

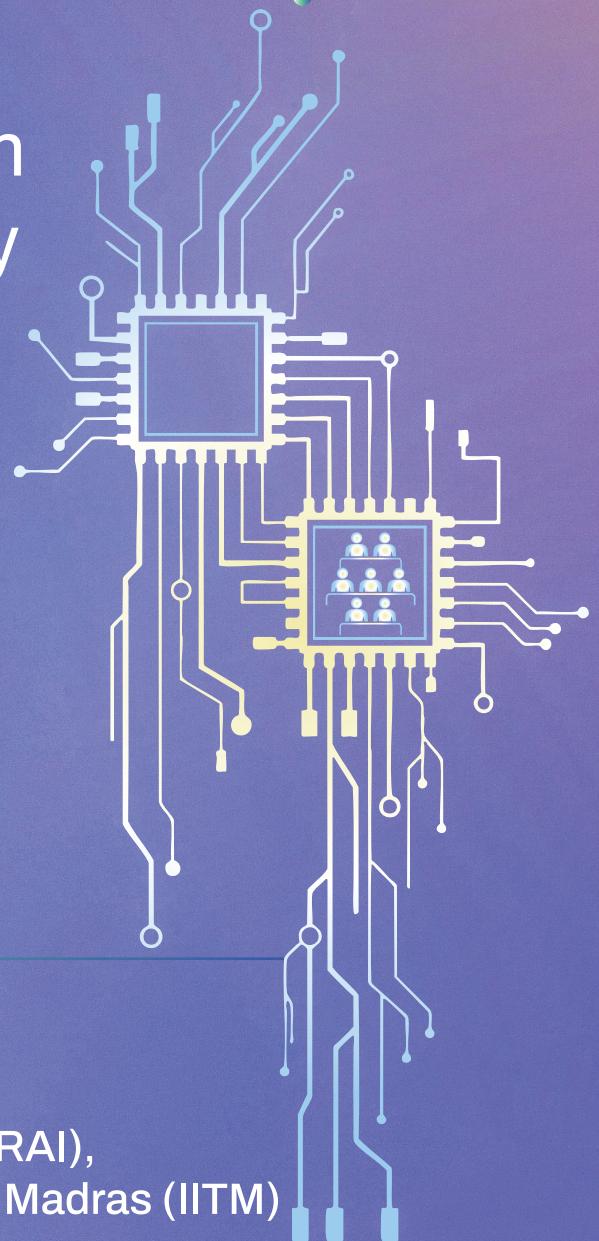
MAKING AI SELF-REGULATION WORK

Perspectives from
India on Voluntary
AI Risk Mitigation

AMLAN MOHANTY

APRIL 2025

Centre for Responsible AI (CeRAI),
Indian Institute of Technology, Madras (IITM)



EXECUTIVE SUMMARY

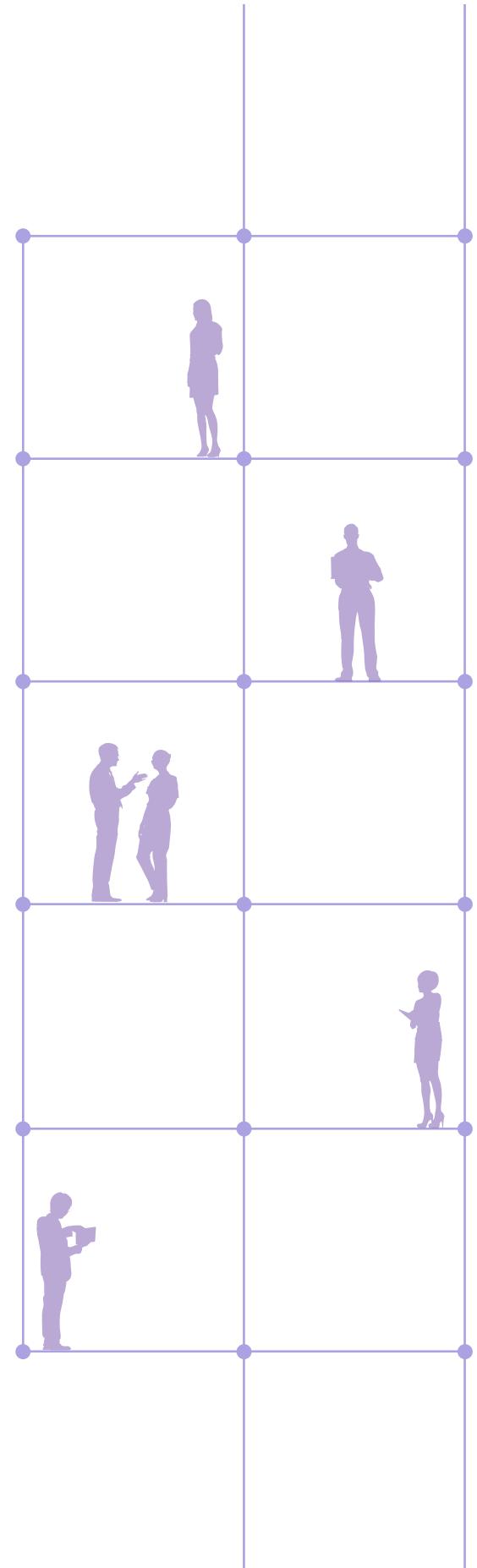
CONTEXT

Self-regulation is increasingly being positioned as a key pillar of artificial intelligence ('AI') governance around the world. Major economies, such as the United States and United Kingdom, are generally opposed to introducing any regulations that could stifle innovation. Others, such as Japan and the European Union, are beginning to make some concessions on binding rules as well. As a result, voluntary efforts by organisations to mitigate the risks of AI have assumed greater significance.

On its part, the Indian government has indicated strong support for AI self-regulation as it aligns with its strategic objectives of fostering AI-driven economic growth while addressing the risks. An expert committee set up by the Principal Scientific Advisor and convened by the Ministry of Electronics and Information Technology ('MeitY') has specifically recommended the adoption of 'voluntary commitments' to promote trust and transparency in the AI ecosystem.

However, skepticism persists, particularly from civil society organisations who question the effectiveness of self-regulation. There is also a lack of clarity on what constitutes self-regulation, what its objectives are, and how it can be made effective in the Indian context.

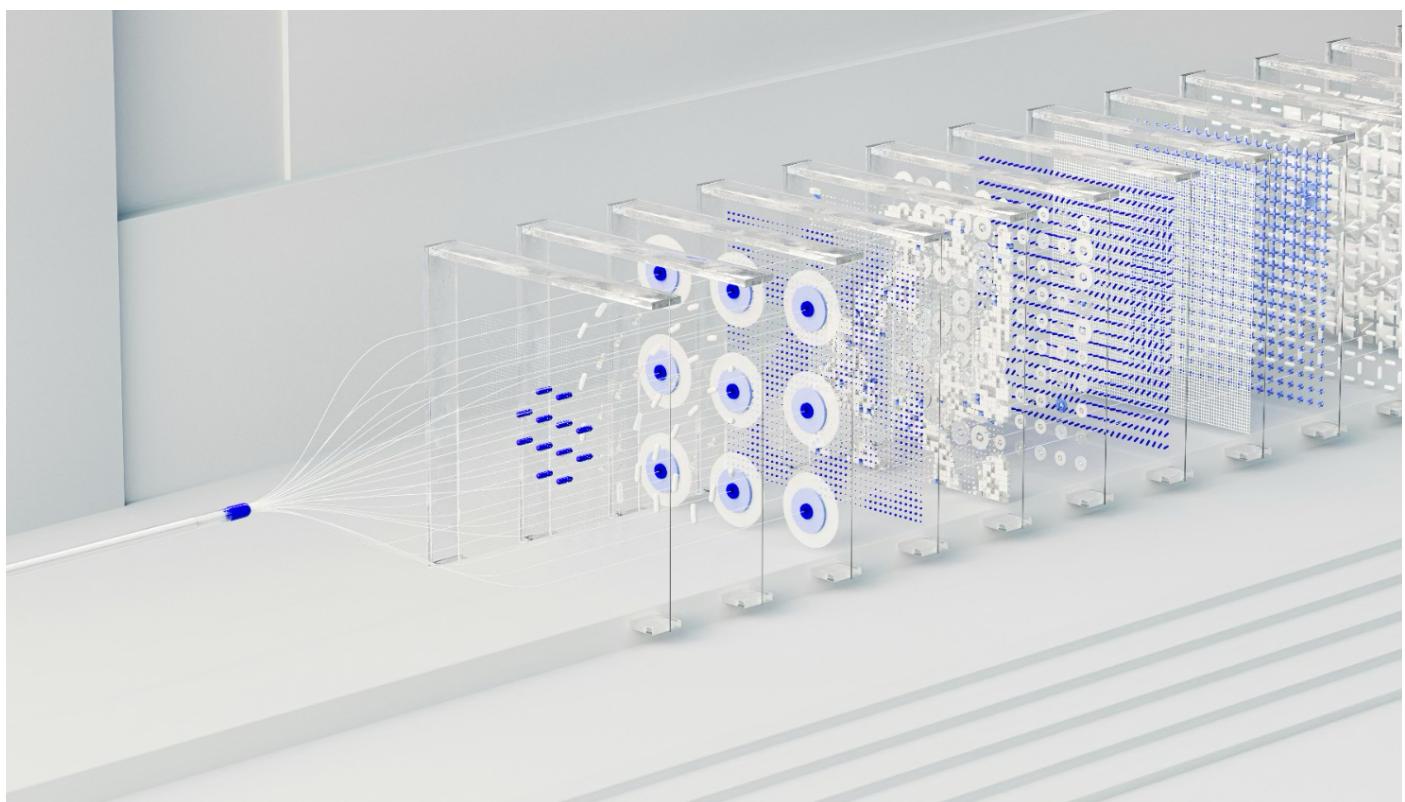
This paper presents a structured approach to explore these questions around AI self-regulation with reference to conceptual frameworks, market factors and the views of key stakeholders in India.



DEFINING AI SELF-REGULATION

AI self-regulation refers to the adoption of principles, standards, and frameworks by industry actors to mitigate AI-related risks on a voluntary basis, without the threat of legal enforcement. From a conceptual lens, the following preconditions must be met for it to be considered ‘self-regulation’:

- **Voluntary participation:** Firms must be able to opt in freely without coercion.
- **Non-enforceability:** These frameworks should not create binding legal obligations.
- **Limited penalties:** Non-compliance should not result in legal action but may carry reputational or industry-level consequences.
- **Flexibility:** Guidelines should be flexible and tailored to specific issues, actors and sectors, while allowing for iterative changes to ensure that they are relevant in view of evolving technological developments and regulatory objectives.
- **Multi-stakeholder input:** Participation from government, industry, and civil society is crucial to ensuring credibility and widespread adoption of these frameworks.



OBJECTIVES OF AI SELF-REGULATION

AI self-regulation serves as a proactive mechanism to balance innovation with risk-mitigation. Specifically, the objectives of such a framework are:

- **Enhancing trust and safety** in the AI ecosystem.
- **Developing industry norms** through transparency and accountability measures.
- **Ensuring regulatory predictability** for businesses to drive innovation.
- **Providing evidence of market behaviour**, regulatory gaps and risks to policymakers that can help inform future regulatory interventions.
- **Reducing the administrative burden** on government agencies by shifting compliance to industry-led initiatives.

STAKEHOLDER PERSPECTIVES

Key stakeholders across government, industry, and civil society generally agree on the need for self-regulation, but several contentious issues remain:

- **Risk classification:** Some favour self-regulation for both low-risk and high-risk AI applications, while others insist on binding rules for high-risk applications. There is also ongoing debate on whether risk classifications should apply to AI models or applications, or both.
- **Localising AI principles:** Policymakers want organisations to adapt global AI principles to India's socio-cultural context, while industry stakeholders say it is operationally challenging.
- **Targets of regulation:** There is a lack of consensus on whom the voluntary commitments should apply to, though there is a preference for a phased approach starting with a few anchor firms.
- **Inputs and Outputs:** Most stakeholders agree that harmful AI-generated outputs should be covered by self-regulation, but it is unclear how inputs relating to data and copyright should be governed. Further, the responsibilities of different platforms will need to be determined, especially social media companies on whose platforms the outputs of generative AI systems are widely distributed.

