Introduction

On 5 November 2025, the Ministry of Electronics and Information Technology (MeitY), issued the AI Governance Guidelines (Guidelines) setting out India's framework for safe and trusted AI. The stated objective of the Guidelines is to balance pro-innovation adoption of AI with measures to mitigate risks to individuals, communities and national security. The Guidelines are sector agnostic and outline the core principles, key recommendations, an action plan and practical guidance for industry actors and regulators to support compliance, transparency and accountability.

Background

The Guidelines have emerged from a multi-year policy effort led by the Government of India. In 2023, a high-level advisory group set up a sub-committee on AI governance, which prepared a draft for public consultation. MeitY published that draft in early 2025 and received more than 2,500 submissions from stakeholders. To finalise the framework, MeitY constituted a drafting committee (Committee) in July 2025 with representation from the government, academia and industry, resulting in framing of the Guidelines.

Proposed Liability & Accountability Framework

➢ ***Adaptation of existing laws to cover AI regulation***

The Committee has concluded that a substantial portion of the risks associated with artificial intelligence can be effectively managed within the framework of existing legislation. It recommends undertaking a systematic analysis to identify any legal gaps and, where necessary, undertake targeted amendments to ensure regulatory coverage.

- *Information Technology Act, 2000 (IT Act)*

  The Committee has recommended that there is a need to amend the legislation to effectively address novel concerns thrown up by the deployment and use of AI systems such as classification of digital entities, apportionment of liability and extension of safe harbour protection to AI systems.

  Classification: The IT Act defines an intermediary to mean an entity that, *"on behalf of another person receives, stores or transmits (an electronic record) or provides any service with respect to such record"*. From telecom service providers to cyber cafes - many are considered to be intermediaries under the broad contours of the current definition. However, it is questionable whether the definition would cover AI systems such as generative AI models, which generate data basis user prompts. The Committee has recommended that clarity be provided on classification of such AI systems under the IT Act.

  Liability attribution: The Committee has recommended defining the role of various stakeholders in the AI value chain (such as developers, deployers, users etc.) and providing clarity on how they will be governed. Attribution of liability i.e. identifying who is at fault when an AI system malfunctions or for the AI generated output is complicated, owing to the autonomous nature of AI and the degree

of unpredictability involved in its functioning. For example, in an AI medical diagnosis error, it would be challenging to determine if the AI developer, the hospital which procured the AI system or the physician would be liable. Current liability frameworks are not fully equipped to handle AI's complexity and the Committee's recommendation to provide clarity in this regard is welcome.

Safe Harbour protection: Intermediaries presently enjoy legal immunity as long as they exercise due diligence in relation to user generated content. To claim safe harbour, the intermediary ought not to have initiated the transmission of data, modified it or selected its recipient. However, many modern AI systems such as generative AI chatbots modify user data or autonomously generate data. The Committee recommends providing clarity on what kind of exposure such AI systems will have under the prevalent laws.

- *Copyright*

  The Committee has observed that implications of using copyrighted material in training and development of AI models may not be covered under the Indian Copyright Act, 1857. As Department for Promotion of Industry and Internal Trade (DPIIT) has formed a committee to deliberate on this issue, the Guidelines recommend that a balanced approach be adopted by the DPIIT formed committee, which enables text and data mining while protecting the rights of copyright holders.

  The legal question concerning the intersection of AI and copyright law—specifically, whether the training of AI models on publicly available data falls within the scope of the fair use doctrine under Indian law—is currently under consideration by the Delhi High Court in a case filed by ANI Media against OpenAI. In the absence of regulatory or legislative clarity, the outcome of this litigation will have significant implications for AI developers, particularly in assessing the legal risks and potential costs associated with model development in India. To ensure certainty and to foster the growth of indigenous AI models, as envisioned in the Guidelines, it is imperative that the Government provides clear guidance on this aspect at the earliest.

- *Data Protection*

  There has been a growing concern that training and deployment of AI models may give rise to potential friction with core data protection principles imbibed in the Digital Personal Data Protection Act, 2023 (DPDPA). For example, AI may make it difficult to facilitate the exercise of data principal rights, including the right to access, correction or deletion of personal data already embedded in the AI model or adhere to the requirement of purpose limitation - as data collected by AI systems is often repurposed for secondary or evolving use cases beyond its original purposes.

  The Committee has recommended that key issues regarding how AI model development and risk mitigation will be dealt with under the DPDPA be clarified and if required legislative amendments be carried out to deal with the same.

- *Content Authentication*

  The growing misuse of AI for creation of deepfakes has flooded courts with cases and disproportionately affected the most vulnerable in society. To tackle this issue, the Committee has recommended developing global standards for content authentication and data provenance. A review of the regulatory framework in India to tackle the issue of deepfakes through techno-legal solutions, such as the use of watermarks to trace origin or dataset provenance tools to identify training data sources, have also been recommended.

  In this context, recently on 22 October 2025, MeitY published draft amendments to the *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* to deal with "*synthetically generated information*" (**SGI**) including within its ambit deepfakes, AI-generated, and algorithmically modified content. The draft amendments amongst other things, propose that intermediaries who enable or facilitate the generation or alteration of information as SGI, are required to ensure that such content is labelled or embedded with a permanent unique metadata or identifier.

