Nishith Desai Associates
LEGAL AND TAX COUNSELING WORLDWIDE

MUMBAI SILICON VALLEY BENGALURU SINGAPORE NEW DELHI NEW YORK GIFT CITY

Research

Unmasking Deepfakes

Legal, Regulatory and Ethical Considerations

October 2024

Research

Unmasking Deepfakes

Legal, Regulatory and Ethical Considerations

October 2024

DMS Code: 100042.1

FT INNOVATIVE LAWYERS 2019
FINANCIAL TIMES
ASIA-PACIFIC WINNER

Ranked as the 'Most Innovative Indian Law Firm' in the prestigious FT Innovative Lawyers Asia Pacific Awards for multiple years. Also ranked amongst the 'Most Innovative Asia Pacific Law Firm' in these elite Financial Times Innovation rankings.

Chambers AND PARTNERS

IFLR1000

The LEGAL 500

BENCHMARK LITIGATION

asialaw

WWL

ASIAN LEGAL BUSINESS

INDIA BUSINESS LAW JOURNAL

INDIA BUSINESS LAW JOURNAL THE A LIST 2023 2024 LAW ASIA

INDIA BUSINESS LAW JOURNAL DEALMAKER 2023 LAW ASIA

Mergermarket An Acuris company

ITR WORLD TAX

iel ELITE

BUSINESS TODAY

THE NATIONAL LAW REVIEW

equalFi

ILA IDEX LEGAL AWARDS

LEGAL ERA BY THE PEOPLE. FOR THE PEOPLE. OF THE PEOPLE

LawWiser .com

Disclaimer

Contact

For any help or assistance please email us on **concierge@nishithdesai.com**
or visit us at **www.nishithdesai.com**.

Acknowledgements

Rhythm Vijayvargiya
rhythm.vijayvargiya@nishithdesai.com

Purushotham Kittane
purushotham.kittane@nishithdesai.com

Vaibhav Parikh
vaibhav.parikh@nishithdesai.com

We would like to thank Saurav Kumar, Hiranya Bhandarkar, and Nishka Kapoor for their contribution to the research paper.

Contents

Introduction 1

Dimensions of Deepfakes: Meaning, Types and the Technologies 3

A. Meaning 3

B. Types of Deepfakes 4

C. Technologies behind Deepfakes 7

The Impact of Deepfakes 10

A. Use Cases of Deepfake Technology 10

B. Misuses of Deepfake Technology 12

Legal and Regulatory Implications 16

A. International Regulatory Interventions 16

B. Indian Legal and Regulatory Implications 21

Way Forward 38

Introduction

We are in the era of artificial intelligence **("AI")**, and it is safe to say that today, the speed of technological breakthroughs is directly proportional to the speed of transmission of information as well as misinformation. Content alteration or manipulation is an age-old concept,[1] but the easy accessibility of various tools has contributed to the growth rate of online orchestrated content increasing by 400% every year.[2]

At present, deepfakes are one of the most advanced forms of synthetically generated media and it is predicted that they could account for up to 90% of the online available content in the upcoming years.[3]

One of the first technologies that produced deepfake-like results was the Video Rewrite Program in 1997,[4] which automated facial reanimation in videos. Based on a similar concept, the Generative Adversarial Network[5] **("GAN")** was introduced in 2014, which was further improvised by Nvidia[6] in 2017 to produce good quality forged images.

With the GAN algorithms slowly catching traction, later in 2017, the term "deepfakes" was coined when an unidentified user on the social media platform Reddit had developed an algorithm,[7] that the user used to transpose celebrity faces onto pornographic content.[8] The likeness of the celebrities was superimposed in the pornographic content to the extent that it appeared to be true. Owing to the nature of the content being shared, it instantly became viral and widespread. The unidentified user used to go with the username *deepfakes*, and hence, the technology came to be commonly referred to as deepfake technology.

Essentially, deepfakes refer to "fake" content that is created using "deep learning" technology.[9] Apart from this oversimplified meaning of deepfakes, it has also been defined by the Oxford University Press[10] as: *"a video of a person in which their face or body has been digitally altered so that they appear to be someone else, typically used maliciously or to spread false information."* With the advances in AI-synthesized techniques, deepfakes are now also capable of creating highly realistically sounded voices.[11]

Today, the technology is widely known for creating realistic-looking images and videos of people and objects that may or may not exist. About ninety-five percent of the deepfake content was in the form of non-consensual porn till December 2018, and Rana Ayyub's case[12] was one of the biggest examples in the deepfake history to depict the depth of revenge porn plotting through this technology.

1 Data manipulation has been in practice since the 1890s; Please see: https://www.loc.gov/collections/spanish-american-war-in-motion-pictures/articles-and-essays/the-motion-picture-camera-goes-to-war/remember-the-maine-the-beginnings-of-war/, (last accessed October 10, 2024).

2 Please see: https://www.globenewswire.com/en/news-release/2022/10/27/2542944/0/en/Deepfake-content-on-the-internet-is-growing-at-the-rate-of-a-whopping-400-year-on-year.html, (last accessed October 10, 2024).

3 Please see: https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html, (last accessed October 10, 2024).

4 Please see: http://chris.bregler.com/videorewrite/VideoRewrite.pdf, (last accessed October 10, 2024).

5 Please see: https://arxiv.org/pdf/1406.2661.pdf, (last accessed October 10, 2024).

6 Please see: https://developer.nvidia.com/blog/generating-photorealistic-fake-celebrities-with-artificial-intelligence/, (last accessed October 10, 2024).

7 The Reddit user's software was called FakeApp, and was used by BuzzFeed in 2018 to create a deepfake video of Barack Obama addressing this issue. Please see: https://www.buzzfeed.com/craigsilverman/obama-jordan-peele-deepfake-video-debunk-buzzfeed.

8 Please see: http://www.jatit.org/volumes/Vol97No22/7Vol97No22.pdf, (last accessed October 10, 2024).

9 Please see: https://iimk.ac.in/uploads/faculty/CAIS_20220810062848.pdf, (last accessed October 10, 2024).

10 Please see: https://languages.oup.com/google-dictionary-en/, (last accessed October 10, 2024).

11 Please see: https://arxiv.org/pdf/2005.13770.pdf, (last accessed October 10, 2024).

12 Please see: https://www.huffpost.com/archive/in/entry/deepfake-porn_in_5c1201cfe4b0508b213746bd, (last accessed October 10, 2024).