POLICY RECOMMENDATIONS

Given these divergent views and challenges relating to risk classification, consensus building, voluntary participation and accountability, the paper makes the following policy recommendations:

1. Develop a Risk-Based Classification

A structured framework is necessary to differentiate AI use cases based on various risk factors and to apply self-regulation to appropriate use cases. These factors include the potential for malicious use, algorithmic discrimination, systemic risks, loss of control, and threats to national security and public safety. The risk classification framework should be informed by empirical data, case studies, and incidents of harm specific to the Indian context. Additionally, it is crucial to incorporate input from multiple stakeholders to ensure a comprehensive and balanced approach to risk assessment.

2. Ensure Government Involvement

For AI self-regulation to be effective, the government must take an active role in encouraging and incentivising industry to participate. By initiating, facilitating, and endorsing voluntary frameworks, the government can enhance the credibility and legitimacy of these regulatory measures. Active government participation will also help align self-regulation with national priorities and regulatory objectives.

3. Introduce Market Incentives

Voluntary compliance with AI commitments can be strengthened by introducing a variety of financial, regulatory, technical, and reputational incentives. For example, adopting AI self-regulatory frameworks could be made a prerequisite for public procurement contracts and funding opportunities under the IndiaAI mission. Additionally, regulatory benefits such as access to sandboxes can serve as incentives for organisations to adopt self-regulatory measures. At the same time, ensuring that these voluntary frameworks are practical and technically easy to implement is particularly important for smaller firms with limited resources.

4. Adopt Accountability Measures

Given the lack of legal enforceability of voluntary codes, alternative accountability measures will need to be implemented. Organisations should be encouraged to adopt a variety of compliance measures, such as publishing transparency reports, updating their platform policies, adopting self-certifications and international standards, monitoring actions of industry peers, and conducting voluntary audits where feasible.

5. Provide Institutional Support

Institutional support is essential for sustaining AI self-regulation efforts. The proposed AI Safety Institute for India can play a pivotal role in guiding industry initiatives, developing benchmarks, and promoting the widespread adoption of AI safety tools. Additionally, a Technical Advisory Council may be established to provide expertise to government agencies, facilitate risk assessments, and support compliance efforts. In the long-term, exploring the feasibility of industry-led Self-Regulatory Organisations ('SROs') can help ensure that actors in specific high-risk sectors adhere to voluntary AI codes, creating a structured and nuanced approach to self-regulation.



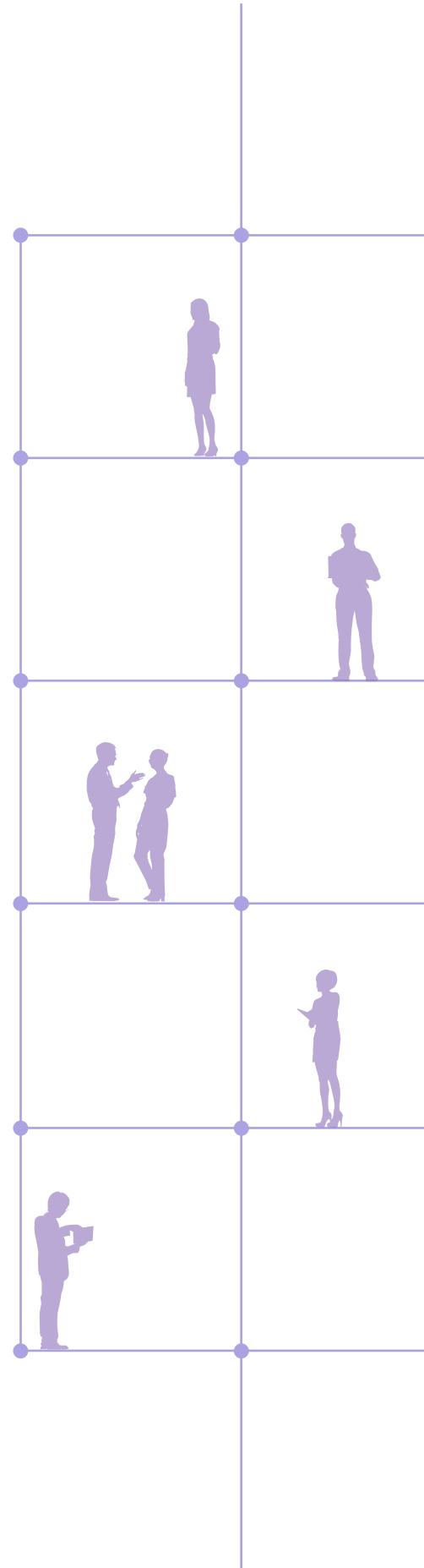
INTRODUCTION

- Self-regulation is a popular approach to govern artificial intelligence ('AI'). It is embedded in AI policy frameworks around the world, including the United States², Singapore³, Japan⁴ and the United Kingdom⁵. At the same time, there is deep skepticism about its effectiveness⁶. For that reason, self regulation is considered a polarising, controversial, and often misunderstood approach to AI governance.

On its part, the Indian government has endorsed self-regulation to support its 'pro-innovation' approach to AI governance⁷. A report published by the Ministry of Electronics and Information Technology ('MeitY') in January 2025 suggests that "meaningful initiatives by the industry to demonstrate self-governance can significantly enhance trust in the use of AI". The report specifically recommends the adoption of 'voluntary commitments'⁸. In Parliamentary discussions, the Minister of State for Electronics and IT has stated that "the government's aim is to create a supportive environment that encourages organizations to follow good practices voluntarily."⁹ AI self-regulation is clearly a priority for Indian policymakers.

However, despite a rich history of AI self-regulation – from national principles¹⁰ and government advisories¹¹, to responsible AI practices¹² and industry handbooks¹³ – there is a dearth of authoritative literature on what AI self regulation is, why it is useful, and how it can be effectively implemented in India.

This paper seeks to fill that gap. It suggests that self-regulation is an important step in India's path to developing a comprehensive AI governance framework. It explains through a conceptual lens what AI self-regulation is, evaluates the sentiments of key stakeholders, and recommends a structured policy framework for effective implementation of AI self-regulation in India.



AIMS AND OBJECTIVES

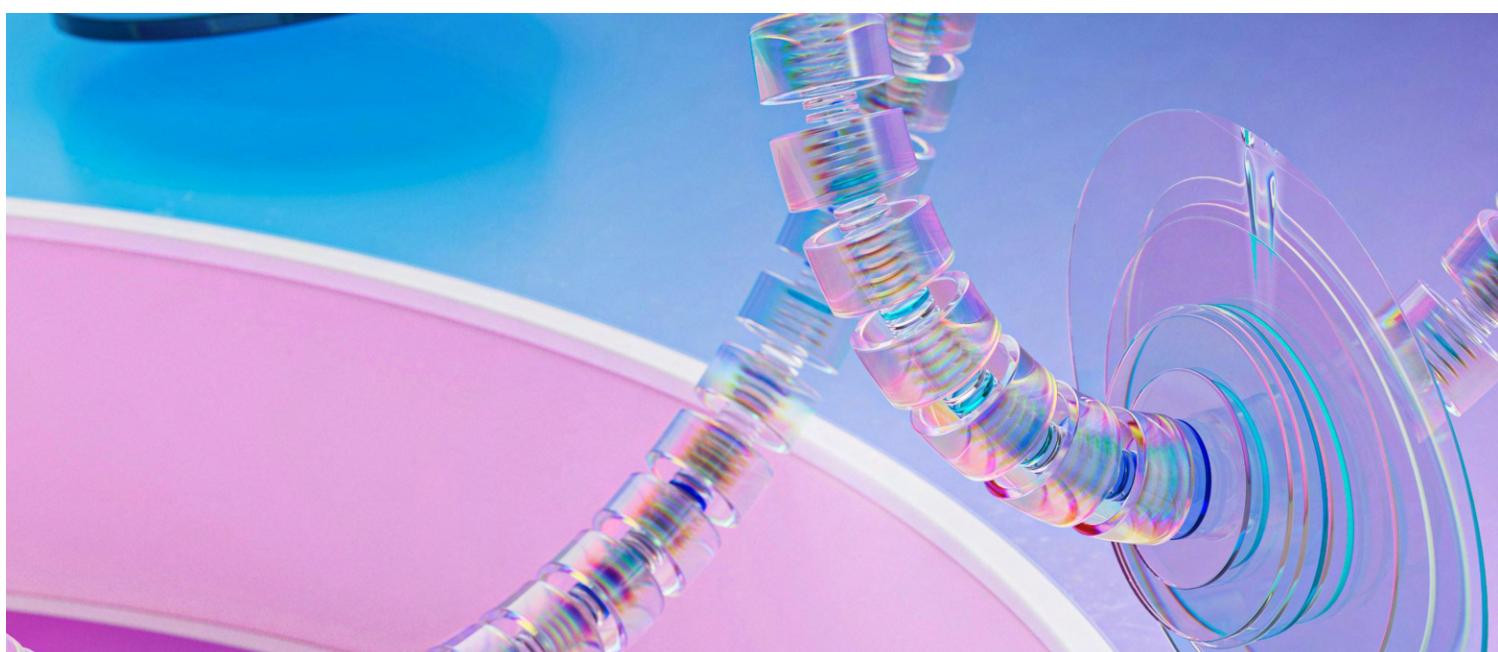
The aim of this paper is to study the features, objectives, and institutional mechanisms of self-regulation as they are applied to AI governance, grounded in the Indian context. The specific aims are to:

1. Provide a clear conceptual understanding of AI self-regulation
2. Explain how AI self-regulation can help achieve India's stated policy objectives
3. Recommend a framework for effective implementation of AI self-regulation in India

The objective of this paper is to inform the development and implementation of self-regulatory AI governance frameworks in India, including the role of voluntary commitments in this process.

RESEARCH METHODOLOGY

The observations, analysis and recommendations contained in this paper are drawn from extensive interviews with more than a dozen domain experts. To ensure their privacy and safeguard the confidentiality of their ongoing engagements, all interviewees have been granted anonymity. Additionally, the author has surveyed academic and policy literature contained in white papers, policy documents, government memos, academic journals and news reports.



STRUCTURE OF THIS PAPER

The paper is divided into three parts:

PART I

Establishes a conceptual framework for understanding AI self-regulation by defining the concept, highlighting its key features and objectives, and examining its various types with relevant examples.

PART II

Evaluates the suitability of self-regulation in India, considering the country's socio-economic and technological landscape and stated policy objectives based on qualitative interviews with domain experts.

PART III

Outlines strategies for institutionalising an effective AI self-regulatory framework in India through design principles, institutional requirements, and actionable policy recommendations.



PART I

CONCEPTUAL FRAMEWORK FOR AI SELF-REGULATION

WHAT IS AI SELF-REGULATION?

Self-regulation is an approach wherein a group of firms in a particular industry or the entire sector voluntarily adopt certain principles, frameworks, standards or guidelines in order to meet certain regulatory objectives¹⁴. In the context of AI governance, self-regulation typically involves adopting certain principles, such as fairness, transparency, accountability, privacy and safety, which together constitute a set of ‘responsible AI’ practices (sometimes also called ‘ethical’, ‘trustworthy’ or ‘safe’ AI).

In that respect, self-regulation is different from how regulation is commonly understood, which generally refers to a set of rights and obligations established and enforced through a legal framework¹⁵. It is distinct from other approaches to regulation, namely *co-regulation* (joint cooperation between industry players and the government); and *binding regulation* (norm-setting by the government, backed by the force of law)¹⁶.

