DEEPFAKES THROUGH GENERATIVE AI

Ms. Arma Malik, Ph.D. Scholar, School of Law, Bennett University Greater Noida.

Dr. Deepak Kaushik, Assistant Professor of Law, Bennett University, Greater Noida.

Mr. Ashutosh Mishra, Ph.D. Scholar, School of Law, Bennett University, Greater Noida.

ABSTRACT

Deepfake has become the new reality. A technology that slowly blurs authenticity into a scripted illusion backed by the emergence of AI dissolving the thin line left between truth and imagination. This paper concerns the social implications of deepfakes to analyze the basic and advanced methods and tools used for their creation, the popular use of such deepfakes to spread deception, and legal bottlenecks to curb its maneuver. Legislations both at national and international levels have started responding to these concerns by establishing the authorities and instruments through which an attempt can be made to restrict these subterfuges. The paper also analyses the accountability of intermediaries in deterring such proliferation through the use of its platforms and assisting the domestic legal system in fighting back against these miscreants.

1. INTRODUCTION:

A camera is said to be incapable of lying. But it's becoming very evident in this digital age that it doesn't always reflect reality. With the help of affordable, user-friendly, and readily available video editing software, advanced artificial intelligence and machine learning enable an increasing number of people to produce so-called deep fake audio, video, and picture content. The concern over these clips which show photoshopped, manipulated, and manufactured footage of people and objects in human society is growing. While pornographic deepfakes have been an issue for some time, political deep fakes are a recent concern. These frequently claim to depict a well-known actress, model, or other woman engaging in a sexual act. Still, they only feature the subject's face superimposed over the body of another woman. This feature, known as face-swapping, is regarded as the most straightforward way to produce a deep fake. Face-swapping can be done with various software programs, and the technology involved is highly developed and widely available. Deepfakes cast doubt on one's ability to express oneself freely while also raising issues of personal reputation and image control. This will significantly affect the security and privacy of users. Governments everywhere are responding to these privacy-evading apps; for example, TikTok was banned in India, and the USA is looking into TikTok's privacy practices and is currently passing legislation to lessen the influence of deepfakes in society. Deepfakes are digital representations of videos or other content that have been altered by advanced artificial intelligence to create fake sounds and images that seem authentic. [1] Artificial intelligence (AI) refers to machines that respond to sensory stimulation in a way that is consistent with traditional human responses since humans are capable of thought, judgment, and intention.[2] Deepfake blends the terms fake and deep learning. Artificial intelligence is used to create deepfake content, which is audio and video that has been altered to make someone appear to be speaking or acting in an unreal manner for instance, speech synthesis is modeling a voice so that it may be employed in a video to make a person appear to be saying a thing they are not, and face swapping is sewing the image of another person's face over another. Deepfake is more hazardous than COVID because it is being used far too often to incite hatred and create pornographic videos for retaliation, false news, and financial fraud. Deep learning and generative adversarial networks are the two techniques that underpin

[1] Grace Shao, What 'Deepfakes'Are and How They MayBe Dangerous, CNBC (Mar. 1, 2024, 10:29 PM), https://www.cnbc.com/2019/10/14/what-is-deepfake-and-how-it-might-be-dangerous.html
[2] [Darrell M. West], [What Is Artificial Intellgence] BROOKINGS [(Apr. 24, 2018)] https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/.

deepfake technology. Deep learning is a subfield of machine learning that uses artificial neural networks—algorithms inspired by the structure and functions of the brain—to process and analyse enormous amounts of data. This process is used by neural networks to classify video, sound, or images, creating realistic imitations or modifications. [3] Deep learning has proven beneficial in many domains, such as visual robotics, recognition of speech, and natural language processing. Deepfake content is currently a major concern for people, companies, and governments everywhere.

2. Historical Background:

The foundation dates to the 1990s, but it wasn't given an official title until 2017, three years after Ph.D. candidate Lan Goodfellow who also works at Apple—created Generative Adversarial Network (GAN), which is a crucial component of today's technology. 2017 The term "deepfake" was first used by an anonymous Reddit user who created and shared pornographic videos using Google's open-source, deep-learning technology. User u/deepfakes posted a video on Reddit that allegedly shows an actress Gal Gadot having sexual relations with her stepbrother. Those who enjoyed deepfake's video created a subreddit dedicated to deepfake videos only. which attracted over 15,000 subscribers in just two months and eventually boasted 90,000 On Reddit, user U/deepfakes and other Redditors posted deepfake pornography featuring female celebrities such as Scarlett Johansson, Maisie Williams, Taylor Swift, and Aubrey Plaza.[4] April 2018 Comedian Jordan Peele created a deepfake video of former President Barack Obama who insulted President Donald Trump in his speech.[5] Using video footage of a speaker that already existed, Christoph Bregler, Michele Covell, and Malcolm Slaney's Videos Rewrite program changed it so that the speaker mouthed lines from an audio track. This system is the first to fully automate this kind of facial reanimation. It achieved this by employing machine learning techniques to create associations between the video subject's sounds and facial shape. When filming a scene, the actors' lips could be synchronized with an innovative soundtrack thanks to a program that was created for movie dubbing. The general public can now overlay existing images and videos onto source images