However, over time, in around 2019, the use cases of deepfake technology started to come into the limelight too.[13] Big Tech organizations like Microsoft,[14] Google,[15] and Samsung[16] adopted GAN for content generation. Currently, the uses of deepfake technology can be seen in the medical,[17] entertainment,[18] marketing,[19] and fashion,[20] industries among others; setting some noteworthy examples and depicting that the technology can have an array of beneficial uses too.

Irrespective of the events that brought deepfake technology into the focus of attention, the possibilities that arise with its use are endless. Additionally, the reason for its significant breakthrough is how convincing these media employing deepfake technology are to the perceptible human mind, and as time progresses, these altered media are becoming increasingly closer to reality.[21]

Content manipulation has now become mainstream and easily accessible, with the quality of forged content being so high that it becomes impossible to filter out with a bare eye. However, it is important to acknowledge that with the quality of this fabricated content rising, the implications and in turn liabilities would rise too.

In this paper, we have systematically examined the types and underlying technologies behind deepfakes, followed by the use and misuse cases of such technology. Additionally, we have discussed the legal, regulatory, and ethical implications of deepfakes in India and other jurisdictions.

13 Please see: https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/, (last accessed October 10, 2024).

14 Please see: https://mspoweruser.com/microsoft-has-made-their-own-ai-powered-image-generator-and-its-pretty-meh/, (last accessed October 10, 2024).

15 Please see: https://ai.googleblog.com/2020/11/using-gans-to-create-fantastical.html, (last accessed October 10, 2024).

16 Please see: https://news.samsung.com/global/behind-the-snapshot-how-the-galaxy-s21s-ai-improves-your-photos-in-the-blink-of-an-eye-single-take, (last accessed October 10, 2024).

17 Please see: https://reader.elsevier.com/reader/sd/pii/S0300571222002676?token=5DA9F8AD14B7D8634EBC69B7FCBB9E0414B10BAEF3C33865833CC51F36FD35144A6558CDA050505AE667AC4F6C17900C&originRegion=eu-west-1&originCreation=20230225021628, (last accessed October 10, 2024).

18 Please see: https://www.youtube.com/channel/UCi38HMIvRpGgMJ0Tlm1WYdw, (last accessed October 10, 2024).

19 Please see: https://www.news18.com/news/tech/cadburys-new-ai-tool-will-let-you-create-free-ad-with-shah-rukh-khans-face-and-voice-4358408.html, (last accessed October 10, 2024).

20 Please see: https://www.forbes.com/sites/katiebaron/2019/07/29/digital-doubles-the-deepfake-tech-nourishing-new-wave-retail/?sh=2d473fbc4cc7, (last accessed October 10, 2024).

21 Please see: https://www.researchgate.net/publication/338144721_Deepfakes_Trick_or_treat, (last accessed October 10, 2024).

Dimensions of Deepfakes: Meaning, Types and the Technologies

A. Meaning

Deepfakes are typically understood to be manipulated media.[1] However, to just say that deepfakes are synthesized media created by using identity-swapping algorithms would not do justice to the depth of this concept. In a world where every image or video on the internet may be altered to some degree using deep learning, it becomes difficult to agree on a common definition of deepfakes since the line where a picture or video becomes "manipulated" or "fake" is blurred.

Nevertheless, the usage of *deep learning technology* to produce *fake* content contributed towards the portmanteau – Deepfakes.[2] Even though there is no globally agreed upon definition for deepfakes, they came to be understood conventionally as *"fake images created by an advanced image editing, deep learning software."*[3]

To understand the nature of deepfakes, it is crucial to establish that deep learning is a machine learning tool that forms the basis of powerful algorithms. These algorithms are designed to use enormous volumes of data to learn to perform specific tasks. Simply put, deep learning is a type of AI, and it uses artificial neural networks to mimic the learning process of the human brain.[4] Deep learning is a fundamental element of the deepfake phenomenon, as the algorithms feed onto the existing and available data[5] to create fake images and videos that humans cannot distinguish from authentic ones.[6]

There is no distinguishable division that separates 'acceptable' forms of augmentation or special effects to media from those which may typically be seen as manipulated. Although, the manipulated content that is not generated using AI or deep learning technology usually bears a lower quality and is easier to point out from the genuine content. These relatively poorly engineered photos and videos are usually referred to as *cheapfakes* or *shallowfakes.*[7]

A July 2021 study undertaken for the European Parliament has defined deepfakes as *"manipulated or synthetic audio or visual media that seem authentic, and which feature a person that appears to say or do something they have never said or done, produced using artificial intelligence techniques, including machine learning and deep learning."*[8]

1 Please see: https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/, (last accessed October 10, 2024).

2 Please see: https://d1wqtxts1xzle7.cloudfront.net/65354450/IRJET_V7I1265-libre.pdf?1609941112=&response-content-disposition=inline%3B+filename%3DIRJET_A_Brief_Study_on_Deepfakes.pdf&Expires=1677852603&Signature=Ybf6GhM91ZkpLWPUmRAI6WrpJBbQXXFkHxuLIPwsmhlwde99WwNaj-i64CUtQLrfiJ2aZLYjXMWAa75rcTHqXp3lClYx5yJAPx-q9od8qTUwTeV2VPIViuq7SeHMDzUzL40mWtmQW7VuPlt0MUkiq3DLlBz0xhCaVqqoc5d2OiCStVaufRq5Jh9gCulojTMAzl15b9uLPTVeMcZFtadesEcjyxhJMSTOX0MiZStQosOp81sy91AFmeUTzBMOQbVk-r-QZiEBxA2frryvmMHJrGzg6kpZ1ssHncdwam-P7j1SaeYXU-BLG0PEYIYGjryJgxIrcxUFIYWa6f5R9XdYcA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA, (last accessed October 10, 2024).

3 Please see: https://digitalcommons.georgiasouthern.edu/cgi/viewcontent.cgi?article=1742&context=honors-theses, (last accessed October 10, 2024).

4 Please see: https://www.ibm.com/in-en/topics/deep-learning, (last accessed October 10, 2024).

5 Please see: https://www.accenture.com/nl-en/blogs/insights/deepfakes-how-prepare-your-organization, (last accessed October 10, 2024).

6 Please see: https://arxiv.org/pdf/1909.11573.pdf, (last accessed October 10, 2024).

7 Please see: https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf;: https://www.lexisnexis.co.uk/legal/guidance/deepfakes#:~:text=A%20deepfake%20is%20a%20form,a%20realistic%20but%20fake%20video, (last accessed October 10, 2024).