Many organisations continue to practice self-regulation in the field of AI governance through responsible AI principles, voluntary commitments, and other models, as explained later on in this paper. Compared to other regulatory approaches, self-regulation in the context of AI governance, helps leverage industry expertise, affords flexibility in a dynamic industry, and promotes innovation¹⁷. To that extent, self-regulation should *not* be considered a means to an end or a precursor to binding rules, but rather a comprehensive approach to risk mitigation in itself, with certain advantages over other regulatory models.

ESSENTIAL FEATURES OF SELF-REGULATION

The key features of self-regulation are often misunderstood, leading to its conflation with other regulatory models such as co-regulation. This lack of clarity can create misplaced expectations and undermine the effectiveness of self-regulation in achieving its intended objectives.

To ensure a precise understanding, this paper outlines five essential conditions that must be met for a policy framework to be classified as self-regulation, drawing upon qualitative interviews with industry experts and comparative analyses with other regulatory models.

- 1. Voluntary participation:** Self-regulation means that participation is purely voluntary, i.e. on an ‘opt-in’ basis and should not be driven by coercion or compulsion. While policymakers may encourage participation through various incentives, as explained later in this paper, the use of any legal authority or mandate to compel participation contradicts the core feature of self-regulation.
- 2. Non-enforceable:** Self-regulatory frameworks cannot be legally enforced, i.e. they do not create any rights or obligations which have the force of law¹⁸. Indeed, self-regulation framework may be introduced or adopted under new or existing legal instruments (for example, through advisories or voluntary guidelines issued under existing laws)¹⁹, but they cannot be enforced in a court of law.
- 3. Limited penalties:** Penalties for ‘non-compliance’ with self-regulatory frameworks does not include civil or criminal penalties. To the extent required, industry actors who have committed to self-regulatory measures may be subject to other accountability measures, such as administrative repercussions, expulsion from self-regulatory bodies, and other non-statutory measures.
- 4. Flexibility:** Self-regulatory measures are generally geared towards a specific industry or sector. Therefore, these measures must be tailored to specific actors based on their role, function and value addition. Accordingly, these self-regulatory measures should also be practical and flexible so they can be adapted to the needs of an evolving technology ecosystem.
- 5. Multistakeholder:** Self-regulation primarily relies on industry leadership, but active participation from government bodies and civil society enhances its credibility and effectiveness.

A clear understanding of these essential conditions or prerequisites helps distinguish self-regulation from other regulatory models and ensures that it is applied in a manner consistent with its core principles. With this foundation in place, the next section explores the specific objectives of self-regulation—examining what it seeks to achieve and how it aligns with broader policy and industry goals.



WHAT ARE THE OBJECTIVES OF AI SELF-REGULATION?

The primary objectives of self-regulation for AI governance can be summarised as follows:

- 1. Increasing trust and safety:** An effective self-regulatory framework assures policymakers that business will act responsibly in the development and deployment of their products and comply with local regulations. It also enhances the credibility of participating organisations, which helps facilitate wider adoption of AI²⁰. Some forms of self regulation can also contribute to advancing notions of safety through technical scholarship in areas such as testing and certification of models. Lastly, transparency reporting creates accountability which can increase trust in the ecosystem²¹.
- 2. Developing industry norms:** Self-regulation helps in establishing industry-wide norms for AI governance in areas such as transparency, fairness, privacy, security, and accountability²². Given the fragmented AI ecosystem in India, where firms operate at varying levels of AI maturity²³, and are subject to different sectoral regulations, they help standardise and elevate best practices²⁴.
- 3. Protecting consumers:** When organisations voluntarily establish grievance redressal mechanisms, for example allowing users to report AI misuse in the form of deepfakes or biased decision-making, it increases consumer awareness around these issues and enforceability of their legal rights²⁵.
- 4. Supporting businesses:** By providing clear guidelines on responsible AI principles, self-regulation helps businesses navigate regulatory uncertainties and fosters predictability in the market²⁶. This prevents unnecessary regulatory overreach, assuring firms that binding rules will be introduced only to address critical market failures. Additionally, alignment with global best practices in responsible AI enables Indian firms to expand into international markets²⁷.
- 5. Reducing administrative burden:** A self-regulatory approach alleviates the strain on government bureaucracy and regulators by allowing them to focus on high-risk scenarios and strategic priorities²⁸. By establishing voluntary codes of conduct and industry-led compliance mechanisms, firms can proactively address concerns without direct government intervention. This reduces the need for prescriptive regulations and new institutional frameworks²⁹. Further, it provides policymakers with evidence of market behaviour, regulatory gaps, and risks to inform future regulatory interventions.

WHAT ARE THE DIFFERENT TYPES OF AI SELF-REGULATION?

Since self-regulation is often conflated with other regulatory approaches, it is useful to develop a clear taxonomy for AI self regulation. For this purpose, this paper analyses three types of self-regulation that are currently prevalent in the domain of AI governance, although certain overlaps are possible:

- 1. Responsible AI principles**
- 2. Voluntary commitments**
- 3. Voluntary standards**

These main types of self-regulatory frameworks can be differentiated on the basis of a three-part test outlined in a report published by the Department for Business, Energy and Industrial Strategy, Government of the United Kingdom³⁰. This is explained in the table below:

TABLE 1: THREE MAIN TYPES OF AI SELF-REGULATION FRAMEWORKS

This table provides a classification framework for three main types of AI self-regulation frameworks based on the role of public authorities in the process and whether there is statutory backing.

TYPES OF AI SELF-REGULATION	CLASSIFICATION FRAMEWORK
Responsible AI principles: Formulated by an organisation internally to guide safe and ethical deployment or use of AI by the organisation and its partners.	No statutory obligation and limited role for public authorities and law.
Voluntary Commitments: Formulated by collaborative efforts involving businesses, industry bodies, governments, civil society, etc, and adopted individually or by a group of firms.	No statutory obligation, significant role for public authorities and law.
Voluntary Standards: Commissioned or developed by government agencies, regulators, standard setting bodies or self-regulating organisations (SROs) and adopted voluntarily.	Statutory backing with significant involvement for public authorities.

In this paper, the author outlines three additional factors on the basis of which the three types of AI self-regulatory frameworks described in Table 1 can be analysed, as explained in the table below:

- 1. Government involvement:** What is the role of the government in drafting, collaborating, endorsing, supervising, and implementing the framework?
- 2. Scope and granularity:** What is the breadth (scope) and depth (granularity) of the issues covered in the self-regulatory framework?
- 3. Regulatory pressure:** What is the actual or perceived pressure to comply with the framework?

TABLE 2: EXPLANATION OF DIFFERENT TYPES OF AI SELF-REGULATION FRAMEWORKS

This table explains how the main types of AI self regulatory frameworks compare across three factors, namely the level of government involvement, scope and granularity, and regulatory pressure.

Description	RESPONSIBLE AI PRINCIPLES	VOLUNTARY COMMITMENTS	VOLUNTARY STANDARDS
	<p>Adopted at an organisational level in the form of non-binding guidance for internal governance without any external oversight.</p> <p>Helps differentiate organisational practices, promote consumer trust, and enhance reputation.</p> <p>National governments may also adopt responsible AI principles, for example NITI Aayog's Responsible AI for All principles.</p>	<p>Voluntary AI commitments originate from governments, industry, and multilateral bodies and help shape responsible AI practices. Government-led initiatives establish high-level policy principles; Multilateral agreements align global AI governance efforts across nations; Industry-driven commitments focus on self-regulation; Civil society organizations ensure AI development remains aligned with broader societal values.</p>	<p>Voluntary standards provide detailed, technical and usually prescriptive guidelines for responsible AI. Some standards are developed by international bodies, others are region-specific, tailored by governments or regulatory bodies to address local industry needs. Even when issued by regulators, these standards fall within the scope of self-regulation as long as they remain voluntary and are not legally binding.</p>

	RESPONSIBLE AI PRINCIPLES	VOLUNTARY COMMITMENTS	VOLUNTARY STANDARDS
Examples	Responsible AI principles from companies such as Google, Meta, Microsoft, Amazon, Anthropic and Indian companies such as Corover.ai, Wipro, Infosys, TCS.	Initiatives with government signatories include the Bletchley Declaration, G7 Hiroshima Process, and ASEAN AI Governance Guide. Initiatives with industry signatories include the Seoul AI Business Pledge and Partnership on AI. Civil society contributions include the Asilomar AI Principles by the Future of Life.	Global frameworks like ISO/IEC 42001 and AI testing standards developed by ISO and IEEE. Regional initiatives include the NIST Risk Management Framework (NIST-RMF) in the US and AI Verify in Singapore. In India, the TEC has published draft standards for 'Fairness Assessment and Rating of Artificial Intelligence Systems'.
Government Involvement	No government involvement, except in the case of National AI principles developed by the government (eg. NITI Aayog).	Level of involvement extends from co-development (eg. the US White House Voluntary Commitments) to organising summits and drafting the declaration (Bletchley Summit).	Technical representatives from government agencies are generally involved in drafting. In other cases, independent standard setting bodies or technical groups may lead.
Scope and Granularity	High-level principles on fairness, safety, privacy, transparency, and accountability, with details on implementation in extraordinary cases (for e.g. safety commitments for frontier models).	Voluntary AI commitments could be in the form of high-level principles around human oversight, safety against misuse, privacy and security, inclusivity, sustainable development, etc. and may also address specific issues like child safety, deep fakes, and election integrity.	Largely limited to key responsible AI principles, such as fairness, safety, privacy, transparency, and accountability. However, they may be highly technical and prescriptive, providing detailed guidance on compliance.

	RESPONSIBLE AI PRINCIPLES	VOLUNTARY COMMITMENTS	VOLUNTARY STANDARDS
Regulatory pressure	Pressure to comply with responsible AI principles and practices is generally considered to be low given the lack of external oversight. However, depending on the level of transparency in reporting and compliance, there may be peer-pressure from other industry actors, consumers who may file lawsuits, and from civil society.	Pressure to comply with voluntary commitments is generally considered low to medium given the lack of legal enforceability. However, some types of commitments, especially those developed in partnership with the government have a higher likelihood of compliance.	Pressure to comply with voluntary standards depends on the level of prescriptiveness and the nature of the body that issues the standards. For example, regulatory pressure is deemed to be higher if issued by the Reserve Bank of India (RBI) compared to the Telecom Engineering Center (TEC).

The decision of policymakers to adopt a specific type of regulatory model depends on a variety of factors, including the role of the state, the desire for compliance, the level of detailing required or expected, the extent of government involvement in developing and supervising the framework, and other such factors.

In the next section, the author examines these factors in the context of India's socio-economic and technological landscape and stated policy objectives in relation to AI governance.

PART II

STAKEHOLDER SENTIMENTS ON AI SELF REGULATION

GENERAL SENTIMENTS

An analysis of stakeholder sentiments suggests there is broad consensus on how India should approach AI governance, with some divergence³¹. The support for AI self-regulation stems from a variety of reasons:

1. A developing country like India should be focused on harnessing the benefits of AI, rather than over-indexing on the potential risks, which may result in strict regulations that could stifle innovation³²;
2. India's unique socio-economic features, such as its cultural diversity, necessitate context-specific solutions that are better addressed through flexible, industry-led voluntary initiatives³³.
3. Voluntary commitments can co-exist with other regulatory models so it is not a binary decision.

Accordingly, below are some developing narratives that reflect the views of government, industry and civil society respectively in the context of India's increasing reliance on self-regulation for AI risk mitigation:

SELF-REGULATION ALIGNS WITH THE GOVERNMENT'S AI OBJECTIVES

The Indian government has expressed broad support for self-regulation, through media interviews³⁴, committee reports³⁵ and parliamentary responses³⁶. Sectoral regulators such as the Reserve Bank of India³⁷ and the Indian Council of Medical Research³⁸ also support self-regulation in their respective domains. Meanwhile, MeitY continues to implement the “Safe and Trusted AI” pillar of the IndiaAI Mission through indigenous AI safety tools, self-assessment frameworks, and voluntary guidelines³⁹.