[3] [Will Knight, [Real or Fake? Al is Making it Very Hard to Know, MI] [(May. 1, 2017)] https://www.technologyreview.com/2017/05/01/152061/real-or-fake-ai-is-making-it-very-hard-to-know/
[4] [Samantha Cole] Reddit Just Shut Down the Deepfakes Subreddit] [VICE: MOTHERBOARD] [(Feb.8, 2018)] https://www.vice.com/en/article/neqb98/reddit-shuts-down-deepfakes.
[5] [James Vincent] [Watch Jordan Peele Use AI to Make Barack Obama Deliver a PSA about Fake News, Verge] (Apr.17, 2018) https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed.

or videos using generative adversarial networks, a method for machine learning, thanks to advancements in artificial intelligence and computer vision. This technology's accessibility also encouraged the use of it in political or pornographic parodies. The words "deep learning" as well as "fake" were combined to create the term "deepfake" in 2017.

2.1. Historical development of Deepfake

The creation and detection of deepfakes have advanced significantly throughout technological history. Although technology has a lot of potential applications, like in the entertainment sector, it also has a lot of risks, most notably the possibility of disseminating misleading information and swaying public opinion. From 2014 and 2023, deepfake technology rapidly grew and became more well-known every day:2014: Research on neural network-based recognition of faces was the first to mention advances in deepfake technology.[6] 2015: Voice actor Val Kilmer passed away from throat cancer, but he was able to speak again recently thanks to Sonantic's deepfake technology."[7]2016: Face2Face: Real-time facial recognition and RGB video reenactment, along with the first notable viral deepfake video featuring President Obama.[8]2017: The creation of sophisticated face-swapping algorithms and the introduction of sophisticated tools like DeepFaceLab.[9]2018: On the internet, deepfake pornographic videos[10] begin to surface. Researchers create techniques for identifying manipulated and deepfaked videos that impact political campaigns. Deepfake software is being created by AI companies, and social media platforms are beginning to take action against deepfake videos.2020: Deepfake voice technology has shown that phoney technology is still evolving, and AI companies are creating cutting-edge detection software to counter deepfakes.2021: composed content, including news articles, can now be faked thanks to technological advancements in deepfake text. Deepfakes are therefore increasingly commonly used in scams and other forms of fraud, and methods and software for spotting deepfakes are continuously being

[6] [Thanh Thi Nguyen, Quoc Viet Hung Nguyen, Dung Tien Nguyen, Duc Thanh Nguyen, Thien Huynh-The, Saeid Nahavandi, Thanh Tam Nguyen, Quoc-Viet Pham, Cuong M. Nguyen], [Deep Learning for Deepfakes Creation and Detection: A Survey], [Computer Vision and Image Understanding, 223 (2022) 103525] [(2019)]
[7] [WIPO magazine], [https://www.wipo.int/wipo_magazine/en/] [(last visited Mar. 2, 2024)].
[8] [Thies, J.; Zollhofer, M.; Stamminger, M.; Theobalt, C.; Nießner, M.], [Face2face: Real-time face capture and reenactment of videos], [In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA], [2016].
[9] [Perov, I.; Gao, D.; Chervoniy, N.; Liu, K.; Marangonda, S.; Umé, C.; Dpfks, M.; Facenheim, C.S.; RP, L.; Jiang, J.; et al.], [Deep Face Lab: Integrated, flexible and extensible face-swapping framework. arXiv[(2020)].
[10] Harris, D], [Deepfakes: False Pornography is here and the Law Cannot Protect You Duke L. Tech. Rev.] [2018].

developed.2022: The problem of creating fake news as well as other written content has been addressed by the advancement of deepfake technologies security. 2023: In November of 2023, a deepfake video for Mandana went viral on social media. The Bollywood actor's face was substituted for the face of a British-Indian influencing woman wearing a black workout dress in the video.