8 Please see: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf, (last accessed October 10, 2024).

There has been a plethora of attempts to define deepfakes, and it is an arduous task to find an agreeable definition for them. Although, the following elements seem to be largely accepted as being fundamentally part of what deepfakes are:

a. Alteration of media using deep learning technological tools;

b. Altered media that involves the assumption of an identity of a person; and

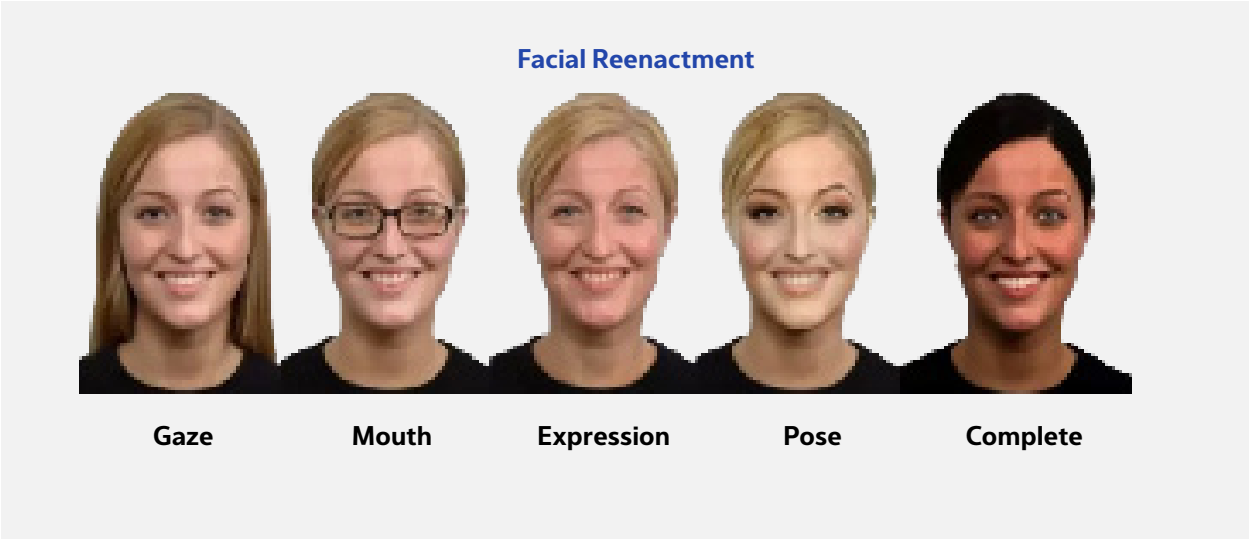
c. An untrue altered media depicted as true to the casual viewer.

B. Types of Deepfakes

Most deepfake videos involve facial manipulation, where a face, faces, or parts thereof are manipulated and then superimposed or inserted on another face or part thereof.[9] The manipulation may be static, as in the case of an image, or dynamic as in a video. The deepfake technology may be used to create doctored content through reenactment, replacement, editing, and audio or visual synthesis.[10]

Reenactment[11]

Reenactment, commonly understood as expression swap, is essentially where the images are fed into the algorithm to drive the expressions of a person and create a deepfake. A reenactment deepfake allows the creator to impersonate an individual's appearance and in turn, control what they appear to say or do. This is one of the most dangerous types of deepfakes, as they can be used to perform acts of defamation, spread misinformation, and tamper with evidence.



Source: Medium[12]

9 Please see: https://iimk.ac.in/uploads/faculty/CAIS_20220810062848.pdf, (last accessed October 10, 2024).

10 Please see: https://arxiv.org/pdf/2004.11138.pdf, (last accessed October 10, 2024).

11 Please see: https://ro.ecu.edu.au/cgi/viewcontent.cgi?article=2530&context=theses_hons, (last accessed October 10, 2024).

12 Please see: https://medium.com/voxel51/have-deepfakes-influenced-the-2020-election-c0fc890aca0f., (last accessed October 10, 2024).

Replacement[13]

Replacement, commonly understood as face or identity swap, is when the images are fed into the algorithm to replace the face of one person in an image or video with that of another person. This method poses danger too, as it can be used for the production of revenge pornographic videos, cheating, and financial fraud. Some of the use cases of this method include generating memes or satirical content, or face swapping for anonymization of one's identity in public content.

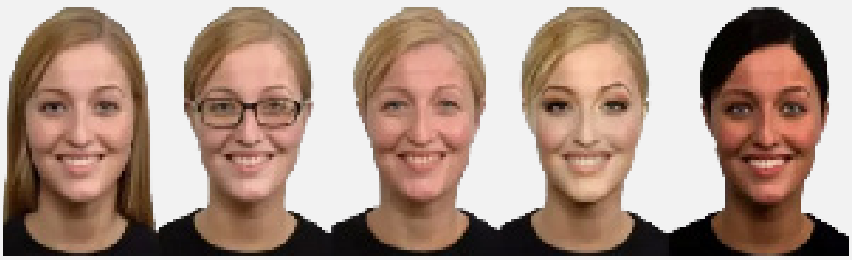
Face Replacement

Source: InfoQ[14]

Editing[15]

Through editing or attribute alteration, a portion of the face is manipulated to achieve a different result. Media retouching done through this method comprises altering facial features such as gender, age, ethnicity, etc. This method is misused widely, for instance, for the removal of a victim's clothes for humiliation or entertainment. At the same time, this method of creating deepfakes is used for entertainment and easy editing purposes too.

Face Editing



Source: InfoQ[16]

13 Please see: https://insights.sei.cmu.edu/blog/how-easy-is-it-to-make-and-detect-a-deepfake/, (last accessed October 10, 2024).

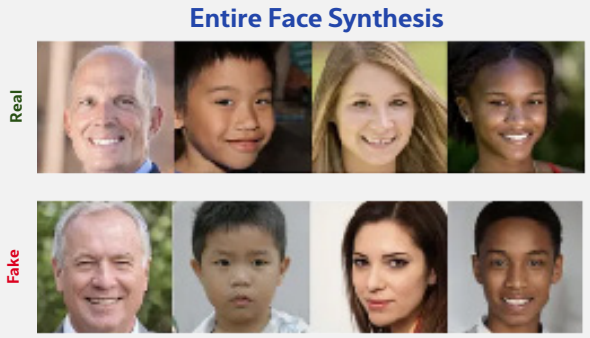
14 Please see: https://www.infoq.cn/article/u7jtw13waskch2mpl918.