The rationale is that self-regulation closely aligns with the Indian government's strategic priorities of fostering innovation, driving economic growth and democratising access to AI. Rather than adopting a rights-heavy, prescriptive model, as seen in the EU, India prioritises economic impact, innovation, and addressing immediate challenges like bias and misinformation through existing laws and advisories⁴⁰.

The government has three objectives it hopes to achieve through AI self-regulatory frameworks⁴¹:

- a. Firms should voluntarily comply with existing laws and regulations and demonstrate compliance when called upon to do so.
- b. They should adapt global AI principles and voluntary risk mitigation efforts to align with India's local socio-cultural context and value systems.
- c. Other strategic objectives, around multi-lingual support, democratising access to compute and related goals should also be reflected in voluntary commitments.

Finally, there is a tacit understanding that the Indian government is closely monitoring the efficacy of self-regulation. Should such self-regulatory efforts prove ineffective, or if there is evidence of market failure, the government will introduce binding rules to address these gaps⁴².

INDUSTRY ACTORS SUPPORT AI SELF-REGULATION

The prevailing sentiment in industry is that India should adopt a two-tiered approach to AI governance⁴³:

Level 1: Self-regulation that enables firms to proactively address the risks of AI through voluntary commitments, self-certifications, and similar models.

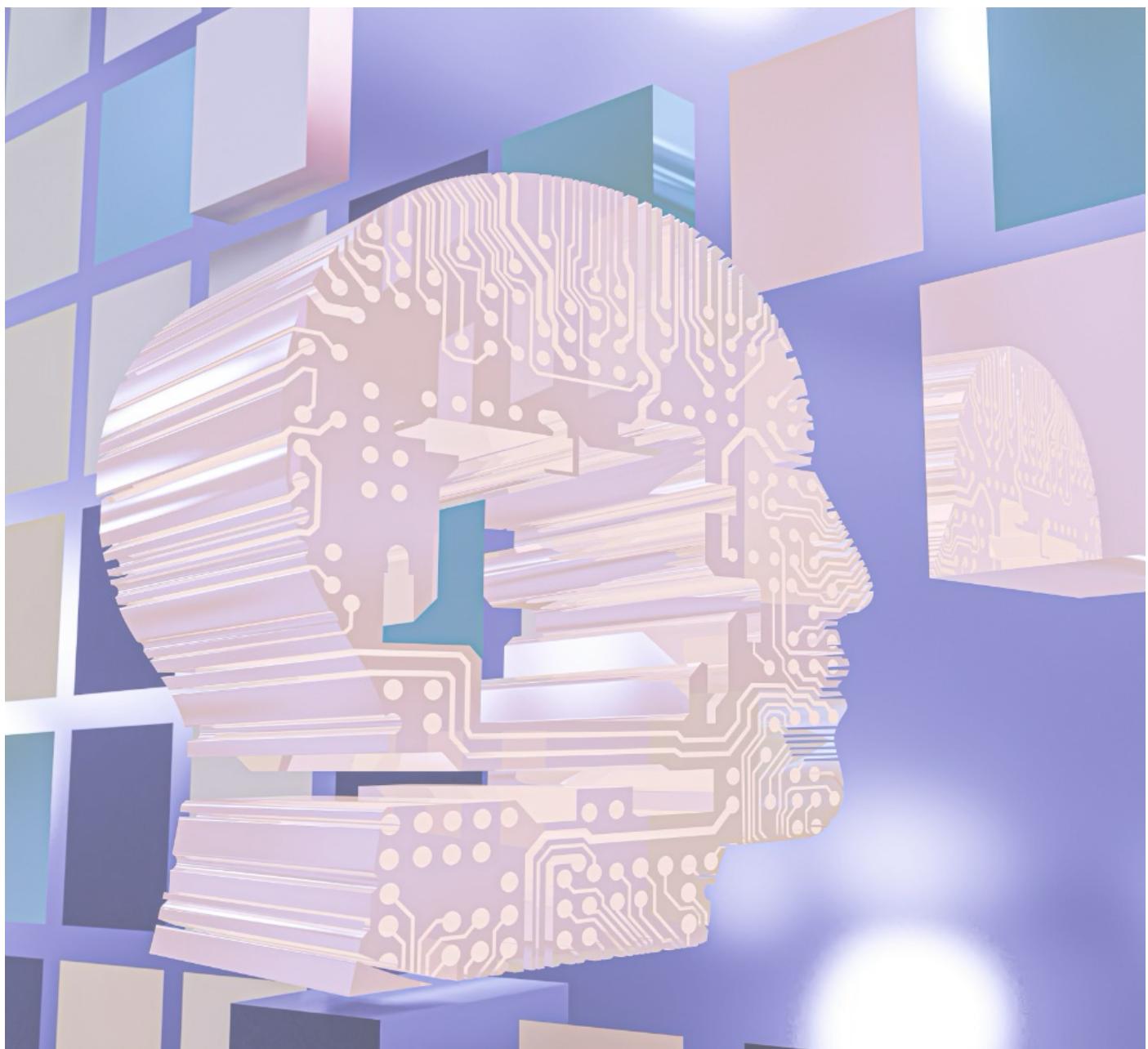
Level 2: Additional regulations may be required to fill the legal vacuum and deal with high-risk AI use cases through bespoke rules, guidelines, and advisories.

Another view is that the government should give credence to various forms of self-regulation that are already being implemented by the industry through responsible AI practices, self-certifications and voluntary commitments⁴⁴. Therefore, any new or proposed framework must add incremental value⁴⁵.

Finally, some executives remain apprehensive about new forms of self-regulation that do not meet the established norms of self regulation, as outlined earlier in this paper. Specifically, they insist that their participation is contingent on these proposed frameworks being purely voluntary and non-enforceable⁴⁶.

CIVIL SOCIETY IS SKEPTICAL OF AI SELF-REGULATION

Skepticism about self-regulation abounds in civil society groups. In a nutshell, they argue that it is self-serving⁴⁷, and unlikely to change harmful behaviour given the lack of accountability⁴⁸. In particular, they point to existing gaps in how these companies handle personal data, deal with bias against marginalised groups and the lack of transparency in how these risks are identified and mitigated at the country level⁴⁹. While many civil society experts believe that self-regulation can be useful to meet some objectives, it must be complemented with binding rules in parallel for effective oversight and accountability⁵⁰.



RISK CLASSIFICATION FOR AI SELF-REGULATION

There are two areas of contention on risk classification and AI self-regulation. One issue is whether, based on their risk profile, voluntary frameworks should apply to models or applications, or both. For example, the voluntary AI commitments secured by the Biden administration applies to certain frontier models⁵¹.

Another polarising issue is whether self-regulation should apply to low-risk, high-risk applications, or both. This is an important distinction because the design of the proposed self-regulatory framework depends on which type of risk profile is attended to. Broadly, there are two viewpoints:

The first camp advocates for self-regulatory frameworks only in low-risk areas for the following reasons:

- Self-regulation in low-risk areas allows for gradual development of best practices, before their effectiveness can be evaluated and extended to high-risk scenarios, if at all.
- The potential for unintended consequences is less severe with low-risk use cases, providing a safer environment for iterative experimentation and refinement.

The second camp argues for self-regulatory frameworks to be extended to *high-risk areas* since:

- Low-risk areas are already governed by existing responsible AI practices.
- There is greater incentive for compliance given the reputational costs involved.
- Regulators will have a better sense of whether binding rules are necessary for certain high-risk cases after supervising voluntary efforts over a period of time.

Given the lack of consensus, industry bodies continue to develop self-regulatory models for both high and low-risk AI systems⁵², while awaiting regulatory guidance on risk classification. Further, as recommended by the MeitY sub-committee on AI governance, organisations should be encouraged to actively contribute to an incidents database to inform the development of risk classification frameworks⁵³. Experts are also of the view that civil society organisations in India should be included in helping to develop a robust risk classification framework given their on-ground experience in tracking AI related harms⁵⁴.

ADAPTING GLOBAL AI PRINCIPLES TO THE LOCAL CONTEXT

A significant challenge in developing AI self-regulatory frameworks lies in reconciling global principles, such as fairness, inclusivity, and safety, with local socio-cultural nuances.

Indian officials have categorically stated that global principles, embedded in multilateral frameworks and voluntary commitments, are insufficient to address the specific risks relevant to India⁵⁵. For instance, the concept of bias in the Indian context, influenced by factors such as caste, religion and linguistic diversity, differs significantly from its treatment in other jurisdictions⁵⁶. Officials expect a more localised approach to bias mitigation, respect for linguistic diversity and training of AI models on locally relevant datasets⁵⁷.

While industry leaders acknowledge the need for local adaptation, they cite the following key challenges⁵⁸:

- It requires a deep understanding of local socio-cultural contexts (hate speech and bias detection across multiple languages, for example), which is operationally challenging.
- Providing multilingual support often requires changes to product roadmaps, which is time consuming and requires a staggered approach.
- The proliferation of self-regulatory frameworks has led to ‘commitment fatigue’, making it difficult for firms to deploy resources to meet local requirements in each and every country.

Industry representatives suggest that policymakers should be cognisant of technical and operational limitations involved in localising global principles, and the need for additional incentives and local evaluation datasets to meet these expectations. Meanwhile companies should be allowed to demonstrate adherence to global AI standards and principles like the G7 Code of Conduct and OECD guidelines⁵⁹.

APPLYING SELF-REGULATION TO SPECIFIC ENTITIES

The scope of any self-regulatory framework is shaped by its intended targets. Experts suggested that voluntary measures should be tailored to the target audience, whether developers, deployers and users⁶⁰.

One expert⁶¹ suggested a more holistic and multi-tiered approach, by tailoring frameworks for:

- Organisations (for eg., AI firms engaged in development and deployment)
- Systems (for eg., broader ecosystems of developers and third-party collaborators).
- Individuals (for eg., manuals aimed at CXOs or other leadership roles tasked with oversight)

Additionally, a phased implementation may be useful, wherein self-regulatory frameworks are anchored around a few key players in the ecosystem, both domestic and global, and gradually expand to include other actors⁶².

In effect, all ecosystem participants should meet certain common minimum requirements to encourage behavioral changes in favor of safe and secure development and deployment of AI. Above these common minimum requirements, certain actors may adopt a broader range of voluntary commitments based on their risk profile. This is similar to negotiations at the World Trade Organisation, where agreements are progressively broadened to accommodate diverse participants, although the end objectives remain the same⁶³. The US government also followed this approach in securing voluntary AI commitments⁶⁴.

REGULATING THE INPUTS OF AI SYSTEMS

There is broad agreement that self-regulation should cover AI-generated ‘outputs’ that are unlawful or harmful. However, experts are divided on whether self-regulation should also address issues relating to ‘inputs’ of AI systems, such as training data, copyright and privacy issues.

Some stakeholders stated that industry codes under the Digital Personal Data Protection Act (DPDP Act) and its accompanying regulations, would be a key to ensuring effective implementation of the law⁶⁵. Self-regulatory codes may be particularly important in cases where input data feeds into high-risk use cases, such as for medical diagnosis — and transparency is expected in relation to what types of data are involved, how it is stored, who it is shared with etc. to enable data collection through informed consent.

At the same time, industry stakeholders expressed concerns that such transparency measures, including the use of model cards, could pose risks to intellectual property

and trade secrets, and might create new security vulnerabilities⁶⁶. In other cases, such as the use of copyright protected data to train AI systems, experts suggested amending the Indian Copyright Act rather than adopting self-regulatory frameworks⁶⁷.

KEY TAKEAWAYS

There is broad agreement within government and industry that self-regulation has an important role to play in the development of safe AI, although civil society remains skeptical about its effectiveness. However, there continues to be friction on how it should be developed and implemented, on the following counts:

- There is disagreement on whether self-regulation should cover both low-risk and high-risk use cases, although it generally tilts in favour of gradual expansion to high-risk use cases after careful evaluation of its effectiveness in low-risk scenarios.
- Government officials have stated that localisation of global AI principles to the Indian context is expected, but this position is being resisted by industry leaders for various operational reasons. Market incentives could help fast track such local adaptations in some areas.
- Ideally, self-regulatory frameworks should be tailored and applied to specific target entities (developer v. deployer, for example), although starting with general principles and applying them to certain ‘anchor firms’ can help mobilise the entire ecosystem.
- To what extent self-regulation should cover the inputs of AI systems remains an open question – with some suggesting legislative amendments, for example in relation to copyright law.

