Dr. Rajeshwar Singh
B.Tech (IIT Dhanbad), M.A, LL.B, Ph.D
M.L.A.
Sarojini Nagar, Lucknow

सदस्य विधान सभा
उत्तर प्रदेश

K-950, Aashiyana,
Lucknow-226012

Ref. No. RS/1185/MIN/23

Date :

18.11.2023

To
Shri Arjun Ram Meghwal,
Union Law Minister,
Ministry of Law & Justice,
4th Floor, A-Wing, Shastri Bhawan,
New Delhi-110001
E-mail- mljoffice@gov.in

SUB: NEED FOR REWORKING THE LEGISLATIVE FRAMEWORK FOR REGULATING OFFENCES RELATING TO FAKE VIDEOS GENERATED VIDE ARTIFICIAL INTELLIGENCE WHICH POSES A SERIOUS THREAT TO THE FUNDAMENTAL RIGHTS OF THE CITIZENS AND THE PHYSICAL, SOCIAL AND ECONOMIC SAFETY OF THE NATION.

Respected Sir,

I am writing this letter in view of the emergent need in the present times for reworking the legislative framework to regulate offences relating to digital forgeries such as fake videos generated vide artificial intelligence with the aim of maligning the reputation of individuals, thereby violating their Fundamental Right to Privacy guaranteed under Article 21 of the Constitution of India. This in turn has a ripple effect on distorting the democratic and social fabric of the nation. Currently there exists many loopholes in the present laws that cannot prevent Deepfakes from harming people and society in general.

1. The term "**Deepfake**" originated in 2017, when an anonymous Reddit user called himself "Deepfakes." This user manipulated Google's open-source, deep-learning technology to create and post pornographic videos. Deepfake in essence is content (video, audio or otherwise) that is wholly or partially

fabricated or existing content (video, audio or otherwise) that has been manipulated to create digital forgeries. There are currently dozens of applications in the market that users can download for nefarious purposes such as scams and hoaxes, celebrity pornography, election manipulation, social engineering, automated disinformation attacks, identity theft and financial fraud amongst others. Facial expressions can be manipulated per frame, pitch, timbre and language can be adjusted; identities of two or more people can be merged, for example by fusing the faces of two people or by giving a figure the face of a well-known person and the voice of another. Deepfake technology has only been around for a few years but its quality has improved dramatically since then, a trend which should continue, making its detection even harder.

2. It is submitted that the existing legal regime is insufficient to protect citizens from the adverse manipulation of this technology and there is an urgent need for strengthening the existing legal framework to meet these challenges. The scale of the turmoil that will ensue has the capacity to create social, political and economic unrest amongst the nation if the same is not addressed swiftly and in a strong manner.

Need For Urgent Protection Of Citizens Fundamental Rights in The Light Of Rampant Increase In Instances Of Digital Forgeries Vide Artificial Intelligence

3. **Risk to Individual Fundamental Rights and Dignity of Women:** Weaponisation of Deepfakes and synthetic media has the impact of invading the personal lives of individuals. The prominence of non-consensual Deepfake pornography, which accounts for 96% of the total Deepfake videos

online, being the most pervasive illustration. Others like Slut-shaming and revenge porn can have very serious consequences. Deepfakes will increase this problem, not only because of their quantity and quality, but also because Deepfake technology allows a person to create a porn video of someone based on a fully dressed photo. Most importantly, in a prismatic society such as India, the social consequences of such an incident are immense, distorting not just the public image of an individual but also their self-image. The obscene Deepfake video of actress Rashmika Mandana which recently went viral is a case in point, which is a malevolent form of the desire to attain virality on the back of celebrities, a desire that can be unfortunately met and enflamed further by the introduction of such technologies. The unabated perpetuation of such reprehensible activities represents a failure of every citizen to fulfill their Fundamental Duty *to renounce practices derogatory to the dignity of women* provided under Article 51A of the Constitution of India, and needs correcting.

