

## Analysing the Impact of Advisory on Deepfakes *User Engagement, Effectiveness, and Future Directions*

### Background and Context

Deepfakes are characterised by the non-consensual generation of artificial intelligence (AI)-driven voices, images, or videos. These synthetic media creations are primarily used for three harmful purposes: producing sexual imagery, committing fraud, and spreading misinformation. The misuse of deepfake technology has increased alarmingly in recent years, impacting individual privacy, financial security, and the integrity of democratic processes. Between 2022 and 2023, the threat landscape saw a staggering 400 percent increase in deepfake sexual content<sup>1</sup> and an unprecedented 3000 percent rise in deepfake-enabled fraud<sup>2</sup>. In India, a survey revealed that over 75 percent of India's online population has encountered deepfake content, with at least 38 percent falling victim to related scams.<sup>3</sup>

MeitY<sup>4</sup> has issued advisories and proposed amendments to the Information Technology Rules, 2021, aimed at combating misinformation and deepfakes. These efforts emphasise strict adherence to Rule 3(1)(b) of the IT Rules, which requires digital platforms to prevent the spread of prohibited content. The advisory suggests using 'consent popups' to inform users about possible fallibilities in AI-generated content and emphasises clear communication about the consequences of dealing with unlawful information on platforms.<sup>5</sup> While the advisory mentions labelling synthetically generated content through metadata, it appears to shift some responsibility for identifying and curbing misinformation onto users. Nevertheless, the rapidly evolving nature of deepfake technology suggests that these broad regulatory measures may be inadequate. While the advisory is a step forward, there is a need for more targeted and technologically informed approaches to effectively address the challenges posed by deepfake.

### Objectives

- Assess the impact of regulatory interventions aimed curbing the spread of deepfake related misinformation.
- Identify alternative mechanisms including global good practices for detecting and mitigating deepfake-generated content without overburdening users.
- Propose balanced regulatory approaches that address deepfake concerns while safeguarding freedom of expression.

### Methodology

- **Literature Review**  
A thorough review of existing research on AI/ML-generated deepfakes will be conducted including policy documents, news reports, and journals, focusing on creation and dissemination methods, potential solutions, and regulatory challenges.
- **Stakeholder Consultations**  
Consultations will include diverse stakeholders such as intermediary platforms, fact-checkers, digital rights groups, consumer advocates, academics, and cybersecurity experts.
- **Documentation**  
Findings from the literature review and consultations will be compiled, analysed, and synthesised into a comprehensive report with key insights and recommendations.

## Outputs

- A policy brief will be developed, based on secondary research and stakeholder consultation, to address strategies for combating online misinformation, including detection technologies and the potential drawbacks of imposing strict responsibilities on users and platforms
- Op-eds will summarise findings and present arguments based on secondary research and stakeholder consultations.

## Envisaged Outcome

- Provide evidence-based insights on the effectiveness and drawbacks of current regulatory efforts in combating deepfakes and related misinformation.
- Identify alternative mechanisms for detecting and mitigating deepfake content without overly burdening users or platforms.

---

## Endnotes

- <sup>1</sup> <https://www.securityhero.io/state-of-deepfakes/#key-findings>
- <sup>2</sup> <https://onfido.com/landing/identity-fraud-report/>
- <sup>3</sup> <https://economictimes.indiatimes.com/tech/technology/75-indians-have-viewed-some-deepfake-content-in-last-12-months-says-mcafee-survey/articleshow/109599811.cms?from=mdr>
- <sup>4</sup> Ministry of Electronics and Information Technology
- <sup>5</sup> <https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf>;  
<https://pib.gov.in/PressReleaseframePage.aspx?PRID=1990542>