15 Please see: https://www.mdpi.com/2313-433X/9/1/18, (last accessed October 10, 2024).

16 Please see: https://www.infoq.cn/article/u7jtw13waskch2mpl918, (last accessed October 10, 2024).

Visual Synthesis[17]

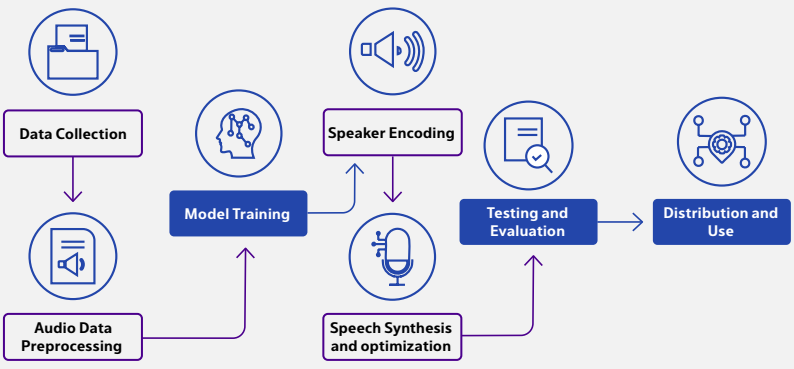
Visual synthesis allows the creation of a deepfake with no target images or videos as a basis. Entire face synthesis is usually based on datasets that are easily available online. This method allows the creation of unreal personas online and may be useful for non-personal communication while at the same time holding the potential for fraud or the spread of misinformation. For e.g., this method also be used to generate non-existing face images for the characters in movies and games.



Source: Medium[18]

Audio Synthesis[19]

In cases of audio synthesis, an audio clip of a person speaking, or even the speech in writing is obtained, and an audio clip is produced out of the speech or text in the tone and style of a chosen target, thereby giving the impression that the target person is speaking. This method poses a threat as it makes phone-call frauds very straightforward for the scammers. Although, speech synthesis has also been relied upon in the medical industry to help patients with impairment.



Source: BotTalk[20]

17 Please see: https://arxiv.org/pdf/1912.04958.pdf, (last accessed October 10, 2024).

18 Please see: https://shivkumarganesh.medium.com/deepfakes-production-detection-using-various-deep-learning-methodologies-3221e6002dd2, (last accessed October 10, 2024).

19 Please see: https://arxiv.org/ftp/arxiv/papers/2111/2111.14203.pdf, (last accessed October 10, 2024).

20 Please see: https://bottalk.io/learn-with-bottalk/everything-about-deepfake-voice/, (last accessed October 10, 2024).

C. Technologies behind Deepfakes

Methods of Creation[21]

There exists an array of machine learning techniques and algorithms which can be used to create deepfakes, and the quality of the result is dependent not only on the quality of the algorithm but also on the quality and quantity of data used to train the algorithm. It is no coincidence that the majority of deepfake videos on the internet involve famous people, as they have large amounts of public images and videos available that can be used to train the algorithm.

Deepfakes are usually created using variation s or combinations of generative networks and encoder-decoder networks. Different kinds of Generative Neural Network **("GNN")** architectures are used for the generation of artificially orchestrated content. A neural network is essentially an algorithmic model which generates content based on an input. Below are the different GNNs used for the creation of deepfakes.

I. Encoder-Decoder Networks ("EDN")

An EDN is made up of two or more algorithms that compress and decompress pictures. Deepfakes are created using autoencoders, which recreate a subject based on the information relative to the data provided to it. The subjects used to create the deepfake must have as many similarities as possible to the intended output so that the shared encoder can identify meaningful features and transfer them appropriately.

The quantity and quality of data presented to the algorithm during the process is directly proportionate to the result's quality. Autoencoders trained on millions of photos of a face from various angles and under a variety of lighting conditions will perform far better and produce more realistic results than encoders trained on a few hundred images with little variation across them.

II. Convolutional Neural Network ("CNN")

A CNN learns pattern hierarchies in data and is hence a very efficient tool for dealing with pictures. CNN has the unique capacity to extract features from images, which may then be utilized in a variety of applications. A CNN extracts a layer of characteristics from a dataset and applies it to an input image; by combining numerous layers of these characteristics, it is possible to construct realistic-looking deepfakes.

III. Generative Adversarial Networks ("GAN")

GAN was first introduced in 2014 and has since become one of the most popular tools for creating deepfakes. GANs are made up of two neural networks that compete with one another. They can "learn from their blunders" and detect patterns in enormous volumes of visual data. Several licensed picture and video editing software, augmented and virtual reality applications, and cutting-edge medical imaging tools use them.[22] There are two popular image translation frameworks that use the fundamental principles of GANs in the creation of deepfakes: (i) image-to-image translation which is often relied upon for generating high-resolution

21 Please see: https://arxiv.org/pdf/2004.11138.pdf, (last accessed October 10, 2024).

22 Please see: https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/, (last accessed October 10, 2024).

imagery with better fidelity, and (ii) CycleGAN which is a combination of two GANs and can be used for object transfiguration and style transfer.

IV. Recurrent Neural Networks ("RNN")

An RNN is a form of neural network that remembers its internal state after processing a dataset and can subsequently be used to process further datasets. RNNs are frequently employed in deepfake generation to modify audio and, in some cases, video.

With the usage of the above-discussed underlying technologies, multiple consumer-grade websites and applications have also been created that allow users to produce deepfakes.
Some of these include FakeApp, FaceSwap, DeepFaceLab, DFaker, FaceSwap-GAN, DeepFake-tf, ZAO, Auto-FaceSwap, FSGAN, FewShotFace, and StarGAN.[23]

Further, video editing software like Adobe After Effects or Wondershare Filmora can be used to create cheapfakes, if not deepfakes. Most of these applications are publicly available and can produce a large range of manipulated data – depending on the quality of the images fed and the platform used.

Methods of Detection

Deepfakes are a result of superimposition which becomes possible through the extensive combinations of datasets and the technologies discussed in the previous section. Since the algorithms put to use for the manipulation are only "editing" pre-existing data, gaps continue to remain in the results. Minute features like eye movements, lighting, and shadows often create discrepancies that allow for deepfake detection.[24] There exist different sets of tools for the detection of deepfakes in images, and in videos.[25] For instance, the detection of deepfake images may be done through the deployment of the following methods:

- Detection of GAN-generated images may be done through an image preprocessing step and differentiating the swapped images from the genuine.[26]
- A two-phased deep learning method may be implemented, in which the first phase uses a feature extractor through the common fake feature network and the second phase categorizes the fake and real images based on the results of the first.[27]
- A CNN model may be used to identify the images where the facial expressions are maliciously tampered with.