In the next section, the author outlines a set of policy recommendations to inform the development and implementation of self-regulatory AI governance frameworks in India, based on a review of the conceptual framework in Part I and the sentiments of key stakeholders outlined in Part II of the paper.

PART III

RECOMMENDATIONS FOR EFFECTIVE AI SELF REGULATION IN INDIA

DEVELOP A RISK-CLASSIFICATION FRAMEWORK TO INFORM AI SELF-REGULATION

To start with, a clear risk classification framework is crucial to inform AI self-regulatory frameworks. It would help policymakers develop frameworks customised based on the risk profile and to evaluate if any additional binding rules are necessary beyond voluntary frameworks.

In order to develop such a framework, policymakers should focus on the following:

- 1. Collect evidence of harm to quantify risk** arising from AI deployments, grounded in the Indian context, through a combination of empirical data, case studies and incident reports⁶⁸.
- 2. Review existing risk frameworks** such as MIT's AI Risk Repository⁶⁹, the International Scientific Report on the Safety of Advanced AI⁷⁰, global benchmarks from NIST⁷¹ and ISO⁷², and frameworks from the TEC⁷³ and IndiaAI to develop a risk classification framework suited for India⁷⁴.
- 3. Classify AI systems or applications** on the basis of specific risk vectors, including for example (1) the risk of malicious uses; (2) algorithmic discrimination; (3) transparency failures; (4) systemic risks; (5) loss of control; (6) national security and public safety risks⁷⁵.

A comprehensive risk-classification framework, once developed, can be incorporated into future regulatory models in the future, including co-regulation and binding rules, should the need arise.

It should be noted that certain risks are qualitatively different from others in that they differ in their urgency, probability and likelihood of causing harm⁷⁶. In that respect, certain categories of risks, such as Chemical, Biological, Radiological, and Nuclear ('CBRN') threats emanating from AI development and use pose a greater risk and will likely require binding rules beyond self regulatory frameworks. On the other hand, certain 'existential risks', such as the risk of 'superintelligence'⁷⁷ may not require any regulatory interventions at this stage because the risks are unclear and there is no tangible evidence of harm.

GOVERNMENT SHOULD BE CLOSELY INVOLVED IN AI SELF-REGULATION

The level of government involvement in developing and implementing AI self-regulatory frameworks is a matter of considerable debate. Drawing from expert opinions in India and global experiences with AI self-regulation, the author recommends that the government be closely involved in at least four key areas:

- 1. Initiating the process:** Proactive action from the government in calling for a self-regulatory framework can act as a ‘forcing function’, especially for firms who prefer the *status quo*⁷⁸. It can also help set expectations and attract a wider group of participants. Moreover, without a clear signal from the government, there is often little incentive for companies to adopt voluntary commitments.
- 2. Drafting voluntary codes:** The government’s involvement in drafting AI self-regulatory codes varies across jurisdictions. In some cases, the government has taken the lead⁷⁹; in many others, they have been jointly developed⁸⁰, and in other cases, they have been largely developed by industry or civil society organisations⁸¹. Based on expert interviews, it appears that government-led drafts are generally more representative of multi-stakeholder views and are therefore more likely to succeed⁸².
- 3. Building consensus:** Consensus-building is critical to resolving conflicts among stakeholders, particularly in contexts where interests diverge, such as between domestic and foreign firms. Industry experts have highlighted that government involvement is often necessary to mediate such differences and prevent monopolisation of the process by a few dominant players⁸³.
- 4. Endorsing frameworks:** In some cases, explicit government endorsement of industry-led voluntary frameworks lends legitimacy to the process and can help in adoption and compliance⁸⁴, although some argue that any formalisation through legal mandates would be a form of co-regulation⁸⁵.

For these reasons, proactive government engagement in conceptualising, developing, facilitating, and endorsing any proposed self-regulatory framework can be useful to ensure its effective operationalisation⁸⁶.

MARKET INCENTIVES ARE REQUIRED TO MAKE AI SELF-REGULATION WORK

Self-regulation is unlikely to work unless adequate market incentives are built in to encourage participation and good behaviour. Based on secondary research and interviews with domain experts, below is a list of incentives that must be incorporated to ensure the success of AI self-regulatory frameworks in India:

- **Financial incentives** are arguably the most important. For example, the government may incentivise adoption by incorporating voluntary commitments into lucrative public procurement contracts⁸⁷. Another method is donor-led, in which venture capitalists and private equity firms partner with the government to offer cloud credits or certifications to investee companies that sign on to voluntary AI commitments. Lastly, grants issued under the IndiaAI mission could also be made conditional upon such participation.
- **Regulatory incentives** also hold promise as they directly relate to certain objectives. For example, participation in regulatory sandboxes can be made contingent on adopting voluntary guidelines. On the other hand, providing clear guarantees to firms about intellectual property rights protections for voluntary disclosures is another critical incentive to maximise participation⁸⁸.
- **Technical incentives** are necessary to promote wider adoption of self-regulatory frameworks. For example, the development of simple, user-friendly tools such as software development kits and the ease with which they can be integrated into their existing applications and services is an important motivator, especially for resource-crunched startups to adopt voluntary AI frameworks⁸⁹.
- **Reputational incentives** play a significant role, as the recognition that comes from publicly showcasing voluntary commitments can enhance an organisation's trustworthiness⁹⁰.

To incorporate these incentives into market dynamics, policymakers will have to address three main issues: *firstly*, a large number of target actors significantly reduces the feasibility of creating effective incentives, whereas a smaller number of participants allow for more tailored and effective incentive mechanisms; *secondly*, incentives must be domain-specific, taking into account factors such as regulatory maturity, the existing regulatory landscape, and the risk level associated with the sector; *thirdly*, different incentives may have to be designed for different actors, namely small versus large companies, domestic versus foreign entities, consumer-facing services versus enterprise companies, and so on.

ALTERNATIVE COMPLIANCE MECHANISMS CAN ENHANCE ACCOUNTABILITY

A central challenge of self-regulation lies in holding firms accountable in the absence of legal enforceability. In other words, firms must face meaningful pressure to conform.

Compliance with self-regulatory frameworks generally depends on the credibility of the supervising entity and the availability of alternative mechanisms to demonstrate compliance. The pressure to conform is often influenced by the perceptions of industry stakeholders regarding these regulators. For instance, MeitY and the RBI command significant authority, given their ability to assert their statutory authority and make legislative interventions, as compared to the TEC⁹¹.

A variety of alternative mechanisms may be utilised to enhance accountability at both the organisational and industry-level, as explained below:

- **Transparency reports:** Organisations publish transparency reports, with details of red-teaming reports, impact assessments and risk mitigation efforts, allowing the public to scrutinise their practices. At the sectoral-level, industry organisations can submit periodic reports to regulators, following models such as the RBI's SRO-FT Guidelines⁹².
- **Internal policies:** Organisations revise their service terms, internal governance frameworks, and operational guidelines to align with voluntary commitments and industry standards.
- **Self-certifications:** Organisations seek certification from independent auditors or standard-setting bodies, who possess established infrastructure and processes for assessing compliance, as currently practiced in the gaming industry.
- **Peer monitoring:** Industry and civil society monitor the actions of ecosystem participants and report violations, creating informal but effective accountability mechanisms⁹³.
- **Techno-legal measures:** Organisations adopt measures wherein regulatory principles are embedded into the design of the system itself, (for eg. consent managers)⁹⁴.
- **Committee hearings:** Regulators and parliamentarians conduct hearings and ask for information to evaluate compliance with voluntary commitments.

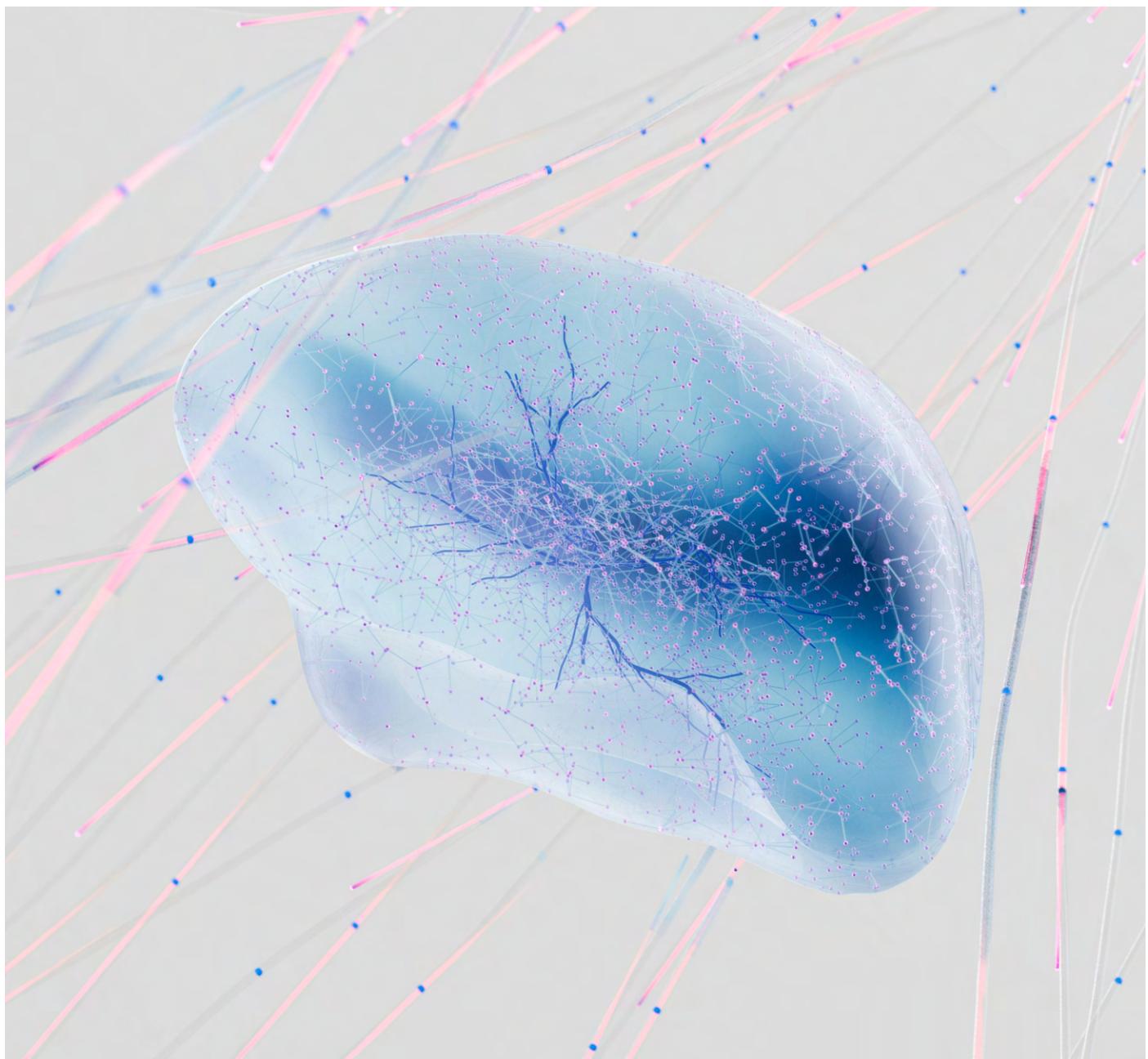
These mechanisms collectively enhance accountability by fostering transparency, encouraging industry-wide cooperation, and incentivising compliance through non-legal mechanisms.