4. **Risk to Political Stability:** Two landmark cases from Gabon and Malaysia that received minimal Western media coverage saw Deepfakes linked to an alleged government cover-up and a political smear campaign. One of these cases was related to an attempted military coup, while the other continues to threaten a high- profile politician with imprisonment. Seen together, these examples are possibly the most powerful indications of how Deepfakes are already destabilizing political processes. Without defensive countermeasures, the Fair and Free elections guaranteed under part XV of our Constitution cannot be ensured as Deepfakes can be used to spread false or misleading information about political candidates and can be used to manipulate public opinion and influence the outcome of an election. With the onset of Parliamentary elections in early 2024 in India, unless Deepfakes

and Synthetic Media is urgently regulated, it could potentially lead to a democratic disruption causing a social and political turmoil and consequently putting the very integrity of our democracy at risk.

5. **Risk to Economic Stability:** Deepfakes can also be used for financial gain, such as manipulating markets in various mannerisms such as: Identity theft - Voice cloning or face-swap video is used to impersonate an individual and initiate fraudulent transactions on their behalf; Impostor Scam - Voice cloning or face-swap video is used to impersonate a trusted government official or family member of the victim and coerce a fraudulent payment; Stock manipulation via fabricated events to falsify a product endorsement that can alter investor sentiments; Malicious bank run - Synthetic photos and text are used to construct human-like social media bots that spread false rumors of bank weakness; Fabricated government action - Voice cloning or face-swap video is used to fabricate an imminent interest rate change, policy shift, or enforcement action. A case in point is the Deepfake message circulated on Whatsapp in 2019, stating that Metro Bank based in the United Kingdom was out of liquidity, which saw people rushing to the Metro Bank to claim all their money and jewellery. This eventually led to its share price falling by 9%. According to Europol Deepfakes are significant in *'perpetrating extortion and fraud, facilitating document fraud, falsifying online identities and fooling KYC [Know Your Customer] mechanisms, falsifying or manipulating electronic evidence for criminal justice investigations, disrupting financial markets and, for example, the theft of trade secrets."* In the case of *State of Gujrat v. Mohanlal Jitamalji Porwal and Anr.*, the Supreme Court has stated that *"[...] the entire Community is aggrieved if the economic offenders who ruin the economy of the State are not brought to books.* Untold damage to the economy and security of the

nation beckons, unless the threat that accompanies the advent of Deepfake technology is not nipped in the bud.

6. **Risk to Rule of Law:** According to the National Judicial Data Grid, as of September, 2021, 4.5 crore cases are pending across all courts in India making the Case Clearance Rate extremely low. Advent of Deepfakes will create a multitude of challenges and foist a greater burden on courts which will struggle to uphold the Rule of Law because of the undue prolonging and vitiation of trials. The courts will struggle to contend with parties' submissions that the evidence against them is fake and fabricated; Deepfakes also increase the risk that a court will falsely assume evidence to be true; Further, in the case of fake news that is distributed to a wider public, but later debunked, the initial (often sensational) fake message will generate significantly more attention than the subsequent rectification, thereby maligning their dignity.

7. **Risk to Society at Large:** Deepfakes may accelerate the move to a post-truth era and increase segregation and stratification, because different Deepfakes will circulate in different echo-chambers or filter bubbles on the Internet. Previously, no commonly available technology could have synthetically created media with comparable realism, so we treated it as authentic by definition. With the development of synthetic media and Deepfakes, this is no longer the case. Every digital communication channel our society is built upon, whether that be audio, video, or even text, is at risk of being subverted. Deepfakes have the potential of distributing disinformation and manipulating public opinion, inciting acts of violence towards minority groups, supporting the narratives of extremist or even terrorist groups, and, stoking social unrest and political polarization.