Similarly, deepfakes videos may be detected through the implementation of the following methods:

- Detection of inconsistency in temporal features across video frames, as the video manipulation is carried out through the fabrication in every single frame. This method also uses CNN technology to extract frame-wise features, which are later fed into another program to differentiate between manipulated and real content. [28]

23 Please see: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse62922020.pdf, (last accessed October 10, 2024).
24 Please see: https://www.warse.org/IJATCSE/static/pdf/file/ijatcse62922020.pdf, (last accessed October 10, 2024).
25 Please see: https://arxiv.org/pdf/1909.11573.pdf, (last accessed October 10, 2024).
26 Please see: https://download.arxiv.org/pdf/2007.10466v1.pdf, (last accessed October 10, 2024).
27 Please see: https://www.scirp.org/pdf/jcc_2021051813373227.pdf, (last accessed October 10, 2024).
28 Please see: https://arxiv.org/pdf/1905.00582.pdf, (last accessed October 10, 2024).

- Upon observing signs such as a person in deepfakes who may have a lot less frequent blinking than that in untampered videos, biological signals analysis became one of the methods to detect manipulated videos. This method focuses on unnatural movements of lips, eyes, or entire faces inserted into deepfake videos.[29]

For the detection of deepfake audios/voice, technologies like usage of voice biometrics are used. The technology uses a person's unique voice pattern to verify their identity. It includes spectral analysis consisting of audio signal analysis to detect voice patterns, deep-learning algorithms that analyze an individual's voice and recognize unique characteristics that are difficult to replicate in deepfakes, etc.[30]

In addition to the technical methods used for the detection of deepfakes, Big Tech[31] has taken some steps too:

- Facebook, Microsoft, and Amazon conducted a Deepfake Detection Challenge in partnership with leading academic institutes. The idea behind this Challenge was to accelerate the creation of methods to detect deepfakes. To facilitate this research on deepfakes for this challenge, a huge dataset of high-quality manipulated videos was created.[32]
- To mitigate the potential harm and abuse, Google released a dataset of manipulated audio to support the development of fake audio detectors.[33] Google also published a dataset of visual deepfakes to support the development of deepfake detection tools.[34]
- X (Twitter) released a synthetic and manipulated media policy in an effort to detect and remove deepfake content from the platform.[35]
- While acknowledging that deepfakes are a growing threat, Intel released a software called FakeCatcher which has the ability to detect a deepfake with around 96% accuracy in less than a second.[36]
- OpenAI announced the launch of its disinformation detector which primarily seeks to detect images created by its generative AI platform Dall-E, which would first be shared with disinformation researchers during the test phase.[37]

A mechanism to enable the easy detection of deepfakes will be vital in ensuring the authenticity of photos, videos, and audio, especially as manipulated media grows more advanced. The ability to swiftly identify fake content will contribute towards the reduction in the spread of misinformation and scams carried out through the deployment of deepfake technology. Further, developing reliable and accessible deepfake detection tools will help foster trust in digital media while protecting against identity theft, financial fraud, and other harmful activities.

29 Please see: https://arxiv.org/pdf/1806.02877.pdf, (last accessed October 10, 2024).

30 Please see: https://www.mobbeel.com/en/blog/voice-deepfake/., (last accessed October 10, 2024).

31 Please see: https://www.wsj.com/articles/tech-companies-step-up-fight-against-deepfakes-11574427345, (last accessed October 10, 2024).

32 Please see: https://ai.facebook.com/blog/deepfake-detection-challenge-results-an-open-initiative-to-advance-ai/, (last accessed October 10, 2024).

33 Please see: https://www.blog.google/outreach-initiatives/google-news-initiative/advancing-research-fake-audio-detection/, (last accessed October 10, 2024).

34 Please see: https://ai.googleblog.com/2019/09/contributing-data-to-deepfake-detection.html, (last accessed October 10, 2024).

35 Please see: https://help.twitter.com/en/rules-and-policies/manipulated-media, (last accessed October 10, 2024).

36 Please see: https://www.intel.com/content/www/us/en/newsroom/news/intel-introduces-real-time-deepfake-detector.html#gs.rjwz79, (last accessed October 10, 2024).

37 Please see: https://www.msn.com/en-my/news/other/openai-launches-disinformation-detector-over-fears-on-faked-images/ar-BB1mghdk, (last accessed October 10, 2024).

The Impact of Deepfakes

Deepfakes, as well as its underlying technology, have had an array of positive and negative implications across different business sectors throughout the world. In light of the same, we have discussed how select industries have adapted the deepfake technology; while there exist a range of misuse cases too.

A. Use Cases of Deepfake Technology

Owing to the nature of deepfakes, they were mostly criticized as a concept for the longest time, until organizations including BigTech utilized the technology to set an example of how deepfakes create a pool of opportunities for multiple industries. Examples of deepfake technology in some of these industries are discussed below.

Film and Advertising Industry

In the recent few years, the film and advertising industry has started to rely on deepfakes for various purposes, like de-aging to show an actor's past in the movie,[1] dubbing the movies with more accuracy,[2] protecting the identities of individuals in films with sensitive issues,[3] and producing personalized advertisements for the consumers to interact with.[4]

A documentary[5] showcasing the torture and murder of queers in Chechnya and the struggle of the activists who fought to help them escape used deepfake technology to superimpose the faces of queer activists from around the world onto the faces of the queer from Chechen, in order to protect the identities of those who were affected and cut down on the repercussions of the documentary on them. Soon after, an AI-based software[6] emerged which allowed a smooth dubbing of movies with high accuracy, which is said to be better than the traditional dubbing methods. Although the technology still operates on a combination of automated and manual dubbing, it has opened doors for discussion of advancements and their impact.

Similar technologies have been used in several advertisements to create deepfakes of Shah Rukh Khan for Cadbury,[7] Lionel Messi for Lays,[8] Salman Khan for Pepsi,[9] and Sachin Tendulkar for Ageas Federal Life Insurance.[10]

1 Please see: https://www.creativebloq.com/news/the-irishman-deepfake; https://fortune.com/2023/02/14/tech-forward-everyday-ai-hollywood-movies/, (last accessed October 10, 2024).