➢ ***Voluntary Frameworks***

In alignment with the Government's pro-innovation stance, the Committee has recommended voluntary measures to address AI-related risks, rather than imposing compliance-heavy regulations. These voluntary measures should be proportionate to the level of risk involved. To promote widespread

adoption, the Committee has proposed enabling access to regulatory sandboxes, offering public recognition, and providing financial and technical incentives.

- ***Grievance Redressal***

  The Committee has recommended that organizations deploying AI systems establish accessible grievance redressal mechanisms in order to ensure timely resolution of issues, thereby mitigating AI-related risks and harms at the earliest possible stage. Experience demonstrates that mandated grievance mechanisms at the organizational level have often led to positive outcomes, enhancing user experience and reducing the likelihood of unnecessary disputes. For instance, the grievance redressal framework required for social media intermediaries under the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 has enabled the timely resolution of a significant number of content moderation issues directly at the platform level, thereby minimizing the need for external escalation.

Institutions

The Committee has recommended that sectoral regulators take the lead in providing guidance and enforcing regulations in their respective domains - with MeitY, as the nodal ministry responsible for overall adoption and regulation of AI systems. The following bodies have been recommended for playing an instrumental role:

- **AIGG**: Constituting an AI Governance Group for coordination regarding cross-government policy, studying regulatory gaps and need for legal amendments, taking measures to improve accountability for compliance with local laws.

- **TPEC**: Creating a Technology & Policy Expert Committee with the primary goal of providing domain expertise to AIGG in order to enable it to perform its functions effectively.

- **AISI**: AI Safety Institute for conducting research on AI safety issues, evaluating regulatory gaps, developing draft guidelines and standards with sectoral regulators, building capacity, and anchoring participation in global networks.

Other Important Recommendations

- **Infrastructure:** The Committee's recommendations on AI infrastructure lay emphasis on not only increasing the adoption of AI but also making it more inclusive and locally relevant. Access to quality datasets relevant to the Indian context and affordable computing power have been identified as key factors required for accelerating adoption - especially in tier-2 and 3 cities, and in sectors such as agriculture, healthcare, etc. which are lagging in adoption. To achieve these goals, leveraging Digital Public Infrastructure (DPI) (such as Aadhar, UPI, etc), introducing market incentives and appropriate data governance to encourage contributions to national data platforms like AIKosh and offering targeted incentives such as tax rebates, AI linked loans and subsidised GPU access to MSMEs, have been recommended.

- **Training Programmes:** To support the safe, sustainable and responsible scaling of AI adoption, the Committee recommends implementing targeted training programs that address both the risks and opportunities associated with AI technologies. In light of the growing prevalence of AI-enabled crimes such as deepfake scams, algorithmic trading frauds, and advanced phishing attacks, the Committee further advises strengthening the capacity of law enforcement agencies to effectively investigate and respond to such threats.

- **Incident Reporting:** The Committee has proposed the establishment of a national database to record AI related incidents, aimed at informing effective policymaking and mitigating associated harms. To foster broad and candid participation, it recommends allowing organizations to report confidentially, without fear of penalties.

- **Techno-legal measures:** The Committee has emphasized the use of techno-legal instruments to effectively operationalize policy objectives. Key measures include the deployment of privacy-enhancing technologies, mechanisms for bias detection and mitigation, explainability frameworks, content provenance tools, and Data Empowerment and Protection Architecture (DEPA)-style consent architectures to support responsible AI training.

- **Human Oversight:** The Committee recommends mitigating loss of control *inter alia* by requiring human-in-the-loop safeguards at critical decision points. Where human oversight is ineffective such as in the case of high-velocity algorithmic trading, adoption of circuit breakers and system-level constraints have been recommended.

Conclusion

The AI Governance Guidelines provide much needed clarity on the approach that the Government is expected to take in relation to AI adoption and risk mitigation in India. The intent is clear that as a short and mid-term strategy - the Government is leaning towards amending existing laws, relying on sectoral regulations and adoption of voluntary frameworks as opposed to enacting new pieces of legislation for governing AI in India.

For industry, the immediate implications are to map applicable laws, embed transparency and grievance redressal processes, adopt voluntary standards, and operationalise techno-legal safeguards across the lifecycle.

For regulators, the focus is on applying existing statutes consistently, clarifying classification and liability, building incident reporting and infrastructure, and enabling innovation through sandboxes and DPI integration.

Pending amendments to the relevant legislations and issuance of necessary clarifications, it will be advantageous for businesses to try and proactively align with the principles and practical measures set out in the Guidelines. These will assist in demonstrating trustworthiness, managing compliance risks better and also in meaningfully participating in the emerging AI governance ecosystem.

- *Anushka Sharda (Partner); Harsh Walia (Partner); Supratim Chakraborty (Partner) and Shobhit Chandra (Counsel)*

About Khaitan & Co

Khaitan & Co is a top tier and full-service law firm with over 1300+ legal professionals, including 300+ leaders and presence in India and Singapore. With more than a century of experience in practicing law, we offer end-to-end legal solutions in diverse practice areas to our clients across the world. We have a team of highly motivated and dynamic professionals delivering outstanding client service and expert legal advice across a wide gamut of sectors and industries.

To know more, visit **www.khaitanco.com**