INSTITUTIONAL CAPACITY MUST BE STRENGTHENED TO MAKE SELF-REGULATION EFFECTIVE

Another key element to making self-regulation effective is a strong institutional framework. In the context of AI self-regulation specifically, there are at least three types of bodies that may be necessary to supplement the existing capacity constraints:

- 1. AI Safety Institute:** MeitY recently announced the setting up of an AI Safety Institute for India (AISI)⁹⁵. The proposed AISI can play a pivotal role in strengthening self-regulation by establishing benchmarks, evaluating and comparing transparency reports, and providing guidance to participants on compliance with voluntary guidelines⁹⁶. From an India perspective, focussing on post-deployment impacts on vulnerable communities, gathering evidence of harm in the local context, and enhancing AI literacy could be useful⁹⁷. Moreover, the initial focus of the AISI should be on research, testing and standardisation rather than rulemaking or enforcement⁹⁸.
- 2. Technical Advisory Council:** There is a gap in technical expertise within existing institutions that could compromise the goals of AI self-regulation⁹⁹. For that reason, MeitY's draft report on AI governance suggests the creation of a 'Technical Secretariat' to support industry self-regulation¹⁰⁰. There is merit in this recommendation. A government-led body with sufficient technical expertise could issue clear and timely guidance and support the development of benchmarks and standards to guide voluntary industry compliance. MeitY's draft report has limited details on what the structure and functions of the proposed Technical Secretariat could be. NITI Aayog's proposal for a Council for Ethics and Technology (CET) is instructive in this regard. It suggests the creation of an independent, multi-disciplinary advisory body to guide AI policy formulation, produce research, and develop risk assessment frameworks and guidelines, similar to the Campus for Research Excellence and Technological Enterprise (CREATE) in Singapore or Innovate UK¹⁰¹.
- 3. Self Regulatory Organisations (SROs):** Industry-led SROs can promote accountability by enforcing self-regulatory codes through bye-laws and other internal governance mechanisms. While in India, there is a suggestion for

SROs in the financial sector¹⁰², there is considerable skepticism on the need for SROs for effective AI governance, for a few reasons: *first*, SRO's in sectors like online gaming, digital media and OTT streaming have faced multiple challenges with the government¹⁰³; *second*, there is a lack of clear legal authority to set up SROs, for example in the case of SROs to tackle misinformation relating to non-government information under the IT Rules¹⁰⁴; and *third*, SROs are sometimes seen as gatekeepers, having disproportionate power in relation to market entry, for example in the online gaming industry¹⁰⁵. For these reasons, it may be premature to establish an SRO for AI governance in India, but could be taken up at a later point in time.



CONCLUSION

Self-regulation presents both an opportunity and a challenge. This paper outlines a structured approach to making AI self-regulation work by defining its core principles, assessing stakeholder sentiments, and providing a roadmap for effective implementation. While self-regulation remains a useful strategy, its success hinges on the right mix of incentives, institutional frameworks, and accountability measures.

As **Part I** illustrates, self-regulation must be clearly distinguished from co-regulation and binding rules to ensure clarity in expectations and responsibilities. This paper identifies the essential conditions for self-regulation and highlights the critical role of voluntary participation and flexibility.

Part II captures the perspectives of key stakeholders—revealing both alignment and friction on the scope and effectiveness of self-regulation. While industry largely supports voluntary commitments, civil society remains skeptical of their sufficiency. The role of the government in fostering self-regulation while preparing for potential future interventions remains a key point of discussion.

As the comparative analysis in **Part III** demonstrates, India must craft its own self-regulatory model tailored to its unique socio-economic and technological landscape. While global principles provide a reference point, their local adaptation is essential for meaningful implementation. The government's involvement—whether in initiating frameworks, facilitating multi-stakeholder dialogue, or endorsing voluntary commitments—will be crucial in shaping AI self-regulation in India.

This paper recommends a phased approach to self-regulation, beginning with voluntary commitments and gradually expanding to sector-specific standards. Voluntary commitments should be strategically designed to support norms development, standardisation and wider adoption. To ensure effectiveness, they must adopt a light touch approach, remaining principle-based and practical in their implementation. AI risk classification should serve as a guiding principle throughout, ensuring that self-regulation is applied where appropriate while allowing space for more structured oversight where necessary.

Finally, and most critically, AI governance in India must be inclusive and participatory. The voices of industry leaders, policymakers, researchers, and civil society must shape the trajectory of AI self-regulation. Most immediately, a transparent consultation process should be undertaken demonstrating both government and industry leadership to make AI self-regulation work for India.

COPYRIGHT

© 2025 CeRAI – Centre for Responsible AI. All rights reserved.

This document and its contents are the intellectual property of CeRAI and are intended solely for informational and non-commercial use. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means without prior written permission from CeRAI. All referenced trademarks and materials are the property of their respective owners.

ABOUT THE AUTHOR

Amlan Mohanty is an Associate Research Fellow at the Centre for Responsible AI (CeRAI), IIT Madras. He is an independent lawyer, policy advisor, researcher and teacher with over a decade's experience working in India's technology sector. He is also a Non-Resident Fellow at Carnegie India, Non-Resident Fellow at NITI Aayog and Visiting Professor at the National Law School of India University. Previously he led public policy and government affairs for Google in India across privacy, platforms and AI.

ACKNOWLEDGEMENTS

The author would like to thank the interviewees who graciously contributed their time and knowledge to this paper without attribution. The author would also like to thank Professor Balaraman Ravindran for his feedback on a draft of this paper, Kevin Preji for research assistance, Omir Kumar and Geetha Raju for copy editing, and Tricada for visual design.

DISCLOSURE

This is an independent piece of scholarship informed by primary and secondary research. However, in the interests of transparency, the author would like to disclose that he consults with various organisations who could benefit from the findings of this paper. A full statement on ethics and disclosures is available on the author's website: www.amlanmohanty.com

DISCLAIMER

The views expressed in this paper do not reflect the views or positions of IIT Madras, Carnegie India, NITI Aayog, National Law School of India University or any other organisation to which the author is currently or formerly affiliated.

© 2025 Centre for Responsible AI, Indian Institute of Technology, Madras. All rights reserved.

REFERENCES

- ¹Amlan Mohanty is an Associate Research Fellow at the Centre for Responsible AI (CeRAI), IIT Madras.
- ²White House Office of Science and Technology Policy, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (The White House, October 2022), <https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/>. See also “Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI” (The White House, July 21, 2023), <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> and SZCZEPANSKI Marcin, “United States approach to artificial intelligence” European Parliament Research Service, January 17, 2024, [https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA\(2024\)757605](https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2024)757605).
- ³“Model Artificial Intelligence Governance Framework Second Edition,” InfoComm Media Development Authority and Personal Data Protection Commission Singapore, January 21, 2020, <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf>. See also “Singapore’s Approach to AI Governance,” Personal Data Protection Commission Singapore, November 03, 2023, <https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework>.
- ⁴“Social Principles of Human-Centric AI,” Cabinet Secretariat, Government of Japan, 2019, <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>. See also Hiroki Habuka, “New Government Policy Shows Japan Favors a Light Touch for AI Regulation,” Centre for Strategic and International Studies, February 25, 2025, <https://www.csis.org/analysis/new-government-policy-shows-japan-favors-light-touch-ai-regulation>.
- ⁵Department for Science, Innovation & Technology, UK Government, AI Regulation: A Pro-Innovation Approach (White Paper, March 2023), <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.
- ⁶Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, 19 <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>. See also “Recommendations on Leveraging Artificial Intelligence and Big Data in Telecommunication Sector,” Telecom Regulatory Authority of India, Government of India, July 20, 2023, 37 onwards, https://www.trai.gov.in/sites/default/files/2024-09/Recommendation_20072023.pdf.
- ⁷Ministry of External Affairs (MEA), “G20 New Delhi Leaders’ Declaration” (New Delhi: MEA, 2023), 24 <https://www.meaindia.gov.in/Images/CPV/G20-New-Delhi-Leaders-Declaration.pdf>.
- ⁸“Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025, 17-18, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf>.
- ⁹Government of India, Ministry of Electronics and Information Technology, Guidelines for Organisations Working on AI, Rajya Sabha, Unstarred Question No. 1343 (December 6, 2024), https://sansad.in/getFile/annex/266/AU1343_RxE4hh.pdf?source=pqars.
- ¹⁰National Strategy for Artificial Intelligence #AIForAll, report, (NITI Aayog, Government of India, June 2018), <https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf>. See also Responsible AI #AIForAll, Approach Document for India: Part 1 – Principles for Responsible AI,

report, (NITI Aayog, Government of India, February 2021), <https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf> and Responsible AI #AIForAll, Approach Document for India: Part 2 – Operationalizing Principles for Responsible AI, report, (NITI Aayog, Government of India, August 2021), <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>.

¹¹Government of India, Ministry of Electronics and Information Technology, Cyber Law and Data Governance Group, Due Diligence by Intermediaries/Platforms under the Information Technology Act, 2000, and Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, (Issued on March 15, 2024), <https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf>. See also “Recommendations on Leveraging Artificial Intelligence and Big Data in Telecommunication Sector,” Telecom Regulatory Authority of India, Government of India, July 20, 2023, https://www.trai.gov.in/sites/default/files/2024-09/Recommendation_20072023.pdf and Sanjeev Sanyal, Pranav Sharma, and Chirag Dudani, “A Complex Adaptive System Framework to Regulate Artificial Intelligence,” Economic Advisory Council to the Prime Minister White Paper, No. 26, January 2024, https://eacpm.gov.in/wp-content/uploads/2024/01/EACPM_AI_WP-1.pdf.

¹²“Our Principles,” Google AI, <https://ai.google/responsibility/principles/>. See also “Principles and Approach,” Microsoft, <https://www.microsoft.com/en-us/ai/principles-and-approach> and “Connect 2024: The responsible approach we’re taking to generative AI” Meta, September 25, 2024, <https://ai.meta.com/blog/responsible-ai-connect-2024/#:~:text=Meta%27s%20five%20pillars%20of%20responsible%20AI%20that%20inform%20our%20work&text=Everyone%20should%20be%20treated%20fairly,equally%20well%20for%20all%20people.&text=AI%20systems%20should%20meet%20high,behave%20safely%20and%20as%20intended> and “Embracing Responsible AI: A Framework for Ethical Innovation,” Infosys, <https://www.infosys.com/services/data-ai-topaz/insights/embracing-responsible-ai.html> and “Four Dimensions of Responsible AI,” Wipro, <https://www.wipro.com/ai/four-dimensions-of-responsible-ai/>.

¹³“The Developer’s Playbook for Responsible Ai in India,” NASSCOM, November 2024, <https://nasscom.in/ai/pdf/the-developer%27s-playbook-for-responsible-ai-in-india.pdf>

¹⁴OECD, “Industry Self Regulation: Role and Use in Supporting Consumer Interests”, OECD Digital Economy Papers, No. 247 (2015), https://www.oecd.org/content/dam/oecd/en/publications/reports/2015/03/industry-self-regulation_g17a260e/5js4k1fjqkwh-en.pdf

¹⁵Margot Priest, “The Privatization of Regulation: Five Models of Self-Regulation,” Ottawa Law Review 29, no. 2 (1998): 233-302 <https://heinonline.org/HOL/Page?handle=hein.journals/ottlr29&id=239&collection=journals&index=->.

¹⁶Florence G’sell, “Regulating under Uncertainty: Governance Options for Generative AI,” October 6, 2024, <https://doi.org/10.2139/ssrn.4918704>.

¹⁷Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, 19 <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>.

¹⁸Philip Eijlander, “Possibilities and Constraints in the Use of Self-Regulation and Co-Regulation in Legislative Policy: Experiences in the Netherlands - Lessons to Be Learned for the EU?” Electronic Journal of Comparative Law 9, no. 1 (January 2005) <https://pure.uvt.nl/ws/files/704952/possibilities.pdf>.