2 Please see: https://www.theverge.com/2021/5/18/22430340/deepfake-dubs-dubbing-film-tv-flawless-startup, (last accessed October 10, 2024).

3 Please see: https://www.vox.com/recode/2020/6/29/21303588/deepfakes-anonymous-artificial-intelligence-welcome-to-chechnya, (last accessed October 10, 2024).

4 Please see: https://analyticsindiamag.com/deepfakes-will-change-advertising-forever/, (last accessed October 10, 2024).

5 Welcome to Chechnya, HBO Films.

6 Flawless AI's synthetic film dubbing technology.

7 Please see: https://www.outlookindia.com/website/story/india-news-how-the-notjustacadburyad-campaign-us-using-ai-to-bring-joy-to-local-c/398614, (last accessed October 10, 2024).

8 Please see: https://www.fritolay.com/lays-i-messi-messages-laysunited, (last accessed October 10, 2024).

9 Please see: https://brandequity.economictimes.indiatimes.com/news/advertising/salman-khan-pepsi-ad-prem-hum-aapke-hain-kaun-bollywood-deepfake-technology/90048824, (last accessed October 10, 2024).

10 Please see: https://brandequity.economictimes.indiatimes.com/news/advertising/ageas-federal-life-insurance-uses-deepfake-tech-to-recreate-young-sachin-tendulkar-in-latest-campaign/90354065, (last accessed October 10, 2024).

While the purpose of these advertisements has been to create personalized messages as well as portray de-aged versions of the actors in them, the deepfake technology has generally showcased several benefits for the film and advertisement industries combined as it offers low-cost production and yields a better quality of content. Additionally, as discussed above, deepfake technology can be utilized for the portrayal of sensitive content through synthesized identities – which would ease the content creation in areas that may be deemed as triggering by some sections of the society.

Gaming Industry

The gaming industry can be highly benefitted through deepfake technology, as it allows the game developers to create immersive experiences for players by creating realistic characters through real-life actors.[11] The players' immersion in the game would increase exponentially when their in-game avatar tracks their facial and bodily movements to mimic them in the game – which again is now possible through the employment of deepfake technology.[12]

Online games such as "AI Dungeon," "Project December" and "Cogmind" have deployed AI-generated content to create dynamic and contextual narratives for the plays and provide an immersive experience.[13] Further, Epic Games[14] and Roblox[15] have indicated that they will be integrating AI-generated content in their games to create art, text, and other elements including avatars for a real-time experience for the players. Interestingly, there is another aspect as well, where the voice of a famous video game "Genshin Impact" characters' voice actor was cloned to create audio and video deepfakes that were widely circulated over social media.[16]

Additionally, the technology also allows players to either sound like their preferred character for entertainment purposes or mask their voice to sound like another gender or age group for safety purposes.[17] Similar to the advantages in the film and advertising industry, deepfakes allow for low-cost production, higher content quality, and implementation of safety measures in the gaming industry.

Healthcare Industry

The Healthcare Industry has shown various innovative uses of deepfake technologies, ranging from the creation of tools to research on the ways to cure diseases,[18] to tools for cancer detection.[19]

---

11 Please see: https://medium.com/predict/why-deepfakes-will-make-you-play-video-games-instead-of-movies-99ee5c2d7c9e, (last accessed October 10, 2024).

12 Please see: https://www.wipo.int/wipo_magazine/en/2022/02/article_0003.html, (last accessed October 10, 2024).

13 Please see: https://aicontentfy.com/en/blog/ai-generated-content-for-video-game-narratives#:~:text=In%20the%20context%20of%20video,on%20vast%20amounts%20of%20data, (last accessed October 10, 2024).

14 Please see: https://blusharkmedia.medium.com/a-new-era-in-gaming-epic-games-vs-steam-on-ai-integration-7972de74ec77, (last accessed October 10, 2024).

15 Please see: https://www.wired.com/story/roblox-generative-ai-gaming-universe/, (last accessed October 10, 2024).

16 Please see: https://www.resemble.ai/video-game-voice-actor-deepfakes/, (last accessed October 10, 2024).

17 Please see: https://www.wipo.int/wipo_magazine/en/2022/02/article_0003.html#:~:text=What%20is%20a%20%E2%80%9Cdeep-fake%E2%80%9D%20and,generate%20a%20realistic%20human%20experience, (last accessed October 10, 2024).

18 Please see: https://www.siliconrepublic.com/innovation/deepfake-ai-healthcare-diseases-insilico-medicine-pharma, (last accessed October 10, 2024).

19 Please see: https://www.technologyreview.com/2019/07/05/134286/ai-deepfakes-gans-medical-cancer-diagnosis/, (last accessed October 10, 2024).

Detailed research on artificial empathy showed that deepfakes datasets could help train doctors on the patients' facial emotion recognition during clinical interaction and train them to respond appropriately.[20]

One of the pharmaceutical organizations, Insilico Medicine used deepfake technology to break down the rules of drug design and generate medicinal molecules from the same, for its platform Pharma.AI.[21] Similarly, research was published by Retrace, a leading entity in dental AI, which shed light on how deepfake technology could be used to improve the diagnosis of periodontal disease.[22]

The medical industry often relies on deepfake technology for inpainting, which is essentially the process of using a set of visuals to fill in the gaps in the images through superimposition. In addition to this, voice cloning and conversation technology employed through the use of deepfakes can be immensely beneficial to patients suffering from speech disorders.[23]

Social Media Industry

The social media industry has contributed highly to the accelerated use of deepfake technology. Starting with Reddit, now almost every social media platform has the presence of deepfake content in some form.

There exist several ways deepfake technology has been adopted on social media platforms which primarily serve the purpose of entertainment for the users. A TikTok account features deepfake videos impersonating Tom Cruise and has millions of followers,[24] while platforms like Instagram and Snapchat[25] provide filters for the users to get creative with, which are also created using this technology. Microsoft[26] and Google[27] introduced their deepfake technology-based art software which enable the users to create artworks from scratch using pre-existing media.

This technology is useful for the social media industry as it facilitates the creation of a captivating experience for the users. The online environment provided through the deployment of deepfake technology allows the users to get creative without indulging in any illegal or ethically incorrect activity.

B. Misuses of Deepfake Technology

Deepfake technology has been infamous, irrespective of the use cases that have started to emerge lately. The fundamental cause for this notion is the widespread misuse and the threat it poses to individuals, businesses, society, and democracy, in various forms.[28] We have identified the key misuse cases for deepfake technology below and analyzed their impact.

