Aapti Institute for the Ministry for Electronics and Information Technology and IndiaAI

Comments: Report on AI Governance Guidelines Development

27th February 2025

Aapti Institute seeks to submit its suggestions for the AI governance guidelines development report released on 6 January 2025. This document's comments are a result of discussions between 14 February 2025 and 27 February 2025, by researchers at Aapti Institute, and the insights we have developed from our tech policy work on themes like responsible artificial intelligence, AI value chains, upholding human rights in the AI ecosystem, AI-related labor, and datasets. We hope that our inputs can help strengthen India's AI governance efforts and drive responsible innovation for the industry.

The following commentary and suggestions are a product of the contributions of the following people:

1. Ritvik Gupta
2. Priyam Vadaliya
3. Vaishnavi Patil
4. Gautam Misra
5. Rattanmeek Kaur
6. Supratik Mitra
7. Kunal Raj Barua

This document consists of broad recommendations that explore the report's discussions and recommendations, as well as more specific comments, highlighting potential changes in various sections and paragraphs of the report. **Our recommendations touch upon three thematic areas:**

1. The labor that provides the AI sector with data for purposes like training, and the challenges and potential ways forward for India's data workers.
2. The environmental costs and considerations of AI, and how sustainability can be enabled across the ecosystem.
3. The operationalisation of AI governance principles across lifecycles.

The development of India's AI governance guidelines is an encouraging step forward in the country's journey towards inclusion, innovation, and responsibility in AI. We are thankful for the opportunity to offer our comments on the report and hope they help strengthen the development of the guidelines further.

Contents

Thematic Commentary

This section makes thematic suggestions on policy areas and key considerations for AI governance that could be further explored to comprehensively address concerns pertaining to AI. The themes covered here add to the matters discussed in the report and its annexure.

Chapter 1 | AI labor: AI governance thinking needs to extend focus and action to the data workers responsible for data resources critical for AI development.

Excerpt from the Guidelines:

"Multiple actors can be involved across the lifecycle of any AI system. Together, they create an ecosystem. For example, in the context of the lifecycle of a foundation model, multiple sets of actors can be involved, including:

- *Data Principals*
- *Data Providers*
- *AI Developers (including Model Builders)*
- *AI Deployers (including App Builders and Distributors)*
- *End-users (including both businesses and citizens)*

Traditional governance approaches can be limited if they focus on one particular set of actors in isolation. By looking at governance across the ecosystem, better and holistic outcomes are obtained. An ecosystem view of actors could also look to clarify the distribution of responsibilities and liabilities between different actors involved in the ecosystem."

-**II. Governance of AI, B. 2., page 5,** Report on AI Governance Guidelines Development

A growing population of data workers help build datasets for various kinds of AI technologies and services, helping AI developers train models and refine services. **Data workers are key to the growth and well-being of the country's AI value chains and are thus key ecosystem actors.**

Regarding the labor that feeds AI development, India is at a cross-roads, and needs to cross two major milestones: **identifying a definition that recognizes "data work" within the country as a profession, and bringing it into the fold of labor regulations and policy action.** With respect to regulation, the state can look into extending existing laws and frameworks on labor towards data work, as well as introducing new measures for the more unique aspects of data work and AI value chains.

1.1. Building the context for the recommendations

Preparing for a growing labour force: While the report discusses the need for a lifecycle approach to AI systems and an ecosystem view of AI actors, it should aim to provide further

guidance on India's approach to labor, data work and the companies involved in providing datasets and other similar services for AI actors. NASSCOM projected that India's markets for services like data annotation could grow considerably, speculating that India could grow from 70,000 data annotation workers to 1 million workers by 2030. However, such work is often fraught with risks and challenges that make working conditions difficult for people contributing to the services enabling AI development. Labor issues related to AI are seeing greater consideration in recent initiatives and discussions, such as the "Pledge for a Trustworthy AI in the World of Work."

Identifying the potential risk for data workers: Data workers can face many risks, such as a volatile supply of work and low levels of remuneration, and a significant amount of Indian workers are engaged in such work arrangements. To provide a basic level of safety to the growing professions of data work, Indian thinking on AI governance needs to discuss and develop an approach for regulating companies and workplaces that produce AI-related resources like datasets.