¹⁹Government of India, Ministry of Electronics and Information Technology, Cyber Law and Data Governance Group, Due Diligence by Intermediaries/Platforms under the Information Technology Act, 2000, and Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, (Issued on March 15, 2024), <https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf>

²⁰“Actively engaging multiple interdisciplinary stakeholders to integrate a diversity of views and expertise in the design and development of AI self-regulatory tools could increase buy-in and adoption.” See “Exploring ways to regulate and build trust in AI,” Rand Europe, April 25, 2022, <https://www.rand.org/randeurope/research/projects/2022/exploring-ways-to-regulate-and-build-trust-in-artificial-intelli.html>

²¹Joachim Roski, Ezekiel J Maier, Kevin Vigilante, Elizabeth A Kane, Michael E Matheny, Enhancing trust in AI through industry self-governance, Journal of the American Medical Informatics Association 28, no.7, (July 2021): 1582–1590, <https://doi.org/10.1093/jamia/ocab065>.

²² White House, Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI (The White House, July 21, 2023), <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> and “AI Principles,” Organisation for Economic Co-operation and Development, <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>, and “Seoul AI Business Pledge,” AI Seoul Summit, May 05, 2024, <https://aiseoulsummit.kr/press/?mod=document&uid=43> and “Frontier AI Safety Commitments, AI Seoul Summit 2024,” Department for Science, Innovation & Technology, Government of United Kingdom, February 7, 2025, <https://www.gov.uk/government/publications/frontier-ai-safety-commitments-ai-seoul-summit-2024/frontier-ai-safety-commitments-ai-seoul-summit-2024>.

²³ “AI Adoption Index 2.0: Tracking India’s Sectoral Progress in AI Adoption” NASSCOM, August 2024, <https://nasscom.in/knowledge-center/publications/ai-adoption-index-20-tracking-indias-sectoral-progress-ai-adoption>

²⁴ “Because standard setting and identification of breaches are the responsibility of practitioners with detailed knowledge of the industry, this will arguably lead to more practicable standards, more effectively policed,” See Gunningham, Neil, and Joseph Rees. “Industry Self-Regulation: An Institutional Perspective.” Law & Policy 19, no. 4 (1997): 363–414. <https://doi.org/10.1111/1467-9930.t01-1-00033>.

²⁵ “INDUSTRY SELF-REGULATION: ROLE AND USE IN SUPPORTING CONSUMER INTERESTS,” OECD Digital Economy Papers, March 1, 2015, https://www.oecd.org/en/publications/industry-self-regulation_5js4k1fjkw-en.html.

²⁶ Andreja Marusic, Corey Owens and Madelynne Grace Wager, “Businesses for the Business Environment: How Companies like Yum! Brands can Improve Compliance through Self-Regulation” World Bank Blogs, February 20, 2018, <https://blogs.worldbank.org/en/psd/how-companies-yum-brands-can-improve-compliance-through-self-regulation>.

²⁷ “Responsible AI has gathered attention around the world and there is an increasing recognition for international collaboration. The GPAI has a working group on responsible AI” Responsible AI #AIForAll, Approach Document for India: Part 2 – Operationalizing Principles for Responsible AI, report, (NITI Aayog, Government of India, August 2021), 23, <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>. See also “Recommendations for Regulating AI,” Google, <https://ai.google/static/documents/recommendations-for-regulating-ai.pdf>.

²⁸ “Self-regulation complements legal frameworks by raising and harmonizing industry standards. It helps to

reduce the burden on the government while providing industry with a means of enforcing Europe-wide rules to which companies must adhere.” See Julie Bonhomme, “The Origins and Privileges of Self-Regulation.” European Federation of Pharmaceutical Industries and Associations, December 11, 2024.

²⁹ Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>.

³⁰ Department for Business, Energy & Industrial Strategy (BEIS), “Designing self- and co-regulation initiatives: best practice - a literature review,” BEIS Research Paper, no. 25, (2019) <https://assets.publishing.service.gov.uk/media/5da9bd8be5274a5cac42150c/designing-self-co-regulation-initiatives.pdf>

³¹ See Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, 3-7, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>

³² Interviews with a senior government official, an executive at a global technology company and an executive at a leading Indian industry association (Interviews 5, 11, 12). See also “Panel: Intelligent Regulation of AI: The Way Ahead,” Youtube video posted by “Carnegie India,” December 11, 2023, Accessed March 11, 2025, <https://www.youtube.com/watch?v=5HCyWC7gjh0>

³³ Interviews with a partner at an Indian technology law and policy firm, and a leader in a community of technologists and policymakers (Interviews 3, 9). See also “Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025,17-18, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf>.

³⁴ Simone Lobo, “Heavy Regulation on AI like in the US, Europe Not the Way to Go for India: IT Minister Ashwini Vaishnaw,” MediaNama, October 7, 2024, <https://www.medianama.com/2024/10/223-ai-regulation-data-privacy-law-approach-it-minister-ashwini-vaishnaw/>. See also Vaitheeswaran B, “Any Hint of Regulation May Inhibit AI Innovation: MeitY Secretary S Krishnan Says,” Times of India, March 10, 2025, <https://timesofindia.indiatimes.com/business/india-business/any-hint-of-regulation-may-inhibit-ai-innovation-meity-secretary-s-krishnan-says/articleshow/118649374.cms>.

³⁵ “Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025,17-18, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf>.

³⁶ Government of India, Ministry of Electronics and Information Technology, Guidelines for Organisations Working on AI, Rajya Sabha, Unstarred Question No. 1343 (December 6, 2024), https://sansad.in/getFile/annex/266/AU1343_RxE4hh.pdf?source=pqars.

³⁷ See Reserve Bank of India (RBI), “Framework for Self-Regulatory Organisation(s) in the FinTech Sector,” (Mumbai, RBI, 2024), <https://m.rbi.org.in/Scripts/PublicationReportDetails.aspx?UrlPage=&ID=1263&s=08> and RBI voluntary guidelines, speeches by RBI governor and deputy and new committee set up under RBI for AI guidelines ‘FREE-AI’ .”RBI Governor Das advocates for fintech self-regulation to ensure balanced innovation with integrity,” The Economic Times. August 29, 2024, https://economictimes.indiatimes.com/news/economy/finance/new-trinity-of-jam-upi-uli-will-be-a-revolutionary-step-forward-in-indias-digital-infrastructure-journey-rbi-guv-das/articleshow/112867702.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst and “Self-Regulatory Organisation Needs to Keep Fintech Sector in Check: RBI Dy Guv T. Rabi Sankar.” The Economic Times. August 29, 2024 <https://economictimes.indiatimes.com/news/economy/policy/self-regulatory-organisation-needs-to-keep-fintech-sector-in-check->

rbi-dy-guv-t-rabi-sankar/articleshow/112901777.cms?from=mdr. See also Saloni Shukla, “RBI Announces Free AI Committee to Develop AI Framework,” Economic Times, March 10, 2025, <https://economictimes.indiatimes.com/news/economy/policy/rbi-announces-free-ai-committee-to-develop-ai-framework/articleshow/116684195.cms?from=mdr>.

³⁸ “Ethical guidelines for application of Artificial Intelligence in Biomedical Research and Healthcare,” Indian Council for Medical Research, 2023, <https://www.icmr.gov.in/ethical-guidelines-for-application-of-artificial-intelligence-in-biomedical-research-and-healthcare>.

³⁹ “IndiaAI Mission Centre Selects 8 Projects for Driving Ethical AI Development,” Moneycontrol, October 16, 2024, <https://www.moneycontrol.com/technology/indiaai-mission-centre-selects-8-projects-for-driving-ethical-ai-development-article-12843615.html>. See also “MeitY and UNESCO Host Stakeholder Consultation on AI Readiness Assessment Methodology (RAM) in India,” IndiaAI, January 21, 2025, <https://indiaai.gov.in/article/meity-and-unesco-host-stakeholder-consultation-on-ai-readiness-assessment-methodology-ram-in-india> and India Takes the Lead: Establishing the IndiaAI Safety Institute for Responsible AI Innovation,” IndiaAI, January 31, 2025, <https://indiaai.gov.in/article/india-takes-the-lead-establishing-the-indiaai-safety-institute-for-responsible-ai-innovation>.

⁴⁰ Government of India, Ministry of Electronics and Information Technology, Cyber Law and Data Governance Group, Due Diligence by Intermediaries/Platforms under the Information Technology Act, 2000, and Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, (Issued on March 15, 2024), <https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf>

⁴¹ Interview with two senior government officials (Interviews 4, 12).

⁴² Interview with a Partner at an Indian technology law and policy firm and Executive at a global industry association (Interviews 8,9) . See also Jatin Grover, “Technology Will Not Rush in Bringing Digital India Act: MeitY Secretary.” Financial Express, January 07, 2025, <https://www.financialexpress.com/life/technology-will-not-rush-in-bringing-digital-india-act-meity-secretary-3708673/> and “Govt Open to New Law to Regulate AI, Minister Tells LS.” The Times of India, December 13, 2024. <https://timesofindia.indiatimes.com/india/govt-open-to-new-law-to-regulate-ai-minister-tells-ls/articleshow/116265688.cms>.

⁴³ Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>

⁴⁴ Interview with an executive at a global industry association (Interview number 8).

⁴⁵ Interview with an executive at a leading Indian industry association (Interview number 5).

⁴⁶ Interview with an executive at a global technology company (Interview number 7).

⁴⁷ “Yet, in practice, say critics, self-regulation often fails to fulfill its theoretical promise, more commonly serving the industry rather than the public interest” See Gunningham, Neil, and Joseph Rees. “Industry Self-Regulation: An Institutional Perspective.” *Law & Policy* 19, no. 4 (1997): 363–414. <https://doi.org/10.1111/1467-9930.t01-1-00033>.

⁴⁸ Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>. See also Shravani Nag Lanka and Naman Kumar, “IFF’s Submissions on the MeitY Report on AI Governance Guidelines” Internet Freedom Foundation, March 03, 2025, <https://internetfreedom.in/iffs->

submissions-on-the-meity-report-on-ai-governance-guidelines/ and Jyoti Panday and Saumya Jain, “India’s Report on AI Governance Guidelines Development,” Internet Governance Project, February 24, 2025, <https://www.internetgovernance.org/2025/02/24/indias-report-on-ai-governance-guidelines-development/> and “AI Now Joins Civil Society Groups In Statement Calling For Regulation To Protect the Public” AI Now Institute, November 1, 2023, <https://ainowinstitute.org/general/ai-now-joins-civil-society-groups-in-statement-calling-for-regulation-to-protect-the-public>

⁴⁹ Michael Akinwumi and Dominique Harrison, “Responsible AI development in the shadow of profit is a balancing act, these experts say,” Technical.ly, January 24, 2024, <https://technical.ly/software-development/ai-ethics-profit-conflict-solution/>. See also Alyssa Wong, “Tech Self-Regulation and Democratic Oversight,” Schwartz Reisman Institute for Technology and Society, University of Toronto, September 21, 2023, <https://srinstitute.utoronto.ca/news/tech-self-regulation-democratic-oversight> and Stephanie Weeker, “Hot Take: Industry Self-Regulation for AI Is a Good Intention That Just Won’t Work,” Medium, January 29, 2025, <https://scweeker.medium.com/hot-take-industry-self-regulation-for-ai-is-a-good-intention-that-just-wont-work-0579e449b05d>.

⁵⁰ Interview with a partner at a policy consulting firm, See also Shravani Nag Lanka and Naman Kumar, “IFF’s Submissions on the MeitY Report on AI Governance Guidelines” Internet Freedom Foundation, March 03, 2025, <https://internetfreedom.in/iffs-submissions-on-the-meity-report-on-ai-governance-guidelines/>

⁵¹ “Where commitments mention particular models, they apply only to generative models that are overall more powerful than the current industry frontier (e.g. models that are overall more powerful than any currently released models, including GPT-4, Claude 2, PaLM 2, Titan and, in the case of image generation, DALL-E 2.)” See: <https://bidenwhitehouse.archives.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf>

⁵² Interview with an executive at a leading Indian industry association (Interview number 5).

⁵³ “Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025, 17-18, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf>.