3. Methods

There are numerous ways to produce deepfake videos. Among the more modern techniques is a machine learning algorithm that creates lifelike videos from text prompts. With multiple characters, unique motion styles, and accurate background and subject details, Sora can create complex scenes. Not only does the model understand what the user requested from the prompt, but it also understands how those items exist in the world of reality. The first technique for gathering data is to gather a lot of pictures, audio, and videos of the subject you wish to manipulate. You can also use face-swapping to create new videos that mimic the target's original face and appearance. restyle the subject's hair or change their colour; face re-enactment, which entails projecting a subject's expressions from a sole source onto the intended video; and fully manufactured material, which trains an individual's appearance using actual material but produces an entirely artificial image. and Facial swapping and recreation can be applied to pairs of faces. face swape change for the alteration of one's age, gender, appearance, hairstyle, and other physical traits. artificial intelligence alludes to the creation of computer systems that are capable of carrying out tasks that would typically require human intelligence[11]. Rule-based AI is a fundamental technique. making choices. In essence, this approach entails writing a sequence of "if this, then that" instructions. Among the more sophisticated methods are deep learning and machine learning[12]. An icon or figure in a video game, smartphone app, or online platform that symbolizes a specific person[13]. Virtual reality augmentation Technologies that gather, process, and apply data are used to build "hybrid" worlds by superimposing digital layers over the real world. They are both virtual and physical

[11] [Oxford Reference], [oxfordreference.com/display/10.1093/oi/authority.20110803095426960] [(last visited Mar. 3, 2024)].
[12] [Pieter van Boheemen et al], [Cyber Resilience with New Technology - An Opportunity and a Necessity] [Rathenau Instituut], [(2020)]. [https://www.rathenau.nl/sites/default/files/2020-07/REPORT%20Cyber%20resilience%20with%20new%20technology%20-%20Rathenau%20Instituut.pdf]
[13] [Avatar Lexico Dictionaries], [https://www.oxfordlearnersdictionaries.com/definition/english/avatar] [(last visited Mar. 3, 2024)].

at the same time. The technology powers well-known apps like Pokémon Go and Snapchat.

3.1. Tools

Two artificial intelligence models are used by the deepfake technology. Facial expressions, such as smiles, grins, and blinks, are one of those processes that use a dataset of publicly accessible sample images or videos to generate fakes. When a second machine learning system fails to determine the validity of the available images or images of the target individual is fake. At the same time, the other simulation tries to determine whether or not or not the available sample discovered from an auto-encoding is a fake, the deepfake-generated result is probably also convincing enough to human eyes. The term "generative adversarial network" (GAN) describes this technique. An adversarial artificial neural network (also known as a GAN is a distinct type of artificial neural network that can learn from disorganized audio-video data which can be found online and be used to create deepfakes. Variational Autoencoders (VAEs) are a class of less widely utilized techniques. These, in contrast to GAN, rely on two independent networks cooperating. To generate original data, the network of encoders uses a dense but smaller version of the source data as its output. The outcome you want can then be achieved by adjusting and blending the decoder. Adding two Variational Autoencoders would enable you to create a "face swap deepfake," for instance. In this type of deep fake, somebody's face is often superimposed on the body of a celebrity. This face, that the celebrity is using, is encoded using that face encoder. As a result, a new face that is readily believable to the unaided eye is added to the original video.[14] This section discusses the various instruments and techniques for producing and identifying deepfakes in text, images, audio, and video. Furthermore, there are two primary types of deepfake tools: those that generate fakes and those that identify them. A deepfake is impossible to remove once it has been made and widely disseminated online. Deepfake creators can easily reach millions of media consumers worldwide by posting their fake images and videos on popular social media sites like Facebook, Instagram, YouTube, Twitter, and LinkedIn.

[14] [Raina Davis, Chris Wiggins, Joan Donovan], [Deepfakes], [SPRING 2020 SERIES, Rev, 1], [(2020)], [https://www.belfercenter.org/sites/default/files/files/publication/Deepfakes_2.pdf].

3.2. Deepfake Creation tools

TOOLS	**MAIN FOCUS**
Sora	The creator of Chatgpt created 'Sora', an innovative open AI tool that transforms text prompts into lifelike videos. The creator of Chatgpt created 'Sora', a new free artificial intelligence tool that transforms text prompts into lifelike videos.
Faceswap-GAN[15]	It is possible to use the face-swapping and recreation method on pairs of faces.
SimSwap[16]	Randomised face swapping in pictures and videos.
Fewshot FT GAN	Asian face transformations don't make for more consistency. These features include geometric distortion, hair, glasses, and fixed gaze.
FaceShifter	The novel two-stage face-swapping method offers exceptional occlusion sensitivity and precision.
DiscofaceGAN[17]	uses three-dimensional representational learning to be fully controllable and disentangled.

[15] [Shaoanlu,Faceswap-GAN], [https://github.com/shaoanlu/faceswap-GAN], [(last visited Mar. 2, 2024)].
[16] [Neuralchen,SimSwap], [https://github.com/neuralchen/SimSwap], [(last visited Mar. 2, 2024)]
[17] [YuDeng, DiscoFaceGan], [https://arxiv.org/abs/2004.11660], [(last visited Mar.2, 2024)]

Faceapp	permits changing one's age, gender, hairstyle, face, and other physical characteristics
StarGAN[18]	Create regulated and untangled facial image representations using 3D imitative-contrastive learning.
StarGAN-V2	satisfies the need for scalability in multiple domains and a variety of generated graphics.
ATTGAN	Transfer of facial traits under classification limitations
Style-Gan	The style upon which GAN is built produces deepfakes.
Style-Gan2	suggests that to improve the image quality, additional expansion be stopped, the generator be changed, the path length be regulated, and the weight be changed.