20 Please see: https://www.jmir.org/2022/3/e29506, (last accessed October 10, 2024).

21 Please see: https://www.wired.com/story/molecule-designed-ai-exhibits-druglike-qualities/, (last accessed October 10, 2024).

22 Please see: https://www.sciencedirect.com/science/article/pii/S0300571222002676?via%3Dihub, (last accessed October 10, 2024).

23 Please see: https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/191023_LewisNelson_TrustYourEyes_WEB.pdf; https://www.respeecher.com/blog/everything-you-need-know-about-deepfake-voice-its-synthetic-voice-ethical-counterpart, (last accessed October 10, 2024).

24 Please see: https://thenextweb.com/news/deepfakes-taking-over-tiktok-how-to-spot, (last accessed October 10, 2024).

25 Please see: https://freshlybuilt.com/gans/, (last accessed October 10, 2024).

26 Please see: https://mspoweruser.com/microsoft-has-made-their-own-ai-powered-image-generator-and-its-pretty-meh/, (last accessed October 10, 2024).

27 Please see: https://ai.googleblog.com/2020/11/using-gans-to-create-fantastical.html, (last accessed October 10, 2024).

28 Please see: https://www.orfonline.org/expert-speak/debating-the-ethics-of-deepfakes/, (last accessed October 10, 2024).

The spread of misinformation and disinformation

Misinformation could refer to all information that is false, irrespective of the intention to mislead the audience, while disinformation is the intentional propagation of maliciously false information with the intent to deceive or mislead the audience, such that their subsequent actions are guided by such false information.[29]

With easy access to tools used to create deepfakes, the process of content manipulation to weave a false narrative has never been easier. The main cause of concern with these false narratives is the intentional as well as the unintentional harm they pose to the whole society at large.[30] The contribution of deepfakes in spreading misinformation and disinformation also results in the users' lack of confidence when accessing news content, causing a post-truth crisis.[31]

A lot of misinformation and disinformation propagated online is targeted towards re-enforcing existing biases held by people.[32] Amplification of such misinformation is typically motivated by economic and political incentives held by certain actors. At times fake news is also used to harass a particular individual or group of individuals.[33]

In 2019, a slowed-down deepfake video of Nancy Pelosi, the then United States House of Representatives speaker was widely circulated and made it appear like her words were gurned. The video was fact-checked and found to be fake, but Facebook refused to take it down. Later that year, a deepfake of Mark Zuckerberg was circulated on Instagram, which sent a message regarding how Facebook owns its followers.[34] Indian Bollywood actresses have also been targeted through the creation of their deepfake videos. A deepfake video of actress Rashmika Mandana[35] went viral, closely followed by another explicit video of Aishwarya Rai Bachchan.[36] In another viral deepfake video, actresses Anushka Sharma and Aishwarya Rai Bachchan were seen discussing an investment opportunity.[37]

The deepfake technology has adversely impacted the elections across countries. During the Bangladesh general elections in 2023, sexually explicit deepfake videos of female politicians (including that of Rumin Farhana and Nipun Roy) went viral .[38] Further during the 2023 elections in Nigeria, deepfake clips of one of the presidential candidates were circulated where it was allegedly incriminated that he was planning to engage in rig-balloting.[39] A similar instance was reported in Slovakia, where deepfake audio of an opposition candidate was spread over where he was shown plotting a rig.[40]

---

29 Please see: https://www.businessinsider.in/tech/how-to/misinformation-vs-disinformation-what-to-know-about-each-form-of-false-information-and-how-to-spot-them-online/articleshow/80295200.cms, (last accessed October 10, 2024).

30 Please see: https://www.orfonline.org/expert-speak/debating-the-ethics-of-deepfakes/, (last accessed October 10, 2024).

31 Please see: https://www.cogitatiopress.com/mediaandcommunication/article/view/3494/3494, (last accessed October 10, 2024).

32 Please see: 2018, https://theconversation.com/misinformation-and-biases-infect-social-media-both-intentionally-and-accidentally-97148, (last accessed October 10, 2024).

33 Please see our Research Paper on Misinformation: http://www.nishithdesai.com/fileadmin/user_upload/pdfs/Research_Papers/Make_It_or_Fake_It.pdf.

34 Please see: https://www.spiceworks.com/it-security/cyber-risk-management/articles/what-is-deepfake/, (last accessed October 10, 2024).

35 Please see: https://www.businesstoday.in/bt-tv/video/ai-deepfake-video-of-actress-rashmika-mandanna-going-viral-amitabh-bachchan-raises-concern-404762-2023-11-06., (last accessed October 10, 2024).

36 Please see: https://timesofindia.indiatimes.com/videos/etimes/bollywood/after-rashmika-mandanna-aishwarya-rai-bachchans-deepfake-video-in-swimwear-goes-viral-netizens-react-ai-magicvery-dangerous/videoshow/106014255.cms., (last accessed October 10, 2024).

37 https://timesofindia.indiatimes.com/videos/etimes/bollywood/deepfake-videos-of-anushka-sharma-and-aishwarya-rai-bachchan-discussing-investment-opportunities-spark-concerns-on-social-media/videoshow/105927554.cms, (last accessed October 10, 2024).

38 Please see: https://www.tribuneindia.com/news/trending/from-rashmika-mandanna-to-bangladeshi-politician-filmed-in-a-bikini-90-per-cent-of-deepfake-videos-online-are-pornographic-571782, (last accessed October 10, 2024).

39 Please see: https://www.ft.com/content/bd75b678-044f-409e-b987-8704d6a704ea, (last accessed October 10, 2024).

40 Ibid.

It is pertinent to note that the realistic nature of deepfake content eases the weaponized use of this technology[41] for spreading false information as it causes people to believe and remember experiences that never occurred.[42]

Institution of fake evidence in legal proceedings

Deepfake content, when created with the intention of malice and used to portray fake information, poses to be a huge threat if used in a courtroom. The impact of the manipulated information in such a case is very deep as it has the potential to sway judicial decisions and misguide the courts.[43]

In criminal cases, evidence plays an extremely crucial role. With the increasing ease of access to deepfake technology, manipulating the evidence to drive the court case in one's favor or against a certain party may not be an incredibly difficult process.[44]

In 2020, a UK court observed the admission of deepfake evidence in a child's custody case. A heavily doctored recording of the child's father was submitted with the intention to discredit him, as he was making violent threats towards his wife in the audio. The evidence was challenged and upon examination, the recording was found to be manipulated and the court dismissed the same as fake.[45]

The lack of awareness surrounding deepfakes' existence and detection, coupled with the lack of stringent mechanisms to challenge a technologically manipulated piece of court evidence contribute highly to the possibility of employing this technology to take advantage of the grievance redressal systems.[46]

Revenge and social harm

The nearly real depiction of fabricated content produced through deepfake technology has made it very convenient to take revenge on any individual or organization, or to target them with an intention of manipulation, coercion, or blackmail. All that is needed by a bad actor in such a case is – the publicly available software and datasets, along with a few images of their target. The target in such a case could either be a celebrity[47] with a wide range of their pictures and videos on the internet, or any other person who might have posted their photos or videos online, on social media platforms for instance.

