KHAITAN & CO ADVOCATES SINCE 1911

UPDATE

ERGO

Analysing developments impacting business

GROWING CONCERN OVER DEEPFAKES – CERT-IN ISSUES ADVISORY

14 December 2024

On 27 November 2024, the Indian Computer Emergency Response Team (CERT-In) issued an advisory on Deepfake Threats and Countermeasures (**Advisory**), highlighting the escalating risks posed by the increasing sophistication of deepfake technology. This advisory serves as a critical resource for individuals, organizations, and businesses to identify, assess, and mitigate the threats arising from synthetic media.

What are Deepfakes?

According to the Advisory, deepfakes are a form of synthetic media created using artificial intelligence (AI) to generate or manipulate realistic videos, images and audio. These AI-generated media can convincingly alter or fabricate content to appear authentic.

Snapshot of the Advisory

The Advisory highlights the threats that deepfake technology poses such as financial fraud, where scammers impersonate company executives or family members to trick victims into transferring funds or sharing sensitive information; disinformation, which involves the creation of fake videos or audio to manipulate public opinion and erode trust in media and institutions, etc. The Advisory provides essential preventive measures for individuals such as verification of sources before sharing or acting on digital content, cross-reference information with reliable sources, limiting the sharing of high-resolution personal media, enabling privacy settings, and securing accounts with Multi-Factor Authentication (**MFA**).

Proposed Mitigation Strategies for Organizations

- *Watermarking Media*. Embed digital watermarks in official content to deter exploitation.

- *Verification Protocols*. Implement robust verification methods, such as MFA for all digital communications and callback procedures for sensitive transactions.

- *Advanced Detection Tools*. Invest in AI-driven detection tools to identify deepfake content and ensure these tools are regularly updated.

- *Media Provenance Verification*. Utilize provenance tools to trace the origins of media or perform reverse image searches to verify authenticity.

- *Enhanced Digital Forensics*. Enhance cybersecurity teams with cutting-edge tools and forensics training to swiftly respond to deepfake incidents.

- *Monitor Public Channels*. Regularly monitor social media and public platforms for potential deepfake content targeting organizations.

- *Crisis Management Plan*. Develop protocols to respond effectively to deepfake incidents to minimize damage.

- *Legal Framework*. Ensure compliance policies and legal mechanisms address deepfake threats and provide avenues for recourse.

- *Secure Communication*. Adopt encrypted channels for sensitive discussions to prevent manipulation or interception.

- *Regular Security Audits*. Conduct audits to identify vulnerabilities and ensure preparedness against emerging deepfake threats.

Comments

The proliferation of deepfake technology poses a significant risk as a vast digital user base increasingly relies on digital communication for business, governance, and personal use. This risk has been recognized by authorities, as evident from recent advisories issued by the Ministry of Electronics and Information Technology (MeitY).

Previously, MeitY has highlighted the need for organizations and government bodies to exercise due diligence and take expeditious action against deepfakes.

While India does not yet have specific legislation to address deepfakes, existing provisions under the Information Technology Act 2000 concerning impersonation, identity theft, and fraud provide a partial legal framework. Additionally, the Advisory, though not binding, offers guidance on precautionary measures such as watermarking, developing media provenance, developing protocols for incident response, etc. The forthcoming Digital Personal Data Protection Act 2023 could potentially address deepfakes in some form, though the extent remains to be seen.

Globally, countries are implementing targeted legal measures to combat the deepfake threat. In the United States, legislators have proposed the No Artificial Intelligence Fake Replicas and Unauthorized Duplications (No AI FRAUD) Act, which aims to prevent the unauthorized creation and use of AI-generated replicas. The European Union's AI Act provides a comprehensive framework for AI regulation, including mandatory transparency obligations for AI-generated content. Similarly, China has introduced regulations mandating that AI-generated content be clearly labelled to distinguish it from authentic content.

Certain measures outlined in the Advisory, such as watermarking and the establishment of incident response protocols, appear to be influenced by such international laws. As deepfake technology continues to evolve, adopting best practices from global regulatory models will be crucial to mitigating risks.

- *Harsh Walia (Partner); Shobhit Chandra (Counsel) and Vanshika Lal (Associate)*

For any queries please contact: editors@khaitanco.com