4. Accountability of intermediaries and social media on deepfake:

Social media companies and intermediaries must implement content moderation guidelines that stop the spread of dangerous deepfake material. This entails creating precise standards for what types of content are inappropriate, such as deepfakes that might promote violence, disseminate false information, or violate people's right to privacy. Platforms ought to make investments in tools and systems for quickly identifying and eliminating deepfake content. This calls for working together with academics, techies, and law enforcement organizations to create efficient detection systems that can recognize manipulated media. Open disclosure is necessary regarding the techniques employed by middlemen for content moderation, including how they find and remove deepfake content. Transparency encourages user trust and allows the public

[18] Tero Karras, Samuli Laine, Timo Aila], [A Style-Based Generator Architecture for Generative Adversarial Networks], [(2019)] [https://arxiv.org/abs/1812.04948].

and regulatory bodies to hold each other responsible. Social media platforms can inform users about the risks associated with deepfake content. This involves teaching people how to recognize manipulated media and encouraging critical thinking as consuming online content. Intermediaries should collaborate with experts in digital forensics, artificial intelligence, and cybersecurity to stay up to date on the latest advancements in deepfake tactics and defences. Collaborating with law enforcement agencies can also be beneficial for the examination and legal action against individuals or groups that create and disseminate malevolent deepfakes. Governments have the authority to enact laws and regulations to hold intermediaries accountable for the dissemination of harmful deepfake content. To do this, laws requiring platforms to quickly remove illegal or harmful satisfied or face fines may be required.

Intermediaries should consider the ethical implications of their decisions before taking action regarding deepfake content. This means maintaining objectivity when moderating content, upholding individuals' right to secrecy, and finding a middle ground between the right to free speech and the necessity of preventing harm. Research and development spending must be sustained to advance technologies that detect and mitigate the adverse consequences of deepfakes. This means sponsoring academic research, supporting open-source projects, and collaborating with industry partners to develop workable solutions. All things considered, holding social media companies and intermediaries accountable for deepfake content necessitates working with governments, tech companies, researchers, and civil society organizations in a multi-stakeholder approach.

More news, commentary, and debate have centered on the national election than on the role of lies during and after the presidential campaign. Political candidates, mainstream media, interest groups, and other people and organizations that shared stories on social media that were intended to be interpreted as authentic news were all accused of lying. It is difficult to stay on top of the reported lies, much less consider how to respond to them, due to the bewildering variety of them. Given the extent to which lies, fake news, and other forms of misinformation appear to be ubiquitous in today’s public discourse, it is imperative that we first discuss the parameters of what constitutes “fake news,” as defined by this article. As will be seen in the discussion that follows, one of the issues that beset prospective regulatory initiatives is the definitional problem. As others have already noted, it can be challenging to define “fake news,” and there is much disagreement over what exactly constitutes fake news. Recently, a few deepfake videos featuring PM Modi singing and performing the Garbha have gone viral.

5. Legal response on deepfakes by international authority:

Vice President Vera Jourova of the European Commission was told A complex web of laws, including the GDPR, the EU's privacy bill, and numerous national laws, currently protect people in Europe who fall prey to explicit deep fakes. Social media behemoths like X and Meta should assume accountability for streamlining the process by facilitating people's ability to flag and swiftly remove offensive content. The EU's Digital Services Act ought to incorporate these measures (DSA). The United States is drafting new laws to prevent AI-generated deepfakes. The Global Coalition over Digital Safety and the Digital Trust Initiative of the World Economic Forum are working to combat harmful online content. The Federal Trade Commission, or FTC, alerts the public to the alarming rise in deepfakes, which are meant to fool gullible people. AI technologies have made it possible for fraudsters to pose as real people with frightening accuracy and on an even larger scale. The Federal Trade Commission's Chair Lina M. Khan says it's more important than ever to protect Americans from impersonator fraud as voice cloning as well as other AI-driven scams proliferate. That is exactly what our proposed amendments to the final pretending to be someone rule would accomplish—stronger FTC protections against AI-driven identity theft schemes. The goal of these proposed regulations is to stop con artists. Still, they additionally encompass impersonating companies and governments, so if when they become law, they might offer electoral protections. Trump vetoed the 2021 NDAA, but Congress overrode his veto and mandated that the Department of Homeland Security, or DHS, issue a yearly report on deepfakes for the following five years. The report should address every possible negative effect of technology, such as fraud, harm to populations, as well as foreign influence operations. This essentially broadened the purview of the fake information report that the NDAA had asked for the year before. The law mandates that DHS look into possible mitigation and detection techniques in addition to deepfake creation technologies. Lastly, the law mandates that the US Department of Defence look into the potential for adversaries to produce deepfake images with soldiers or their loved ones and suggest adjustments to current procedures. Chinese regulators have released an online manual for handling audio and video content. The potential for novel technologies like "deep fake" to be abused as well as the sharp increase in the number of consumers of online audio and video platforms necessitate the implementation of such regulations. The dissemination of illicit and dangerous content is just one of the offenses listed by the Chinese Internet Administration against individual legitimate rights and interests. The National Television and Radio Administration, an internet watchdog, along with the Ministry of Tourism and Cultural Affairs

