feedback.

2. Review the current legal framework to evaluate risks and regulatory gaps.

3. Consider targeted legislative amendments to encourage innovation (for eg. in copyright and data protection) and to clarify issues around classification and liability.

4. Develop common standards and benchmarks to achieve regulatory objectives (e.g., on content authentication, data integrity, cybersecurity, fairness, etc.).

5. Create regulatory sandboxes to enable the development of cutting-edge technologies in constrained environments affording legal immunities.

6. Support strategic engagements and foreign diplomacy in national, regional and multilateral forums to further India's interests on AI governance issues.

7. Conduct horizon-scanning and scenario planning analysis to anticipate future developments in AI that may require policy or regulatory responses.

RISK MITIGATION:
MANAGE RISKS PROACTIVELY

Develop India-specific risk classification

Launch AI Incident Reporting System

Encourage voluntary safeguards & red-teaming

Embed human oversight in high-risk systems

RECOMMENDATIONS

1. Develop a risk assessment and classification framework that is customised for India's local context, and accounts for risks to vulnerable groups.

2. Establish a robust AI incidents mechanism to encourage individuals and organisations to report harm and create a feedback loop to track and analyse risks.

3. Encourage the adoption of voluntary frameworks to mitigate risks through principles, commitments, standards, audits, and appropriate incentives.

4. Guide the development and deployment of AI systems that are transparent, fair, open, non-discriminatory, explainable, and secure by design.

5. Use techno-legal measures and DEPA-based AI training models, where appropriate, to buttress existing policy choices, regulatory instruments, and voluntary measures.

6. Require human oversight and other safeguards to mitigate loss of control risks especially in sensitive sectors involving critical infrastructure.

ACCOUNTABILITY:
ENSURING PROPORTIONAL RESPONSIBILITY

- Implement graded liability across value chain
- Enforce existing laws visibly and consistently
- Promote transparency reports and grievance redressal
- Use technology-led tools for compliance-by-design

RECOMMENDATIONS

1. Clarify how different entities in the AI value chain (for example, developers, deployers, end-users) are governed under existing regulations, such as the IT Act.

2. Impose obligations for each of these entities that are proportionate to their function and the risk of harm (for example, transparency reporting, content removal, grievance redressal, transparency, and legal assistance).

3. Ensure laws are complied with through timely and consistent enforcement.

4. Mandate grievance redressal mechanisms with adequate feedback loops.

5. Provide guidance on how existing laws will be enforced in relation to AI systems (for eg. a master circular with a list of applicable regulations to support compliance).

6. Develop accountability mechanisms that would support voluntary compliance to mitigate harm (for example, self-certifications, peer monitoring, third party audits).

7. Increase transparency of the AI value chain so regulators have an understanding.

INSTITUTIONS:
ROLES AND RESPONSIBILITIES

AI Governance Group (AIGG)	Policy coordination and oversight
Technology & Policy Expert Committee (TPEC)	Strategic and technical guidance
AI Safety Institute (AISI)	Risk assessment, standards, global engagement
Sectoral Regulators	Domain-specific governance and enforcement

RECOMMENDATIONS

1. Establish an AI Governance Group to coordinate overall policy development and align AI governance frameworks with national priorities and strategic objectives.

2. Create a Technology & Policy Expert Committee which advises the AI Governance Group on matters of national and international importance relating to AI.

3. Provide adequate resources to the IndiaAI Safety Institute to conduct research, develop draft standards and their evaluation metrics and testing methods and benchmarks, collaborate with international bodies, national standard making bodies and provide technical guidance to regulators and industry.

IMPLEMENTATION ROADMAP

Short-term
Focus on establishing **governance structures**, developing **risk frameworks**, and expanding **AI awareness** and safety tools.

Medium-term
Strengthen regulatory frameworks, implement operational AI incident database, pilot **regulatory sandboxes** and enable **integration of DPI**.

Long-term
Continuously refine governance frameworks, build **global engagement**, and **draft new laws** to address emerging AI risks and opportunities.

THANK YOU