GOVERNMENT OF INDIA
MINISTRY OF ELECTRONICS AND INFORMATION TECHNOLOGY
LOK SABHA
UNSTARRED QUESTION. NO. 3091
TO BE ANSWERED ON: 19.03.2025

THREATS OF DEEPFAKES

3091.SMT. KANIMOZHI KARUNANIDHI:

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

(a) whether the Government is aware of the survey conducted by the cybersecurity company McAfee which highlights that 75 percent of Indians have viewed some form of deepfake content and at least 38 percent have been targeted by a deepfake scam;
(b) whether the Government has taken any measures to address the issue of deepfake content and scams, if so, the details thereof and if not, the reasons therefor;
(c) whether the Government has initiated any awareness campaigns or educational programmes to inform the public about identifying and safeguarding against deepfake content and scams, if so, the details thereof and if not, the reasons therefor; and
(d) whether the Government is planning to introduce legislative frameworks or regulatory measures to mitigate the risks associated with deepfakes and if so, the details thereof?

ANSWER

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

(a) to(d): The Government of India emphasizes the concept of 'AI for All,' aligning with the Hon'ble Prime Minister's vision to democratise use of technology. This initiative aims to ensure that AI benefits all sectors of society, driving innovation and growth.

Government is committed to harnessing the power of Artificial Intelligence (AI) for the good of our people in sectors like healthcare, agriculture and education. At the same time, the **Government is cognizant of the risks posed by AI and the need to create guardrails to ensure that AI is safe and trusted**.The Central Government after extensive public consultations with relevant stakeholders has notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules, 2021") on 25.02.2021 which was subsequently amended 28.10.2022 and 6.4.2023. Government engages with and receives inputs from the public and stakeholders, including in respect of changes required to existing legislation and the need to introduce fresh legislation. It has notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules,2021") on 25.02.2021 which were subsequently amended 28.10.2022 and 6.4.2023.

Cyber-security risks arising out of creation, distribution, and dissemination of deepfakes and the technologies that enable creation of deepfakes that may pose risks at an individual level or a national level are currently being addressed by provisions under the Information Technology Act, 2000 ("IT Act") and the rules therein such as the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules, 2021") which do not distinguish between any information that is synthetically generated using AI tools such as deepfakes and those which are authentic/original information. The IT Act and the IT Rules, 2021 collectively provide a framework for removal of harmful content and ensures that users over the internet are safeguarded from its perils which includes content that is artificially generated and patently false.

Further, taking into cognizance the urgency to address the harms and criminalities arising out of widespread circulation of misinformation and deepfakes powered by AI, MeitY has previously conducted multiple consultations with industry stakeholders/ social media platforms to discuss the challenges identified in combating deepfakes. **MeitY has issued advisories dated 26.12.2023, 15.03.2024 and 03.09.2024 (collectively as "Advisories"), through which the intermediaries were reminded about compliance** with their due-diligence obligations outlined under the Rule 3(1)(b) of the IT Rules, 2021 and advised on countering unlawful content including malicious 'synthetic media' and 'deepfakes' for better **compliance to curb deepfakes and promptly remove harmful content online**. The overall objective of these advisory is to empower users with adequate information so that they make informed decisions while using synthetic media tools. These Advisories among other things advise every intermediary and platform to ensure that its computer resource in itself or through the use of AI model(s)/ LLM/ Generative AI, software(s) or algorithm(s) does not permit any bias or discrimination or threaten the integrity of the electoral process.

Government has constituted an **Advisory Group on AI for India-specific regulatory AI framework** under the chairmanship of Principal Scientific Advisor (PSA) to Hon'ble Prime Minister of India with diverse stakeholders from academia, industry and government with an objective to address all issues related to development of Responsible AI framework for safe and trusted development and deployment of AI.

Government has constituted a **committee on the matters related to the issue of deepfakes**with diverse stakeholders from academia, Industry and government with an objective to address issues related deepfakes.

Government has funded two projects namely **"Fake Speech Detection Using Deep Learning Framework"** and **"Design and Development of Software for Detecting Deepfake Videos and Images"**. In the project "Design and Development of Software for Detecting Deepfake Videos and Images" a prototype tool capable of detecting deepfakes, named **FakeCheck**, has been developed to detect deepfakes without the use of the internet. The tool has been shared with select law enforcement agencies for testing and to obtain feedback for further refinement

Under the safe and trusted pillar of IndiaAI Mission, **Expressions of Interest**have been invited to address the need for robust guardrails to ensure the responsible development, deployment, and adoption of AI technologies. These covers a range of critical themes in areas such as Watermarking & Labelling, Ethical AI Frameworks, AI Risk Assessment & Management, Stress Testing Tools, and**Deepfake Detection Tools**.

With the innovation and increase in digital technologies for delivery of services to users, fraudsters are adopting social engineering techniques including deepfakes to trick users and conducting fraudulent activities. The Indian Computer Emergency Response Team (CERT-In) issues alerts and advisories regarding latest cyber threats/vulnerabilities including malicious attacks using Artificial Intelligence and countermeasures to protect computers, networks and data on an ongoing basis. In this context, an **advisory on safety measures to be taken to minimize the adversarial threats** arising from Artificial Intelligence (AI) based applications was published in May 2023. CERT-In has published an **advisory in November 2024 on deepfake threats and measures** that need to be followed to stay protected against deepfakes.
CERT-In has taken following **measures to enhance awareness among users and organisations** for safe usage of digital technologies and tackling digital risks:

i) CERT-In issues alerts and advisories regarding latest cyber threats/vulnerabilities including social engineering, phishing and vishing campaigns and countermeasures to protect computers, mobile phones, networks and data on an ongoing basis.
ii) CERT-In has issued an advisory to various Ministries in November 2023 outlining the measures to be taken for strengthening the cyber security by all entities that are processing the digital personal data or information including sensitive personal data or information.

iii) CERT-In operates the Cyber Swachhta Kendra (Botnet Cleaning and Malware Analysis Centre) to detect malicious programs and provides free tools to remove the same, and also provides cyber security tips and best practices for citizens and organisations.

iv) CERT-In provides leadership for the Computer Security Incident Response Team-Finance Sector (CSIRT-Fin) operations under its umbrella for responding to and containing and mitigating cyber security incidents reported from the financial sector.

v) Security tips have been published for users to secure their desktops and mobile phones and to prevent phishing attacks.

vi) CERT-In is regularly carrying out various activities for awareness and citizen sensitization with respect to cyber-attacks and cyber frauds. CERT-In is observing the Cyber Security Awareness Month (NCSAM) during October of every year, Safer Internet Day on 1st Week Tuesday of February Month every year, Swachhta Pakhwada from 1 to 15 February of every year and Cyber JagrooktaDiwas (CJD) on 1st Wednesday of every month by organising various events and activities for citizens as well as the technical cyber community in India. CERT-In conducted several awareness activities such as Quiz, webinars, Capture the Flag event in collaboration with Government and industry partners during NCSAM 2024 with the theme "SatarkNagrik, Secure our World".

****