

GOVERNMENT OF INDIA  
MINISTRY OF ELECTRONICS AND INFORMATION TECHNOLOGY  
**RAJYA SABHA**  
**UNSTARRED QUESTION NO. 717**  
TO BE ANSWERED ON: 25.07.2025

**REGULATION OF AI VIS-A-VIS PROTECTION OF CITIZENS' RIGHTS**

**717. SHRI K.R.N. RAJESHKUMAR:**

Will the Minister of ELECTRONICS AND INFORMATION TECHNOLOGY be pleased to state:

- (a) the measures being taken by Government to protect citizens' rights, particularly the right to privacy, in the context of emerging technologies such as Artificial Intelligence (AI) and deepfakes;
- (b) the manner in which Government proposes to balance the need for technological innovation with the necessity of regulation and oversight;
- (c) whether Government is planning to regulate the AI Sector in India to address the unchecked proliferation of deepfake technologies and the spread of misinformation; and
- (d) the steps being taken by Government to address the spread of misinformation and disinformation through digital platforms?

**ANSWER**

MINISTER OF STATE FOR ELECTRONICS AND INFORMATION TECHNOLOGY  
(SHRI JITIN PRASADA)

(a) to (d): The policies of the Government of India are aimed at ensuring an open, safe, trusted and accountable cyberspace for users in the country amidst the growth of AI.

The following acts regulate the information generated by users or otherwise, including generated using AI tools or any other technology, and ensure individual privacy.

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

- IT Rules, 2021 framed under the act impose due diligence obligations on intermediaries to ensure users do not host, upload, or share unlawful content on their platforms
- Failure to observe this due diligence may lead to loss of 'safe harbour' under Section 79 of the IT Act, making intermediaries liable for third-party content
- Intermediaries are required to publish monthly compliance reports, mentioning the complaints received, actions taken, and content removed or disabled, including those identified through automated monitoring tools

**Digital Personal Data Protection Act, 2023 ("DPDP Act")**

- Act mandates Data Fiduciaries to adopt reasonable security safeguards and technical and organizational measures to comply with its provisions of the Act
- Comprehensive framework for providing individual rights like right to access, correction & erasure, grievance redressal, etc.

- Robust framework of accountability mechanisms with the Data Protection Board as an independent adjudicatory body
- Board is empowered to investigate complaints, conduct inquiries, and impose penalties of up to ₹250 crore for non-compliance of the provisions of the Act

### **Bhartiya Nyaya Sanhita, 2023**

Section 353 penalize the act of making false or misleading statements, rumours, or reports that can cause public mischief or fear

India's Ai strategy aims to create economic and employment opportunities for India's citizens while mitigating their negative impacts. IndiaAI mission advances this through its seven pillar strategy.

### **"Safe and Trusted AI" pillar**

- This is an important pillar of IndiaAI Mission's which focuses on ensuring ethical and responsible AI development and deployment in India.
- Aims to create frameworks, tools, and guidelines for safe and trustworthy AI systems.
- Details of the projects undertaken are at Annexure-I

### **Steps taken to check misinformation and disinformation**

- Multiple consultations between MeitY and industry stakeholders/social media platforms to address emerging harms in cyberspace resulting from misinformation and deepfakes
- Advisories issued by MeitY reminding the intermediaries about their due diligence obligations outlined under IT Rules, 2021, and on countering unlawful content including malicious 'synthetic media' and 'deepfakes'.
- Regular alerts & advisories issued by Indian Computer Emergency Response Team (CERT-In) regarding cyber threats/vulnerabilities, including malicious attacks using AI, and countermeasures to protect computers, network's and data

\*\*\*\*\*

## Annexure I

Details of projects selected against the first Expression of Interest (EoI) floated under the **Safe and Trusted AI Pillar**

<b>Name Of the Theme</b>	<b>Selected Applicant</b>	<b>Title Of the Project</b>	<b>About the Project</b>
Machine Unlearning	IIT Jodhpur	Machine Unlearning in Generative Foundation Models	To development novel method for targeted unlearning in open-source generative foundation models while minimising negative impact on overall model performance.
Synthetic Data Generation	IIT Roorkee	Design and Development of Method for Generating Synthetic Data for Mitigating Bias in Datasets; and Framework for Mitigating Bias in Machine Learning Pipeline for Responsible AI	To develop the algorithm and method for handling the bias at the model training and in-processing stage of ML model development
AI Bias Mitigation Strategy	National Institute of Technology Raipur	Development of Responsible Artificial Intelligence for Bias Mitigation in Health Care Systems	To develop responsible AI algorithms that reduce biases in medical system applications, image analysis, and diagnostic decisions
Explainable AI Framework	DIAT Pune and Mindgraph Technology Pvt. Ltd.	Enabling Explainable and Privacy Preserving AI for Security	Create AI models that provide accurate and interpretable results for human activity analysis for effective security in crowded environment.
Privacy Enhancing Strategy	IIT Delhi, IIIT Delhi, IIT Dharwad and Telecommunication	Robust Privacy-Preserving Machine Learning Models	To develop robust distributed/federated learning algorithms which perform well in an

<b>Name Of the Theme</b>	<b>Selected Applicant</b>	<b>Title Of the Project</b>	<b>About the Project</b>
	Engineering Center (TEC)		adversarial environment susceptible to attacks
AI Ethical Certification Framework	IIIT Delhi and Telecommunication Engineering Center (TEC)	Tools for assessing fairness of AI model	To develop a three-step certification process involving bias risk assessment, processing for metrics, and bias testing to ensure fairness for AI Systems in the Indian context
AI Algorithm Auditing Tool	Civic Data Labs	ParakhAI - An open-source framework and toolkit for Participatory Algorithmic Auditing	The proposed framework and toolkit will enable involving citizens in the responsible design, development, and deployment of algorithmic decision-making systems.
AI Governance Testing Framework	Amrita Vishwa Vidyapeetham and Telecommunication Engineering Center (TEC)	Track-LLM, Transparency, Risk Assessment, Context & Knowledge for Large Language Models	To identify and address the specific gaps in the existing governance testing frameworks related to LLM;s downstream use-case and deployment

\*\*\*\*\*