As a critical part of AI development, data work requires consideration in terms of both how this industry can be nurtured, and how its workers can be provided a measure of protection from uncertainty and precarious working conditions. Organizations like the Partnership on Artificial Intelligence (PAI), the Distributed Artificial Intelligence Research (DAIR) Institute, and Fairwork have provided both insights on the precarity of data work, as well as guidance on responsible practices and improving working conditions.

Understanding the context of varying models: Research suggests that AI data work firms can exist across a variety of models: *(i) internal services and platforms serving companies, (ii) generalist and AI data business process outsourcing (BPO) firms, and (iii) generalist and AI data platforms*. A variety of practices and processes exist across different models that make data work a precarious form of work. Policy thinking on data work needs to remain cognizant of the different business models that populate the data work landscape.

Mapping the data work ecosystem in India: While the body of work from academics, labor stakeholders, and reporters is growing, India could benefit from a national-level analysis of the size and conditions of data workers. Policy direction will require actionable information that provides numbers on the number of data workers active in India and the rates at which they prepare datasets. For example, in the past, the NITI Aayog released a report that provided an overview and estimates on India's gig and platform workforce. Visibility around data work could be considered minimal, making it vital that some form of large-scale survey and examination shed light on the role's working conditions and trends in India. Such initiatives will help think about regulating the governance of AI-related labor for both the state and the industry.

1.2. Recommended pathways

- **Resources should be committed to exploring India's data work ecosystem, exploring aspects like the size of the market and worker base, the prevalent business models, and the trends in designing working arrangements.** Such an

examination can help the state and industry decide on suitable policy directions to discuss and pursue.

- **Regulatory provisions should apply to business models and work arrangements in ways that bring a tangible level of safety and certainty to various forms of data work available in India.** Thus, new explorations would help address data work's more novel risks and considerations.
- **The inclusive governance of AI will require extending current policies to data work alongside the introduction of new, context-sensitive measures.** Some topics that could require more specialized provisions and new thinking would be the use of technological surveillance and algorithmic management systems on workers, the calculation of task-based remuneration for different kinds of data enrichment, and the streamlining of waiting periods between new paid work being made available to workers.

1.3. Specific policy recommendations for data work: What existing measures can cover, and what requires new approaches

State intervention on data work would involve two forms of action: **utilising existing laws and policies and introducing new measures**. It is essential to note that there are elements of data work arrangements that can create difficulties in applying existing labor laws.

The table below (Table 1) lists some policy areas that could potentially benefit from expanding existing provisions versus those where novel interventions might be required. A comprehensive policy approach will require all stakeholders to balance existing measures with the need for interventions that engage with the aspects of data work that make it different compared to older, more "conventional" forms of work.

Areas that could build from existing provisions	**Working hours:** Indian labor laws have limits for people's working hours. **Social Security:** The Code on Social Security, 2022, and the recent state-level bills on platform and gig work may provide a potential way to build services for data workers. **Minimum wages:** India has provisions on minimum wages for workers that could be extended to data workers to provide a level of pay that can cover survival and basic necessities.
Areas that will benefit from new measures and thinking	**Algorithmic management:** Technological measures can be used to manage and monitor workers, which may lead to issues like stress due to demanding metrics, and automated termination from work without appeals or a robust explanation. Data workers would benefit from provisions and protections that ensure workplace safety and fairness with respect to the usage of data and algorithms to manage people. **Psychological harm:** Certain forms of data work, like content moderation, can traumatise workers through the materials

	they routinely handle. They require access to mental health services to cope with the job's psychological toll. **Social security:** While there exist frameworks and laws that deal with the questions of social security, newer approaches could help design more context-sensitive solutions. **Minimum wages:** Data work is a profession where some business models have workers enduring substantial periods without paid work, as they await platforms to assign them projects and tasks. In such situations, where pay is linked to the access to tasks, the minimum wage function could be strengthened. **Pay calculation:** Data workers often complete tasks that require a considerable amount of time, effort, and skill, but provide low pay in return. There need to be tools or guidelines that help calculate remuneration in a way that is cognizant of the level of labour that different kinds of data tasks require. **Operating costs:** In situations where companies do not provide necessary assets like computers, internet connectivity, or specific assistance in accessing resources, significant portions of workers' earnings are eroded by the cost of procuring and maintaining these assets. Minimum wage policy also needs to be considerate of the diminishing effects asset-light businesses' policies have on workers' earnings.