8. **Risk to Free Press:** The advent of Deepfake technology has led to a situation of *infodemic* worldwide challenging the old notion of "seeing is believing". The tainting of the credibility of audiovisual evidence produced in journalism and its systematic investigation being the obvious fallout of this new normal. Though often put to benign ends, Deepfakes and related synthetic media are also widely seen as posing a potentially serious threat to the right to Freedom of Speech and expression guaranteed under Article 19(1)(a) of the Constitution of India. All democracies need reliable public information; audio and video recordings have become an increasingly important source of such information; but once this source is undermined, one can no longer distinguish authentic instances from fakes. The worry is that, as fake and misleading audio and video images rapidly become realistic, easy to produce, immune to technological detection, and widely disseminated, even authentic audio and video images become untrustworthy—a variant of the so-called 'liar's dividend' that gives a dangerous new weapon to those hoping to undermine democratic norms. This raises a question mark on the authenticity of the entire Press. A case in point is the viral video on WhatsApp in 2018 in India seemingly taken from a CCTV footage, showing a group of children playing cricket on the street when suddenly, two men on a motorbike grab one of the smallest kids and speed away. This "kidnapping" video created widespread confusion and panic, spurring an 8-week period of mob violence that saw the death of at least nine innocent people across the country. The footage was actually a Deepfake video edited from a video of a public education campaign in Pakistan, designed to raise awareness of child abductions.

Global Redressal Mechanisms To Counter The Deepfake Menace

9. **United States**: On the federal level, the U.S. has introduced the DEEP FAKES Accountability Act in 2019 which will require Deepfake creators to disclose their use, will stop the distribution of Deepfakes intended to deceive viewers during an election or harm an individual's reputation, and will fix potential fines and imprisonment for violators. The bill will also establish a task force within the Department of Homeland Security to analyze and mitigate the impact of Deepfakes on national security and calls for increased funding for research into detecting and mitigating the harm caused by Deepfakes. The bill has also introduced stringent content-provenance norms(C2PA) which will have to be mandatorily followed by all technology corporations, regardless of any outcry concerning unbreakable encryption standards. Several states in the USA like California and Texas have already passed laws that criminalize the publishing and distribution of Deepfake videos (within a short duration before and after the election) that intend to influence the outcome of an election. The law in Virginia imposes criminal penalties on the distribution of nonconsensual Deepfake pornography.

10. **European Union**: The EU has taken a proactive approach to Deepfake regulation, calling for increased research into Deepfake detection and prevention, as well as regulations that would require clear labelling of artificially generated content. The most relevant European Deepfake Policy trajectories and regulatory frameworks are: The AI regulatory framework; The General Data Protection Regulation; Copyright regime; e-Commerce Directive; Code of Practice on Disinformation amongst others. The EU has proposed laws requiring social media companies to remove Deepfakes and other disinformation from their platforms. Updated in June 2022, the EU's Code of Practice on Disinformation_addresses Deepfakes through fines of up to 6 percent of global revenue for violators. The code was initially introduced

as a voluntary self-regulatory instrument in 2018 but now has the backing of the Digital Services Act which increases the monitoring of digital platforms for detecting various kinds of misuse.

11. **China**: In 2019, the Chinese government introduced laws that mandate individuals and organizations to disclose when they have used Deepfake technology in videos and other media. The regulations also prohibit the distribution of Deepfakes without a clear disclaimer that the content has been artificially generated. China also recently established provisions for Deepfake providers, in effect as of 10 January 2023, through the Cyberspace Administration of China (CAC). The contents of this law affect both providers and users of Deepfake technology and establish procedures throughout the lifecycle of the technology from creation to distribution. These provisions require companies and people that use deep synthesis to create, duplicate, publish, or transfer information to obtain consent, verify identities, register records with the government, report illegal Deepfakes, offer recourse mechanisms, provide watermark disclaimers, and more.

12. **Canada**: Canada's approach to Deepfake regulation features a three-pronged strategy that includes prevention, detection, and response. To prevent the creation and distribution of Deepfakes, the Canadian government works to create public awareness about the technology and develop prevention tech. To detect Deepfakes, the government has invested in research and development of Deepfake detection technologies. In terms of response, the government is exploring new legislation that would make it illegal to create or distribute Deepfakes with malicious intent. Existing Canadian law bans the distribution of nonconsensual disclosure of intimate images. Further, the Canada Elections Act contains language that may apply to Deepfakes.

13. **South Korea**: In 2020, South Korea passed a law that makes it illegal to distribute Deepfakes that could "cause harm to public interest," with offenders facing up to five years in prison or fines of up to approximately 43,000 USD.

