GS World
Committed To Excellence
exclusive Editorial Simplified
08 December 2023

REGULATING DEEPFAKES AND AL IN INDIA

The Hindu
Paper - III (Science & Technology and Internal Security)

Last month a video featuring actor Rashmika Mandanna went viral on social media, sparking a combination of shock and horror among netizens. The seconds-long clip, which featured Mandanna's likeness, was manipulated using deepfake technology. Deepfakes are digital media, video, audio, and images, edited and manipulated using Artificial Intelligence (AI). Since they incorporate hyper-realistic digital falsification, they can potentially be used to damage reputations and undermine trust in democratic institutions. This phenomenon has forayed into political messaging as well, a concern in the run-up to the general elections next year.

Have deepfakes been used in politics?

Back in 2020, in the first-ever use of AI-generated deepfakes in political campaigns, a series of videos of Bharatiya Janata Party (BJP) leader Manoj Tiwari were circulated on multiple WhatsApp groups. The videos showed Mr. Tiwari hurling allegations against his political opponent Arvind Kejriwal in English and Haryanvi, before Delhi elections. In a similar incident, a doctored video of Madhya Pradesh Congress chief Kamal Nath recently went viral, creating confusion over the future of the State government's Laadli Behna Scheme.

Other countries are also grappling with the dangerous consequences of rapidly evolving AI technology. In May last year, a deepfake of Ukrainian President Volodymyr Zelenskyy asking his countrymen to lay down their weapons went viral after cybercriminals hacked into a Ukrainian television channel.

How did deepfake tech emerge?

Deepfakes are made using technologies such as AI and machine learning, blurring the lines between fiction and reality. Although they have benefits in education, film production, criminal forensics, and artistic expression, they can also be used to exploit people, sabotage elections and spread large-scale misinformation. While editing tools, like Photoshop, have been in use for decades, the first-ever use of deepfake technology can reportedly be traced back to a Reddit user who in 2017 had used a publicly available AI-driven software to create pornographic content by imposing the faces of celebrities on to the bodies of ordinary people.

Now, deepfakes can easily be generated by semi-skilled and unskilled individuals by morphing audio-visual clips and images. As such technology becomes harder to detect, more resources are now accessible to equip individuals against their misuse. For instance, the Massachusetts Institute of Technology (MIT) created a Detect Fakes website to help people identify deepfakes by focusing on small intricate details. The use of deepfakes to perpetrate online gendered violence has also been a rising concern. A 2019 study conducted by AI firm Deeptrace found that a staggering 96% of deepfakes were pornographic, and 99% of them involved women. Highlighting how deepfakes are being weaponised against women, Apar Gupta, founding director of Internet Freedom Foundation (IFF) says, "Romantic partners utilise deepfake technology to shame women who have spurned their advances causing them psychological trauma in addition to the social sanction that they are bound to suffer."

What are the laws against the misuse of deepfakes?

India lacks specific laws to address deepfakes and AI-related crimes, but provisions under a plethora of legislations could offer both civil and criminal relief. For instance, Section 66E of the Information Technology Act, 2000 (IT Act) is applicable in cases of deepfake crimes that involve the capture, publication, or transmission of a person's images in mass media thereby violating their privacy. Such an offence is punishable with up to three years of imprisonment or a fine of two lakh. Further, Sections 67, 67A, and 67B of the IT Act can be used to prosecute individuals for publishing or transmitting deepfakes that are obscene or contain sexually explicit acts. The IT Rules, also prohibit hosting 'any content that impersonates another person' and require social media platforms to quickly take down 'artificially morphed images' of individuals when alerted. In case they fail to take down such content, they risk losing the 'safe harbour' protection — a provision that protects social media companies from regulatory liability for third-party content shared by users on their platforms.

Provisions of the Indian Penal Code (IPC) can also be resorted for cybercrimes associated with deepfakes — Sections 509 (words, gestures, or acts intended to insult the modesty of a woman), 499 (criminal defamation), and 153 (a) and (b) (spreading hate on communal lines) among others. The Delhi Police Special Cell has reportedly registered an FIR against unknown persons by invoking Sections 465 (forgery) and 469 (forgery to harm the reputation of a party) in the Mandanna case.

Is there a legal vacuum?