introduced the regulation. According to the statement, online services should acquire the necessary certifications by legal and administrative demands before allowing users to access audio and video posts. The document states that protocols for user registration, content review, and information safety management need to be developed and enhanced. It also asked that cyber-security laws enable appropriate user identity verification. According to the agreement, no person or organization may use these services or related information technologies for illicit purposes or to violate the rights and interests of others. Users and online service providers who use cutting-edge technologies like virtual reality or machine learning to create, release, as well as distribute digitally altered audio or video material must prominently label their work. According to the document, no one can use these technologies to produce, publish, or spread false information or news. Union Railway Minister Ashwini Vaishnaw addresses the Rajya Sabha during the Parliamentary Budget Session in New Delhi to hold social media companies more responsible for defamatory content uploaded on them, the government is proposing laws and taking other actions, according to Ashwini Vaishnaw, Minister of Communications, and Information Technology.

6. Domestic legislation of deep fakes (USA, CHINA, EU, SOUTH KORIA):

European Union (EU)

The Digital Services Act[19], which was passed by the European Union, requires social media companies to adhere to labeling regulations, promoting transparency and helping users verify the legitimacy of media. The world's first comprehensive AI law, the AI Act, will regulate deepfakes in the EU. While the use of deepfakes will not be explicitly prohibited by the proposed AI Act, it does aim to regulate them by imposing transparency requirements on their creators under Article 52(3) of the Act. The EU AI Act was unexpectedly agreed upon by negotiators from the European Parliament and Council Presidency in December 2023, it is anticipated that regulators will complete the Act's text in the first quarter of 2024. In 2018, the EU revised and enacted the Audiovisual Media Services Directive (AVMSD) in response to the emergence of online video-sharing platforms in the media landscape. Regulations requiring video-sharing platforms to identify the type of content shared and take action in the best interests of the public, creators, and viewers are called for by the AVMSD. As a result, the

[19] THE DIGITAL SERVICES ACT (DSA) - Regulation (EU) 2022/206[https://www.eu-digitalservicesact.com/][(last visited Mar. 1, 2024)]

AVMSD has clauses that address issues like the dissemination of non-consensual pornographic deepfakes. Following these initial actions, the Code of Practice on Disinformation was published in 2018 as part of the European response to combating online disinformation. Platform operators must distinguish between political and non-political content to accomplish this, and they must demonetize political advertising that spreads false information. In October 2018, Mozilla, Twitter, Facebook, Google, and a few other interested parties signed the Code. Microsoft and TikTok joined later in 2019 and 2020, respectively[20]. According to the European Court of Human Rights, "One of the essential components of personal development is the right to protect one's image, which presupposes the right to control the use of that image." The right to personal life protection, as stated in Article 8 of the ECHR, is closely linked to the right to image protection[21]. Deepfakes are most likely to fall under the purview of the list of copyrighted material, photographic works, and cinematographic works in particular. The creators of those works who own the copyright can file claims and object to their content being used in a deepfake video. General Data Protection Regulation (GDPR): Personal data is usually used in the process of creating a deepfake. These include voice snippets, images, and videos that show specific people and can be used to identify or link back to a specific person. Given that it pertains to an identifiable or identified natural person, a deepfake that portrays a real person may be regarded as personal data.[22]

U.S.A

Legislators in California are considering measures to further prohibit deepfakes, such as more extensive prohibitions on the use of AI in political campaigns and pornographic material used for such nefarious ends. As one of the first states to enact anti-deepfake legislation in 2019 before the current AI frenzy, California already sets the national standard on this issue. Less than a dozen states have since implemented restrictions on the technology, with Michigan being the most recent to do so. Deepfakes are digitally altered pictures or videos that usually inaccurately depict an individual. Most of the Golden State law is comprised of two bills

[20] CODE OF PRACTICE ON DISINFORMATION (2018) [https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation] [(last visited Mar. 1, 2024)]

21 EUROPEAN COURT OF HUMAN RIGHTS (2020) [https://www.echr.coe.int/documents/d/echr/fs_own_image_eng#:~:text=%E2%80%9C%5BA%5D%20person's%20image%20constitutes,essential%20components%20of%20personal%20development] (last visited Mar. 1, 2024).