Table 1. Policy recommendations for data work

Chapter 2 | Natural resources, environmental risks, and AI: Building a framework for managing the environmental costs and challenges posed by the AI sector

Excerpt from the Guidelines:

"• facilitate the development of metrics (e.g., measurement standards for assessing environmental impact of AI in India) & common frameworks (e.g., on data provenance, system cards, security, evaluation data sets, use of open source, transparency reports etc.); and..."

-IV. Recommendations, 2., page 15, Report on AI Governance Guidelines Development

While AI provides tools to address the pressing issues of the climate crisis, it also requires significant amounts of computational power to develop, deploy and diffuse models. This demand will be felt acutely in a developing economy like India. As AI becomes more integrated into varied industries, we need to ensure that its environmental impact does not offset the benefits it offers. There is increasing pressure for industries to adopt eco-friendly practices that address the climate crisis, develop buy-in from consumers, create pathways towards material efficiencies, and ensure the industry's long-term viability. Thus, mitigating AI's environmental risks becomes a primary concern for India's AI governance — which could result in the reduction of both ecological and operational costs of AI systems. **It is recommended that India's conceptualisation of AI governance is cognizant of the material flows of the AI lifecycle, and is capable of encouraging efficient and sustainable management of computational infrastructure.**

The lifecycle approach purported by the report provides a solid foundation to empower stakeholders to instil sustainability across the AI value chain. Importantly, it will encourage governance of AI's energy and water consumption in a comprehensive manner. **By adopting a wide lens when addressing sustainability, AI governance in India could potentially address the fragmentation between sectoral efforts to bring in effective environmental regulations.**

2.1. Building the context and framing the recommendations

A Goldman Sachs report states that each ChatGPT query uses 10 times as much electricity as a Google search. Research also suggests that a GPT-3 query that generates 10-50 medium-length responses consumes a 500ml bottle of freshwater. The energy and water demand is growing exponentially in line with the growing market of new AI products.

2.1.1. Understanding the impact of data centers and adopting 'Decarbonisation' approaches

AI represents the fastest-growing workload on data centers, driving 160% increase in data center power demand. Additionally, according to the International Energy Agency, data centers

themselves are estimated to consume twice as much electricity by 2026 as today, equivalent to the country of Japan.

Decarbonization entails reducing energy consumption, investing in renewable energy, and reducing CO2 emissions. Metrics like Power Usage Efficiency (PUE) provide insight into approaches of decarbonizing the data center by inviting hardware and cooling systems that are energy-efficient.

Examples from the ecosystem

The Telecom Regulatory Authority of India (TRAI) recommends that all India-based data centers must rely on 95% renewable energy sources. Additionally, it has pushed for national standards and codes in line with the NBC (National Building Code), ECBC (Energy Conservation Building Code), and guidelines issued by the Central Pollution Control Board (CPCB) and the Ministry of Environment, Forest and Climate Change (MoEFCC).

States like Karnataka have committed, through their Data Center Policy, to encouraging the use of renewable energy. The policy proceedings claim that some players have committed towards renewable energy usage to fulfil up to 80% of the data center's power requirements.

2.1.2. Recommended pathway

These recommendations should be brought under the purview of AI governance, with its whole-of-government approach. IGBC's existing green data center certifications — paired with national policies that help decentralise data centers, provide for recycling of materials, minimize data redundancies, and encourage aligning data center practices with suitable climatic conditions — can help the sustainable development of Indian AI systems. **While there is a need to decarbonize the entire AI lifecycle, recognising and focusing policy intervention at the data center is a needed starting point.**

2.2. Reducing the water footprint

The increasing water footprint of AI has received less attention as compared to its energy consumption. The urgency to ensure that AI does not exacerbate the global water stresses is felt even by companies like Google and Microsoft, who have pledged to mitigate AI's water consumption. Such commitments can be further supported through national policies that encourage sustainable practices in the AI lifecycle.

A McKinsey report suggests that a modern data center can devote around 40% of its energy usage to cooling systems. Air cooling can be significantly inefficient as compared to liquid cooling, especially in climatic conditions prevalent in India, thereby placing a burden on the data center's power consumption. Thus, water becomes the go-to solution for efficient cooling in data centers. However, this can put stress on fresh water sources in the region.

2.2.1. Recommended pathway

Alternative cooling solutions like immersion cooling should be encouraged through policy. In case such alternatives are infeasible, greywater should be funneled into the closed-loop cooling systems to minimize the data center's water footprint. Additionally, to ensure sustainable and long-term AI innovation, **transparent monitoring and reporting of a data center's water usage should be made mandatory**.