⁵⁴ Interview with an executive at an Indian policy consulting firm (Interview Number 2)

⁵⁵ Interview with a senior government official (Interview number 12). See also “Panel: Intelligent Regulation of AI: The Way Ahead,” Youtube video posted by “Carnegie India,” December 11, 2023, Accessed March 11, 2025, <https://www.youtube.com/watch?v=5HCyWC7gjh0>

⁵⁶ Ganguly, Nabodita. “India AI Dataset Platform AI Kosh Will Help Remove Foreign Bias, Says MeitY Secretary S Krishnan.” Outlook Business, March 6, 2025. See also “Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf> and Bharat S Raj, “Navigating Diversity: Generative AI’s Odyssey in India’s Unique Landscape,” August 21, 2023, <https://indiaai.gov.in/article/navigating-diversity-generative-ai-s-odyssey-in-india-s-unique-landscape> and “Racist, Sexist, Casteist: Is AI Bad News for India?” The Economic Times, September 11, 2023. <https://economictimes.indiatimes.com/tech/technology/racist-sexist-casteist-is-ai-bad-news-for-india/articleshow/103567192.cms>.

⁵⁷ Interview with a senior government official (Interview number 12).

⁵⁸ Interviews with three executives at global technology companies and an executive at an Indian Policy

consulting firm. (Interviews 1, 2, 7, 11).

⁵⁹ Interview with an executive at a global technology company (Interview number 11).

⁶⁰ Interviews with three executives at global technology companies and an executive at an Indian Policy consulting firm. (Interviews 1, 2, 7, 11).

⁶¹ Interview with a partner at an Indian technology law and policy firm (Interview number 10)

⁶² Interview with an executive at a global technology company (Interview number 11).

⁶³ Davis, Christina L., and Meredith Wilf. "Joining the Club: Accession to the GATT/WTO." *The Journal of Politics* 79, no. 3 (2017): 964–978. <https://doi.org/10.1086/691058>.

⁶⁴ The US Administration has personally negotiated with big-tech companies on an individual basis to 'secure' voluntary commitments from them. See "Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI," The White House, July 21, 2023, <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/>. See also, White House, Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Eight Additional Artificial Intelligence Companies to Manage the Risks Posed by AI, September 12, 2023, <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2023/09/12/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-eight-additional-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> and White House, Fact Sheet: Biden-Harris Administration Announces New AI Actions and Receives Additional Major Voluntary Commitment on AI, July 26, 2024, <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/07/26/fact-sheet-biden-harris-administration-announces-new-ai-actions-and-receives-additional-major-voluntary-commitment-on-ai/>.

⁶⁵ Interview with an executive at a global industry association. (Interview number 8).

⁶⁶ Interview with an executive at a global technology company (Interview number 11).

⁶⁷ Interview with a partner at an Indian technology law firm (Interview number 10).

⁶⁸ See "Report on AI Governance Guidelines Development," Ministry of Electronics and Information Technology, Government of India, January 2025, 16-17, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf> and Amlan Mohanty and Shatakratu Sahu, "India's Advance on AI Regulation," Carnegie India, November 21, 2024, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>.

⁶⁹ "The AI Risk Repository," MIT AI Risk Repository, <https://airisk.mit.edu/>.

⁷⁰ "International Scientific Report on the Safety of Advanced AI: Interim Report," Department for Science, Innovation and Technology, and AI Safety Institute, May 17, 2024, <https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai>.

⁷¹ "AI Risk Management Framework," National Institute for Standards and Technology, July 2024, <https://www.nist.gov/itl/ai-risk-management-framework>.

⁷² Tim McGarr, "ISO/IEC 23894 – A new standard for risk management of AI," AI Standards Hub, <https://aistandardshub.org/a-new-standard-for-ai-risk-management>.

⁷³ “Fairness Assessment and Rating of Artificial Intelligence Systems,” Telecommunication Engineering Centre, Ministry of Communications, Government of India, July 2023, https://tec.gov.in/pdf/SDs/TEC%20Standard%20for%20fairness%20assessment%20and%20rating%20of%20AI%20systems%20Final%20v5%202023_07_04.pdf.

⁷⁴ “Risk Identification and Assessment Tool,” IndiaAI, <https://indiaai.gov.in/responsible-ai/governance-framework>.

⁷⁵ Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, 11-13, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en¢er=india>.

⁷⁶ Yoshua Bengio, et al., International AI Safety Report”, available at: <https://www.gov.uk/government/publications/international-ai-safety-report-2025/international-ai-safety-report-2025>

⁷⁷ Nick Bostrom, Superintelligence: Paths, Dangers, Strategies, (Oxford University Press, July 2014).

⁷⁸ Interview with an executive at a global technology company (Interview number 11).

⁷⁹ See “IT Ministry Finalising Voluntary Code of Conduct, Ethics for AI Companies,” Business Standard, November 18, 2024, https://www.business-standard.com/technology/tech-news/it-ministry-finalising-voluntary-code-of-conduct-ethics-for-ai-companies-124111800929_1.html. See also National Strategy for Artificial Intelligence #AIForAll, report, (NITI Aayog, Government of India, June 2018), <https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf> and Responsible AI #AIForAll, Approach Document for India: Part 1 – Principles for Responsible AI, report, (NITI Aayog, Government of India, February 2021), <https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf> and Responsible AI #AIForAll, Approach Document for India: Part 2 – Operationalizing Principles for Responsible AI, report, (NITI Aayog, Government of India, August 2021), <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>. See also “Model Artificial Intelligence Governance Framework Second Edition,” InfoComm Media Development Authority and Personal Data Protection Commission Singapore, January 21, 2020, <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf> and “The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023,” Department for Science, Innovation & Technology, Government of United Kingdom, November 1, 2023, <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>.

⁸⁰ See “Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI” (The White House, July 21, 2023), <https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/> and “The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023,” Department for Science, Innovation & Technology, Government of United Kingdom, November 1, 2023, <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>.

⁸¹ See “Asilomar AI Principles,” Future of Life Institute, August 11, 2017, <https://futureoflife.org/open-letter/ai-principles/> and “Thorn and All Tech Is Human Forge Generative AI Principles with AI Leaders to Enact Strong Child Safety Commitments,” Thorn, July 16, 2024, <https://www.thorn.org/blog/generative-ai-principles/>.

⁸² Interview with an executive at a global technology company and an executive at an Indian policy consulting

firm (Interviews 2, 11).

⁸³ Interview with a partner at an Indian technology law firm (Interview number 10).

⁸⁴ “The Developer’s Playbook for Responsible Ai in India,” NASSCOM, November 2024, <https://nasscom.in/ai/pdf/the-developer%27s-playbook-for-responsible-ai-in-india.pdf>

⁸⁵ Interview with an executive at a global technology company (Interview number 7).

⁸⁶ Fred Krimmelbein, “The Role of Regulatory Bodies in AI Governance and Oversight,” Sogeti Labs, October 24, 2024, <https://labs.sogeti.com/the-role-of-regulatory-bodies-in-ai-governance-and-oversight/>.

⁸⁷ Interview with an executive at a global technology company (Interview number 1).

⁸⁸ Interview with an executive at a global technology company (Interview number 7).

⁸⁹ Interview with a leader in a community of technologists and policymakers (Interview number 3)

⁹⁰ Interview with a partner at an Indian technology law and policy firm (Interview number 9)

⁹¹ Interviews with a partner at an Indian technology law and policy firm, and a leader in a community of technologists and policymakers (Interviews 3, 9).

⁹² Reserve Bank of India (RBI), “Framework for Self-Regulatory Organisation(s) in the FinTech Sector,” (Mumbai, RBI, 2024), <https://m.rbi.org.in/Scripts/PublicationReportDetails.aspx?UrlPage=&ID=1263&s=08>

⁹³ The Code for Self-Regulation of Advertising Content in India, (Advertising Standards Council of India, June 2022 edition), https://www.ascionline.in/wp-content/uploads/2022/11/asci_code_of_self_regulation.pdf.

⁹⁴ “Data Empowerment and Protection Architecture: A Secure Consent-Based Approach” (NITI Aayog, Government of India, August 2020), <https://www.niti.gov.in/sites/default/files/2023-03/Data-Empowerment-and-Protection-Architecture-A-Secure-Consent-Based.pdf>.

⁹⁵ “Emphasising the need for AI-related safety, the Minister announced the establishment of an AI Safety Institute. It will be based on the hub-and-spoke model. It will enable the researchers, developers and other stakeholders to develop tools, frameworks and processes for AI safety. In this regard, 8 projects have been undertaken under Indigenous Governance of AI” See “India Takes the Lead: Establishing the IndiaAI Safety Institute for Responsible AI Innovation,” IndiaAI, January 31, 2025, <https://indiaai.gov.in/article/india-takes-the-lead-establishing-the-indiaai-safety-institute-for-responsible-ai-innovation>

⁹⁶ Amlan Mohanty and Tejas Bharadwaj, “The Importance of AI Safety Institutes,” Carnegie India, June 28, 2024, <https://carnegieendowment.org/posts/2024/06/the-importance-of-ai-safety-institutes?lang=en>

⁹⁷ Sidarth Deb, “What India’s AI Safety Institute could do” The Hindu, December 02, 2024, <https://www.thehindu.com/opinion/op-ed/what-indias-ai-safety-institute-could-do/article68935356.ece>

⁹⁸ Aditi Agarwal, “Govt mulls setting up Artificial Intelligence Safety Institute,” Hindustan Times, October 13, 2024, <https://www.hindustantimes.com/india-news/govt-mulls-setting-up-artificial-intelligence-safety-institute-101728833433153.html>

⁹⁹ Amlan Mohanty and Shatakratu Sahu, “India’s Advance on AI Regulation,” Carnegie India, November 21, 2024, 21-22, <https://carnegieendowment.org/research/2024/11/indias-advance-on-ai-regulation?lang=en&>

center=india.

¹⁰⁰ “Report on AI Governance Guidelines Development,” Ministry of Electronics and Information Technology, Government of India, January 2025, 15-16, <https://indiaai.s3.ap-south-1.amazonaws.com/docs/subcommittee-report-dec26.pdf>.

¹⁰¹ Responsible AI #AIForAll, Approach Document for India: Part 2 – Operationalizing Principles for Responsible AI, report, (NITI Aayog, Government of India, August 2021), <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>.

¹⁰² Reserve Bank of India (RBI), “Framework for Self-Regulatory Organisation(s) in the FinTech Sector,” (Mumbai, RBI, 2024), <https://m.rbi.org.in/Scripts/PublicationReportDetails.aspx?UrlPage=&ID=1263&s=08>

¹⁰³ Das, Shouvik. “Amid Lack of Regulation, Online Gaming in Limbo, Failing to Attract Investment.” Mint, December 4, 2024. <https://www.livemint.com/industry/regulation-online-gaming-meity-it-ministry-gambling-real-money-11733228242999.html>. See also Jha, Lata. “Streaming Platforms Battle Fresh Concerns of Government Control, Regulations.” Mint, September 4, 2024. <https://www.livemint.com/industry/media/ic-814-netflix-ott-platforms-government-control-regulations-it-rules-film-series-cinema-11726987986151.html>.

¹⁰⁴ JE News Desk, “MCA Meets MoS IT Rajeev Chandrasekhar, Proposes Industry-Backed Self Regulatory Body For Indian Fact Checkers” English Jagran, April 06, 2023, <https://english.jagran.com/india/mca-meets-mos-it-rajeev-chandrasekhar-proposes-industrybacked-self-regulatory-body-for-indian-fact-checkers-10072432>.

¹⁰⁵ Interview with an executive at an Indian policy consulting firm (Interview Number 2).

CeRAI at IITM, a premier multidisciplinary, non-profit research centre positioned in the Global Majority, is one among the few global institutions that specializes in both technical and policy research to ensure and enable responsible development and deployment of AI systems.

 cerai.iitm.ac.in

 5th floor, Block II, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology, Madras, Chennai-600036, India.