[22] Article 4 (1) The General Data Protection Regulation (GDPR), https://gdpr-info.eu/art-4-gdpr/ ((last visited Mar. 2, 2024).

sponsored by Assemblymember Marc Berman: A.B. 602, which deals with pornography, and A.B. 730, which deals with political elections. A deepfake victim has the legal right to sue the person or entity disseminating the material under both measures. The US has pushed for the creation of a task force by the Department of Homeland Security (DHS) to handle digital content forgeries, or 'deepfakes'. A lot of states have passed laws of their own to stop deepfakes. U.S. intermediary service providers on the Internet, where unlawful activities like hate speech and defamation occur, are shielded from liability by Section 230 of the Communications Decency Act of 1996[23] for remarks made by users and third parties. This effectively means that only users are responsible for any illegal activity they engage in on the internet; if it is impossible to identify a user, then no one is held liable. The limitation of copyright law is not giving relief to the victim it gives relief only owner, the owner can claim a deepfake video or photograph for copyright. Some acts like the Child pornography prevention act of 1996 were prohibited from showing any child photograph, film, video, picture, or AI-generated image or picture of a minor doing sexually explicit behaviour. Additionally, statements of opinion and statements of fact must be used separately according to defamation law. Libel or slander requires that the falsehood be presented as fact. False light law, a tort that falls under the category of invasion of privacy and concentrates on disseminating malicious falsehoods about a specific individual, frequently intersects with defamation law. The emotional distress brought on by public speech that damages one's reputation is the main focus of false light law, not the actual harm done to one's reputation. Both could be employed as a line of defence against the spread of targeted deepfakes, but they are also open to interpretation because of the subjective definition of reputational harm and the requirement that the defamation be presented as fact rather than opinion. While legal recourse through defamation and false light laws may be available to the targets of these deepfakes, monetary compensation may not be sufficient to repair emotional harm and reputational damage. Legislation is therefore required to deter the creation and dissemination of new deepfakes that have the potential to destroy lives while appearing to be true and to provide criminal penalties for those who abuse the technology. Yoo In-chon, the nation's minister of culture, seemed to be belting out a well-known 'Kim Kwang-seok song' in a deepfake video that was created and distributed by Ryu Ho-jeong, a member of the Committee on Sports, Culture, and Tourism. The demonstration served as a stark reminder of the potential dangers and improper use of deepfake technology, underscoring the critical need for regulation. In February 2024, a deepfake

[23] Communications Decency Act. Sec. 230 [(1996)].

recording featuring the voice of US President Joe Biden was discovered. The audio clip was played during a computerized phone call to supporters of Democrats in the US nation's state of New Hampshire. The phony message instructs recipients not to vote in the nation's primary elections and is delivered in a voice that sounds like Joe Biden. More than 375,000 South Koreans have signed a petition posted on the executive Blue Residence website, calling on the government to intervene against the internet trend known as "deepfake" pornography, in which popular Korean actors' faces are altered to produce explicit pictures that are then circulated. The petition was started just before a Seoul-based company's AI-powered "chatbot" service was forced to close its doors for using vulgar and offensive language, including asking lesbians "disgusting" as well as "creepy."

CHINA

The Procedures for Administrative Law Enforcement by the Cyberspace Administration Departments (the "Enforcement Procedures") were published by the Cyberspace Administration of China ("CAC") on March 18, 2023, and went into effect on June 1st, 2023. The Enforcement Procedures clarify jurisdiction, enhance the due process for the parties involved, and define administrative law enforcement procedures under the framework of "case filing - investigation and evidence collection - hearing - penalty decision - execution" in 58 articles spread across five chapters. The purpose of the Enforcement Procedures is to make CAC's law enforcement operations more methodical and competent. With the authority granted by the Data Security Law (2021), Personal Information Protection Law (2021), Cybersecurity Review Measures (2021), Cybersecurity Law (2016), and National Security Law (2015), the CAC's role has expanded from "content-based" administration (censorship) to data security and privacy protection, enforcement of regulations on data cross-border transfer, and cybersecurity review of network products (2012). People are protected by CAC from being impersonated without their permission by deepfakes, which are photos that are nearly identical to the original and can be readily manipulated or misrepresented. and Personal information protection law expressly forbids the creation of deepfakes without user consent and demands proof that artificial intelligence (AI) was used to create the content. In a similar vein, the People's Republic of China enacted legislation requiring platform operators and app providers to identify, mark, or remove unlabelled content. This law, which went into effect on January 1, 2020, mandates that fake news be removed as soon as it is identified because it prohibits its

creation and dissemination[24]. China's response to the issue of online anonymity was to enact the Cybersecurity Law, which included a highly scrutinized provision requiring real-identity authentication for the registration of online services, such as a valid mobile phone number or official IDs[25].