Chapter 3 | Operationalising AI governance principles

While the report incorporates AI governance principles, effectively operationalizing them requires greater nuance. Without disaggregated mechanisms for implementation, these principles risk having a limited or unfulfilled impact.

Therefore, the report could benefit by considering how these principles are applied across the AI lifecycle. Below are some key suggestions to strengthen this approach.

3.1. Contextualising AI bias - *Fairness in AI*

> **Excerpt from the Guidelines:**
>
> *It is also important to understand that in many decisions there may be a notion of bias, however, only biases that are legally or socially prohibited need to be protected against.*
>
> - **III. Gap analysis, page 11, 4. AI led bias and discrimination**

While the report examines the modalities of AI bias, it would benefit if contextualised to India's socio-political realities. Popular scholarship on AI fairness principles has focused predominantly on two axes of discrimination, race and gender. Its interoperability with different socio-political contexts has already been questioned. Building a definitive understanding of how AI bias can impact Indian users, or how AI systems can produce biased outcomes, specific to the Indian context "socially prohibited biases" could strengthen the effective scope or definition.

Thus such an effort to understand fairness in India could greatly benefit the explorations of the **technical secretariat** and the **committee/group.**

Recommended pathways:

- To develop a more inclusive understanding of AI fairness in India, it is essential to **facilitate multi-stakeholder consultations** that bring together diverse voices, including civil society organizations, academia, and end-users. These spaces should be designed to foster meaningful engagement and dialogue, ensuring that the perspectives of marginalized and underrepresented communities are actively included. **Capturing the lived experiences of these groups will help uncover how AI bias operates in different contexts and inform more equitable AI governance frameworks.**

- **Dedicated funding mechanisms and institutional support must be extended** to advance research on the socio-political implications of AI systems in India. Encouraging interdisciplinary studies that examine AI bias through the lens of caste, language, and regional disparities will provide a more nuanced understanding of its impact. **Strengthening such research efforts will be crucial in shaping policies and interventions that address context-specific challenges and promote fairness in AI systems.**

3.2. Transparency of processes and mechanisms of making an AI incident database- strengthening transparency and accountability

> **Excerpt from the Guidelines:**
>
> *To build evidence on actual risks and to inform harm mitigation, the Technical Secretariat should establish, house, and operate an AI incident database as a repository of problems experienced in the real world that should guide responses to mitigate or avoid repeated bad outcomes.*
>
> - **IV. Recommendations 3, page 11,**

An AI incident database, when designed with a risk-based approach, can play a crucial role in mitigating risks in AI systems. Documenting real-time harms not only helps push for stronger standards but also deepens the understanding of how AI-related failures could unfold. However, without clear, transparent processes and mechanisms, such a database may fall short of its intended impact. Supported by well-defined criteria for AI harms that go beyond legal frameworks, databases should capture broader social, ethical, and systemic consequences that could strengthen its effectiveness. Open access to these databases can further drive accountability, enabling both users and key stakeholders in the AI ecosystem to take more informed and responsible actions.

Recommended pathways

- To ensure the effectiveness of an AI incident database, **clear reporting standards must be established.** This includes defining transparent rules for submitting incidents, specifying the required documentation, and setting clear criteria for what qualifies as reportable AI harm. Additionally, a **well-defined process for verifying the provenance** of incidents is essential to maintain credibility and accuracy. To ensure fairness and consistency, **explicit guidelines should also be outlined for accepting or rejecting reported incidents**, preventing arbitrary decision-making and enhancing trust in the system.
- To balance transparency with privacy, **a well-defined access framework** must be implemented for the AI incident database. This may include establishing tiered access mechanisms that allow different stakeholders to engage with the database in a responsible and meaningful way. Additionally, **privacy protocols** should be established to anonymize sensitive data while preserving the integrity of reported incidents

3.3. Building Knowledge Pools and Talent for AI Cybersecurity - *Safety and security*

As AI becomes increasingly embedded in cybersecurity ecosystems, new risks to critical infrastructure and sensitive data emerge. While AI enhances threat detection, automates responses, and improves predictive analytics, it also introduces vulnerabilities that adversaries can exploit. Malicious actors are leveraging AI for more sophisticated cyberattacks, including deepfake-based social engineering, adversarial machine learning, and automated hacking tools designed to bypass traditional security measures. Addressing these challenges requires a holistic approach that goes beyond legal protections for end-users to encompass broader national security concerns.