"The existing laws are not really adequate given the fact that they were never sort of designed keeping in mind these emerging technologies," says Shehnaz Ahmed, fintech lead at the Vidhi Centre for Legal Policy in Delhi. She, however, cautions that bringing about piecemeal legislative amendments is not the solution. "There is sort of a moral panic today which has emanated from these recent high profile cases, but we seem to be losing focus from the bigger question — what should be India's regulatory approach on emerging technologies like AI?", she says. She highlights that such a regulatory framework must be based on a market study that assesses the different kinds of harm perpetrated by AI technology.

Pointing out a lacuna in the IT Rules, she says that it only addresses instances wherein the illegal content has already been uploaded and the resultant harm has been suffered; instead, there has to be more focus on preventive measures, for instance, making users aware that they are looking at a morphed image.

Agreeing that there is a need to revamp the existing laws, Mr. Gupta points out that the current regulations only focus on either online takedowns or criminal prosecution but lack a deeper understanding of how generative AI technology works and the wide range of harm that it can cause. "The laws place the entire burden on the victim to file a complaint. For many, the experience that they have with the local police stations is less than satisfactory in terms of their investigation, or the perpetrator facing any kind of penalty," he asserts.

What has been the Centre's response?

The Union Minister of Electronics and Information Technology Ashwini Vaishnaw on November 23 chaired a meeting with social media platforms, AI companies, and industry bodies where he acknowledged that "a new crisis is emerging due to deepfakes" and that "there is a very big section of society which does not have a parallel verification system" to tackle this issue. He also announced that the government will introduce draft regulations, which will be open to public consultation, within the next 10 days to address the issue.

However, the Minister of State for Electronics and Information Technology (MeitY) Rajeev Chandrasekhar has maintained that the existing laws are adequate to deal with deepfakes if enforced strictly. He said that a special officer will be appointed to closely monitor any violations and that an online platform will also be set up to assist aggrieved users and citizens in filing FIRs for deepfake crimes. Mr. Gupta points out, "The advisory issued by the MeitY does not mean anything, it does not have the force of law. It is essentially to show some degree of responsiveness, given that there is a moral panic around generative AI sparked by the Rashmika Mandanna viral clip. It does not account for the fact that deepfakes may not be distributed only on social media platforms."

How have other countries fared?

In October 2023, U.S. President Joe Biden signed a far-reaching executive order on AI to manage its risks, ranging from national security to privacy. Additionally, the DEEP FAKES Accountability Bill, 2023, recently introduced in Congress requires creators to label deepfakes on online platforms and to provide notifications of alterations to a video or other content. Failing to label such 'malicious deepfakes' would invite criminal sanction. The European Union (EU) has strengthened its Code of Practice on Disinformation to ensure that social media giants like Google, Meta, and Twitter start flagging deepfake content or potentially face fines. Further, under the proposed EU AI Act, deepfake providers would be subject to transparency and disclosure requirements.

What next?

According to Mr. Gupta, AI governance in India cannot be restricted to just a law and reforms have to be centred around establishing standards of safety, increasing awareness, and institution building. "AI also provides benefits so you have to assimilate it in a way that improves human welfare on every metric while limiting the challenges it imposes," he says. Ms. Ahmed points out that India's regulatory response cannot be a replica of laws in other jurisdictions such as China, the US, or the EU. "We also have to keep in mind the Indian context which is that our economy is still sort of developing. We have a young and thriving startup eco-system and therefore any sort of legislative response cannot be so stringent that it impedes innovation" she says.

Expected Question

Que. Consider the following statements in the context of international regulation regarding deepfakes:

1. A far-reaching executive order on AI has been brought in the US to manage risks ranging from national security to privacy.
2. There is no provision in the Indian Penal Code (IPC) for cyber crimes related to deepfakes.

Which of the statements given above is/are correct?

(a) Only 1

(b) Only 2

(c) Both 1 and 2

(d) Neither 1 nor 2

Answer : c

Mains Expected Question & Format

Que.: Why are deepfakes dangerous? How are deepfakes being used to spread misinformation in political campaigns? What should be the regulatory response according to experts?

Answer Format :

- In the first part of the answer discuss the danger of deepfakes and the use of deepfakes in political campaigns.
- The second part discusses regulatory responses to address deepfakes.
- Finally give a conclusion giving suggestions.

Note: - The question of the main examination given for practice is designed keeping in mind the upcoming UPSC mains examination. Therefore, to get an answer to this question, you can take the help of this source as well as other sources related to this topic.