14. **United Kingdom**: The UK government has introduced several initiatives to address the threat of Deepfakes, including funding research into Deepfake detection technologies and partnering with industry and academic institutions to develop best practices for detecting and responding to Deepfakes. The UK has funded research and development to support and spread awareness about the harms of revenge or Deepfake porn in its ENOUGH communications campaign. Further, in November last year, the UK announced that Deepfake regulation would be included in its much-anticipated Online Safety Bill.

Existing Legal Regime In India And Suggestions To Curb The Menace Of Deepfake

15. There is no specific legal provision or enactment that can prevent or mitigate the misuse of Deepfake technology. Some laws which can be invoked in the interim include Copyright Violation, Defamation and cyber felonies. Thus keeping in view the extent of the harm that can be caused by the Deepfake technology, it is imperative to either enact new laws or suitably amend existing ones to deter people from misusing the technology.

a. **Identity theft and virtual forgery**. — These crimes can be prosecuted under Section 66 (computer-related offences) and Section 66-C (punishment for identity theft) of the Information Technology Act, 2000. Also, Sections 420 and 468 of the Penal Code, 1860 could be invoked in this regard.

Suggestion

Punishment under Section 66C of the IT Act should be immediately increased to seven years from the existing three years and the fine should also be increased to Rs. Five lacs from the existing Rs. one lac.

Similarly, punishment under Section 420 and 468 IPC should be increased to 10 years from the existing 7 years.

Another desirable change could be inserting creation of a non-consensual deepfake media as an illustration of the term alteration contained within the definition of personal data breach in the Digital Personal Data Protection Act, 2023. To complement this perhaps an omission of the word 'significant' in Section 33(1) of the aforesaid Act, will remove any doubts regarding the Data Protection Board of India's ability to impose hefty fines on data fiduciaries responsible for safeguarding the data of the injured principal.

b. **Misinformation against Governments**. — The spread of false or misleading information can create confusion and undermine public trust and can be used to manipulate public opinion or influence political outcomes. These crimes can be prosecuted under Section 66-F (cyber terrorism) and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 of the Information Technology Act, 2000. Also, Section 121 of the IPC(waging war against the Government of India) can be invoked to deter any miscreants.

c. **Hate speech and online defamation**. — These crimes can be prosecuted under the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 of the Information Technology Act, 2000. Also, Sections 153-A and 153-B (Speech affecting public tranquility), Section 499 (defamation) of the Penal Code, 1860 could be invoked in this regard.

d. **Practices affecting elections**. — These crimes can be prosecuted under Section 66-D(2) (punishment for cheating by personation by using computer resource) and Section 66-F (cyber terrorism) of the Information Technology Act, 2000. Also, Sections 123(3-A), 123 and 125 of the Representation of the People Act, 1951 could be invoked to tackle the menace affecting elections in India.

Suggestion

Punishment under Section 66D of the IT Act should be immediately increased to seven years from the existing three years and the fine should also be increased to Rs. five lacs from the existing Rs. one lac.

e. **Violation of privacy/obscenity and pornography**. — These crimes can be prosecuted under Section 66-E (punishment for violation of privacy), Section 67 (punishment for publishing or transmitting obscene material in electronic form), Section 67-A (punishment for publishing or transmitting of material containing sexually explicit act, etc. in electronic form), Section 67-B (punishment for publishing or transmitting of material depicting children sexually explicit act/pornography in electronic form) of the Information Technology Act, 2000. Also, Sections 292 and 294 (Punishment for sale etc. of obscene material) of the Penal Code, 1860 and Sections 13, 14 and 15 of the Protection of Children from Sexual Offences Act, 2012 (POCSO) could be invoked in this regard to protect the rights of women and children.

Suggestion

Punishment under Section 66E of the IT Act should be immediately increased to seven years from the existing three years. Punishment under Section 67 should be increased to seven years from existing three years.

Similarly, punishment under Section 292 and 294 IPC should be increased to seven years from the existing two years and fine should be increased to five lac rupees and to seven years from existing three months respectively if the Deepfake related offences have been committed.

f. Further, India's IT Rules, 2021 require that all content reported to be fake or produced using Deepfake technology needs to be taken down by intermediary platforms within 36 hours of being flagged by the authorities, failing which they will lose 'safe harbour immunity' and be liable to criminal and judicial proceedings under the Indian laws. The 36-hour period of taking down the content from the platform should be further brought down to 12 hours.

g. Further all the offences in this regard should be immediately made cognizable and non-bailable. These offences should also be made scheduled offences under the Prevention of Money Laundering Act.

h. The intermediary platforms and other concerned entities should be directed to develop such a technology like disappearing messages which could help in remotely deleting the content from the mobile phones and other devices whenever such content is flagged down by the concerned authorities and/or directed to be taken down from the platforms.