SOUTH KORIA

A law was passed in South Korea that forbids the distribution of deepfakes that endanger public safety. Violators face a maximum sentence of five years in prison or a fine of up to 50 million won, or roughly USD 43,000. Recently On Tuesday, January 30, 2024, South Korea's special parliamentary committee approved an amendment to the Public Official Election Act that forbade the use of artificial intelligence (AI)-generated deepfakes in political campaign videos during election season.

INDIA

In India, there is no codified law that criminalizes deepfake directly. There are numerous ways that crimes could be committed with deepfake technology. While technology in and of itself is not dangerous, it can be a tool for crimes against people and society. Deepfake can be used to commit the following crimes: Identity theft and deepfake virtual forgeries are serious crimes that can have a big impact on people's lives and society at large. If deepfakes are used to steal identities, misrepresent people, or manipulate public opinion, they can harm an individual's image and credibility and spread false information. These crimes are punishable under Sections 66-C (punishment for identity theft) and 66 (offences relating to computers) of the Computer Technology Act of 2000. This case may also fall under the purview of Sections 420 and 468 of the Penal Code, 1860, which deal with false information directed towards governments. Deepfakes are a severe problem that could have a significant effect on society if they are used to propagate false information, cast doubt on the authority of the government, or foster resentment and disillusionment with it. Spreading misleading or false information can sway people's opinions, erode their trust, and influence political outcomes. Cyberterrorism-related offences are covered by Section 66-F of the IT Act of 2000 and the Information Technology

[24] [Lavender Au] [*China targets 'deepfake' content with new regulation]* (2019). https://technode.com/2019/12/03/china-targets-deepfake-content-with-new-regulation/][(last visited Mar. 3, 2024)]

[25] [Eliza Gkritsi] [*Dust has yet to settle two years after China's landmark cybersecurity law*] (2019)

(Intermediate Guidelines and the Digital Media Ethics Code) Amendment Regulations, 2022. Furthermore, the Penal Code 1860's Sections 124-A and 121, which address insurrection against the Government of India, may be applicable in this case. Hate speech and deepfake online defamation are grave problems that can affect people individually as well as society at large. Deepfakes have the potential to seriously damage people's reputations, well-being, and online communities when they are used to disseminate polarising or defamatory content. The legal penalties for these offences are outlined in the Information Technology (intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022 of the Information Technology Act, 2000. Furthermore, the Penal Code, 1860's Sections 153-A and 153-B (Speech affecting public tranquillity) and Section 499 (defamation) may be applicable in this situation. porn and content that is offensive or violates privacy. The use of this technology allows for the creation of phony photos or videos that show people in ways that have never been seen or heard of, which could harm people's reputations or disseminate misleading information. [26]Deepfakes could also be used maliciously for non-consensual pornography, political propaganda, and disinformation campaigns. Deepfakes can be harmful to society as well as the people whose images or likenesses are used without permission when they are used to spread misleading information or change public opinion. The penalties for breaching someone's privacy is outlined in Section 66-E, publishing or transmitting pornographic material electronically is prohibited by Section 67, and publishing or transmitting content that contains sexually explicit images is prohibited by Section 67-A.[27] Section 67-B of the Information Technology Act of 2000, which describes the penalties for publishing or transmitting pornographic or sexually explicit content featuring children, may be applied to these offences. It is also possible to use Sections 292 and 294 of the Penal Code 1860 and Sections 13, 14, and 15 of the Protection of Children from Sexual Offences Act, 2012 (POCSO) to defend the rights of women and children.[28]

7. Challenges of deepfakes under the Indian legal system:

The difficulties with deepfakes under Indian law: The absence of precise legal definitions and frameworks about deepfakes in Indian law is one of the main obstacles. Deepfakes make it difficult to categorize them under current laws about defamation, fraud, or privacy because

[26] Indian Penal code,1860, No. 45, Acts of Parliament, 1860 (India)
[27] Information Technology Act,2000, No. 21, Acts of Parliament, 2000 (India)
[28] Protection of Children from Sexual offences Act, 2012, No. 32, Acts of Parliament, 2012 (India)