Recommended pathways

- To strengthen defences against AI-driven cyber threats, **India must actively foster co-learning, co-sharing, and co-creation initiatives with international allies**. Bilateral and multilateral cybersecurity collaborations—such as the India-Australia Cyber Treaty—can play a pivotal role in enhancing collective resilience.
- Additionally, a significant talent gap of cybersecurity professionals with AI expertise could be bridged, making cybersecurity skilling programs a critical priority. **Investing in AI-focused cybersecurity training and establishing international partnerships** for workforce development will create economic opportunities and bolster India's cyber resilience.

Specific Comments

The table below covers specific suggestions for the content and discussions featured in the report, along with the rationale for suggested inputs. These suggestions stem from the themes discussed in the preceding section of this document.

#	Report Content	Suggested Content	Rationale
1	**Development** which involves examining the designing, training, and testing of a given system. *(Page 5 of 22, II., B. 1. Examining AI systems using a lifecycle approach)*	**Development** which involves examining the **data labelling and curation**, designing, training, and testing of a given system.	Data annotation and labelling/content moderation is a critical activity in the development stage of AI systems to produce quality data sets on which models are trained on. Adding the term '*data labelling and curation*' will make the lifecycle approach more inclusive and influence specific considerations while strengthening the AI principles and governance framework.
2	For example, in the context	For example, in the context	Data workers play a key role

#	Report Content	Suggested Content	Rationale
	of the lifecycle of a foundation model, multiple sets of actors can be involved, including: • ***Data Principals*** • ***Data Providers*** • ***AI Developers*** (including Model Builders) • ***AI Deployers*** (including App Builders and Distributors) • ***End-users*** (including both businesses and citizens) *(Page 5 of 22, II., B. 2. Taking an ecosystem view of AI actors)*	of the lifecycle of a foundation model, multiple sets of actors can be involved, including: • ***Data Principals*** • ***Data Providers*** • ***Data Workers*** • ***AI Developers*** (including Model Builders) • ***AI Deployers*** (including App Builders and Distributors) • ***End-users*** (including both businesses and citizens)	in creating high quality data sets which are used for training the AI models. India is a growing market for providing data annotation services which is speculated to grow to 1 million workers by 2030 according to the Nasscom report, where data workers continue to face a range of risks. Including '*Data Workers*' in the list of key actors will ensure addressing the specific risks the workers face and creating safer working environments while developing AI governance policies and guidelines.
3	Consequently, the Technical Secretariat could start such work by anchoring a collaboration with industry to build consensus around voluntary commitments. Such commitments can include elements such as: • disclosures of the **intended purposes** of AI systems and applications; • commitments to release regular **transparency reports** by AI developers and deployers; • commitments to internal and external red-teaming of models or systems in areas; • processes to test and monitor data quality, model robustness, and outcomes; • processes to validate data quality and governance measures; • processes to ensure peer review by third-party qualified experts;	**Add the following commitments to the list:** • Incorporating fair labor and decent work practices through compliance with regulations, as well as adopting labor standards and principles. Some examples of standards that can help mitigate some of data work's precarity include the Fairwork AI Principles and the PAI Data Enrichment Sourcing Guidelines.	The workers that meet the AI ecosystem's data needs face considerable challenges and uncertainty. Meeting certain minimums regarding the treatment of labor, businesses can bring in measures of safety, stability, and fairness to data work. Incorporating this stance could incentivise businesses to strengthen their internal policies and practices.

#	Report Content	Suggested Content	Rationale
	• processes to ensure conformity assessments with accepted responsible AI principles; • security, vulnerability assessment, and business continuity requirements. *(Page 17 of 22, IV., 4.)*		
4	Non-official members could include persons capable of representing the interests of AI developers, AI deployers, data providers, data principals, and end-users – so that the perspectives of the overall ecosystem can be considered. *(Page 14 of 22, IV., 1.)*	Non-official members could include persons capable of representing the interests of AI developers, AI deployers, data providers, data principals, **academia, CSOs** and end-users – so that the perspectives of the overall ecosystem can be considered	As an emerging technology, AI is constantly evolving, rapidly evolving and creating new use cases. Additionally, changes to the technical changes within the architectures and functionalities of models also continue and require dedicated attention. Researchers, CSOs, and academia should be formally recognised as key actors in the AI governance ecosystem. Their contributions can strengthen research and create a robust understanding within the committee.