Enacting A Specific Legislation and ancillary changes

Ref. No. RS/1185/MIN/23 Date :

16. There is no specific legal provision that recognizes the individual rights against such menace caused by Deepfakes and related technological advancements. The existing laws which indirectly combats such menace includes provisions of laws on defamation, cheating by impersonation, etc.

17. Moreover, keeping in mind the extent of democratisation of all forms of technology, alongside the positive application of this technology in various facets such as education, a blanket ban on Deepfake Technology is neither feasible nor desirable.

18. Thus, in consonance with the recommendations of the Lodha Committee in the context of Online Gaming stating that a blanket ban always has negative consequences, thus the need of the hour is strict regulation.

19. Lastly, Legislative, Technological and Media Literacy solutions have to be implanted together to tackle the multifarious issue. Measures like creation of a deepfake zoo in the USA(a regulatory sandbox which will encourage the private sector to share insights and collectively come up with resilient technological solutions) need to be instituted. Proliferation of radioactive data sets of video content for easier detection of deepfake videos is another possibility. The government can also limit public access to specific sufficiently advanced detectors(developed through collaboration with the private sector through investments like those made by the Networking and Information Technology Research and Development program in the USA) to keep them in strategic reserve for catching deepfakes capable of jeopardizing national security. Simultaneously, media literacy can very effectively be an initial buffer for the effects of Deepfakes with a robust, affordable, and effective solution that will be able to buy time for any impending legislation. General media literacy makes Deepfakes less

persuasive overall and once deepfakes lose their persuasiveness, people become less likely to share and spread Deepfakes i.e. resilience has to be cultivated amongst the citizens to disabuse them of any trust they may repose in unsolicited audio-video approaches. Revamping the school curriculum with the inclusion of such lessons and an engaged pedagogy by continuously trained teachers should form the bedrock of any such enterprise. Once there are proper detailed laws on combating Deepfakes and regulatory committees are created to act as watchdogs for large technology companies, such as Google and Facebook, will force them to remove Deepfakes from their platforms. Further, in congruence with the Deterrent Theory, people will be deterred if there are new laws regarding consequences and fines to spreading Deepfakes and disinformation in general. In this regard, incorporating suitable changes to the IPR regime to adequately protect publicity rights of celebrated individuals may prove to be a pecuniary disincentive for miscreants intent on maligning these persons.

20. In the current scenario as detailed above, there is an emergent need to rectify the menace being caused by the misuse of this technology.

21. Therefore, in my view, a Committee must be formed of experts comprising of individuals and organizations with substantial experience and specialized knowledge. They should be tasked with studying inputs from various sources, including the State and Union Governments, and other countries(for example legislations like AB-602 and AB-730 in California). Post analyzing these inputs this committee should formulate a comprehensive legal framework to stop any misuse of the AI within a specific time frame.

Ref. No. RS/1185/MIN/23 **Date :**

22. The Law Commission can also be entrusted with the task to study the inadequacies of the prevailing laws and suggest regulatory changes to arrest the misuse of this nascent technology.

23. We have to collaboratively come up with innovative methods on a war footing or risk the situation spiralling out of control.

The people of India have complete faith in the leadership of NDA which possesses the wisdom to understand the problems of the citizens and the drive to address them with alacrity. As such, it is kindly submitted that on the basis of this representation, necessary instructions may be issued to the authorities concerned to have this matter suitably examined for immediate appropriate action, as deemed fit please.

Thanking you,

Yours faithfully,

(Dr. Rajeshwar Singh)
18/11/23

Copy to:
The Honorable Chief Minister,
The State of Uttar Pradesh,
Lok Bhawan, Lucknow,
Uttar Pradesh – 226001,
E-mail: cmup@nic.in