they blur the lines between truth and falsehood. The difficulty of locating and apprehending the makers of deepfakes is another problem. Deepfake technology frequently permits offenders to remain anonymous or to transfer blame, making it difficult for law enforcement to assign blame. It is possible to produce and spread deepfakes internationally, which presents law enforcement with jurisdictional issues. Effective investigation and prosecution require coordination with international authorities. Deepfakes have the potential to breach people's privacy by using their faces to overlay offensive or dangerous content. To adequately address these issues, data protection and privacy laws in India must be strengthened. Deepfakes have the potential to damage people's and organizations' reputations by undermining confidence in the media and information. It is crucial to have legal safeguards in place to stop the spread of false information and rebuild confidence in digital content. The admissibility of deepfakes as evidence in court cases could face difficulties. Guidelines for evaluating the veracity and authenticity of digital evidence, such as deepfake films, must be developed by courts. Law enforcement organizations, legal experts, and people everywhere need to be made more aware of the existence and risks of deepfakes. To effectively combat this phenomenon, capacity building in cyber investigation and digital forensics is essential. Legal frameworks are necessary, but deepfake technology's ethical ramifications also need to be addressed. Achieving a delicate balance between the right to free speech and the need to avoid deceit and harm is a challenging task that needs careful consideration. addressing these issues calls for a multifaceted strategy that includes public awareness campaigns, technological advancements, international collaboration, and legislative changes. Upholding fundamental rights and values while adapting legal frameworks to the changing landscape of digital manipulation is imperative. In situations where a person's privacy is violated through the taking, publishing, or transmission of their images in the media, Section 6E of the IT Act of 2000 applies. A fine of up to two lakh rupees or up to three years in prison are the possible penalties for this offence. Section 66D of the IT Act is another pertinent section. A sentence of up to three years in prison or a fine of up to 21 lakh rupees is stipulated for those who use computer resources or communication devices maliciously to deceive or impersonate someone[29]. The IT Acts provisions may be used to bring criminal charges against those responsible for deepfake cybercrimes in India. Copyright protection for works, including music, films, and other creative content, is provided by Indian law. If someone uses copyrighted works to create deepfakes without permission, copyright owners may file a lawsuit against that person. Section 51 of the

[29] Information Technology Act, 2000, No.21, Acts of Parliament, 2000 (India)

Indian Copyright Act of 1957 stipulates penalties for several offences, including copyright infringement. It forbids using any property that belongs to someone else and over which that person has the sole right without permission[30]. Provisions of Indian law forbid fraud, including financial and identity theft. Although there are currently no laws in India specifically addressing deepfakes, there are some legal provisions and government initiatives that may be used to address the problem. The Indian government will take more action to address the problem and shield people from harm as deepfakes become more common and complex.

8. Conclusion & Suggestions:

Technology is constantly evolving. There is a new development in the field of technology every day. But laws don't change that quickly, and as it stands, deepfakes are not primarily covered by any laws in India or many other nations. To use technological algorithms to address the deepfake issues, the current laws might not be adequate. Some concerns regarding the regulation of deepfakes may surface, including Real-time detection and identification of deepfakes. It is possible to establish attribution and punish offenders. Acknowledging the lack of consensus on whether published content upholds the idea of free speech or infringes upon an individual's right to privacy. Making sure that during these lawsuits, the benefits for the victims are not disproportionate. The effect of the courts' natural rhythm, must be reinstated to control and lessen the effects of deepfakes as they currently exist. If the police forces are prepared and skilled to look into cases involving deepfake content. Technical expertise is required of the legal representatives to handle these kinds of criminal charges and the rapid removal of deepfake material from the internet. There has been much discussion on these topics about cybersecurity. However, it is also our moral duty as a society to work towards halting the spread of harmful content that lacks consensus. gaining knowledge and raising awareness about manipulations and the damage they can do.

The negative effects of producing, sharing, downloading, or uploading fake content online should be taught to young people. It is recommended that regulators embrace novel approaches to address the problem of deepfakes, pinpoint the origin of the content, and take appropriate action to block it. A person's right to the liberty of speech and expression provided by the 19th article of the Indian Constitution is commonly invoked as the main defence used against fake content. What needs to be taken into account is that Our right to privacy begins where our

[30] Indian Copyright Act, 1957, No. 14, Acts of Parliament 1957 (India)

freedom of speech ends. It is our responsibility to realize that our freedom and actions do not tend to impede the enjoyment of rights by any other person. Article 19 of the Indian Constitution guarantees every individual the right to withhold consent; however, this right cannot be invoked to excuse the production and distribution of fake or altered still images or video content that could influence viewers' opinions about the subject matter. As a result, to stop this, authorities and people alike must act in the interests of the community. Internationally, the regulation of synthetic or manipulative media is still in its infancy or underdeveloped. Similar to the United States is dependent on pre-digital technologies tort, criminal, and election offences for both deterrence and redress. The European Commission adopted a balancing strategy when it released a proposal to align AI technology use with EU citizens' fundamental rights and values. Most notably, the proposal lays out minimal requirements like transparency obligations but permits the permissive use of deepfake technologies. One important element that is absent from the proposal is that it does not provide clear guidelines for altered media disclosures, leaving the content creator free to choose how to disclose. Furthermore, designating the content as manipulated media absolves the bad actors of their actions but fails to take into consideration the harm to the victims' reputations and experiences of intimidation and bullying.