The pornographic material posted by the Reddit user 'Deepfakes' was one of the biggest reasons behind the technology becoming mainstream, especially for the generation of pornographic content.

41 Please see: https://www.wm.edu/offices/global-research/research-labs/pips/white_papers/2019-2020/hogan-final.pdf, (last accessed October 10, 2024).

42 Please see: https://www.sciencedirect.com/science/article/abs/pii/S1053810009000798, (last accessed October 10, 2024).

43 Please see: https://sawanandsawan.com/deepfake-videos-as-evidence-in-court/#:~:text=Digital%20evidences%20such%20as%20images,harder%20and%20uncertain%3B%20losing%20credibility, (last accessed October 10, 2024).

44 Please see: https://surovellfirm.com/criminal-law/prevalence-of-deepfakes-causes-courts-to-question-validity-of-evidence/, (last accessed October 10, 2024).

45 Please see: https://www.thenationalnews.com/uae/courts/deepfake-audio-evidence-used-in-uk-court-to-discredit-dubai-dad-1.975764, (last accessed October 10, 2024).

46 Please see: https://deliverypdf.ssrn.com/delivery.php?ID=283125078096096102095100111116113067025005033082061059074070066094081066107115067111126106001127041058000013099122094007115073025054009058001006113003091114104001074026040032125114021081089087117104113064000085004081092080117089118088108097071093023021&EXT=pdf&INDEX=TRUE, (last accessed October 10, 2024).

47 Please see: https://www.vice.com/en_us/article/bjye8a/reddit-fake-porn-app-daisy-ridley, (last accessed October 10, 2024).

Soon after the wide circulation of these pornographic videos with the faces of celebrities superimposed on them, cases of revenge pornography started to surface.[48]

In 2018, an investigative journalist became a victim of a deepfake pornographic plot. The journalist's face was used in a pornographic video that was widely circulated in an effort to silence her. Following this, the journalist's phone number and address were now being circulated along with the video. The humiliation faced by the journalist not only had repercussions on her emotional, mental, and physical health; but it strongly affected her career too. This issue escalated to a point where the United Nations had to intervene to ensure that the journalist was safe.[49] The deepest concern with revenge pornography on the target individuals remains to be the fact that irrespective of the quality of the manipulated media, its repercussions are bound to follow.[50]

Scams and cyberattacks

Individuals with malicious intentions are capable of using audio and visual deepfakes to scam unsuspecting victims for financial gain. As discussed above, deepfakes can very easily be used to blackmail people. False media, including video and audio, can be used to brainwash someone into giving up money, personal information, or favors.[51]

Deepfake audio has the strongest utility when it comes to scamming and cyberattacks. The attacker only has to ensure that the data comprising vocal samples of the mimicked subject are fed into a computer algorithm, which can be accessed from public sources such as speeches, presentations, corporate films, and interviews to create a voice model. Once a sufficiently solid deepfake audio profile has been created, it can be used in conjunction with specialist text-to-speech software to generate scripts for the fake voice to read.[52]

This voice model can then be used for scams such as voice phishing.[53] In 2019, a CEO of a UK-based energy firm was tricked into believing that he was on a phone call with his boss from Germany, and he transferred approximately $243,000 to the bank account of a Hungarian scammer, who used the deepfake technology to undertake this attack.[54]

This attack, being one of the first of its kind, raised high concerns over the heightened possibilities of cybercrime as technology improves. It was also highlighted that the above scam came into notice when the victim realized that the phone number was from Australia. However, it may be difficult for victims in other cases to avoid such scams, especially when they are contacted through a video call, with a familiar voice and a face on the other end of the screen.[55]

48 Please see: https://towardsdatascience.com/deepfakes-harms-and-threat-modeling-c09cbe0b7883, (last accessed October 10, 2024).

49 Please see: https://www.huffpost.com/archive/in/entry/deepfake-porn_in_5c1201cfe4b0508b213746bd, (last accessed October 10, 2024).

50 Please see: https://www.theguardian.com/society/2023/mar/07/laura-bates-for-teenage-girls-escaping-harassment-revenge-porn-and-deepfake-porn-is-impossible, (last accessed October 10, 2024).

51 Please see: https://towardsdatascience.com/deepfakes-harms-and-threat-modeling-c09cbe0b7883, (last accessed October 10, 2024).

52 Please see: https://www.unionbank.com/commercial/insights/fraud-prevention/deepfake-scams-protection-from-deepfake-fraud, (last accessed October 10, 2024).

53 Please see: https://venturebeat.com/security/deepfake-phishing/, (last accessed October 10, 2024).

54 Please see: https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=2500561d2241, (last accessed October 10, 2024).

55 Please see: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402, (last accessed October 10, 2024).

Legal and Regulatory Implications

A. International Regulatory Interventions

Multiple jurisdictions across the globe have been consistently making efforts to regulate the usage of deepfake technology by implementing measures that encompass both criminalizing malicious creation and distribution of deepfakes, safeguarding against their deceptive use in various contexts, and outlining guidelines for responsible development and deployment within legal and ethical boundaries. We have discussed the major jurisdictions that have laws in force, as well as jurisdictions that have proposed laws that seek to tackle the issues posed by deepfake technology.

China

Owing to the rise in AI and deepfake frauds,[1] the People's Republic of China has been one of the few countries at an early stage to actively undertake efforts to curb the misuse of AI with a focus on deepfakes since 2019.[2] In 2021, Internet Information Service Algorithmic Recommendation Management Provisions[3] were introduced to regulate the usage of AI and algorithm-based recommendations made to users on various online platforms. Further, in November 2022, the Cyberspace Administration of China, along with the Chinese Ministry of Industry and Information Technology, and Ministry of Public Security introduced the Administrative Provisions on Deep Synthesis in Internet-Based Information Services **("Deep Synthesis Law")** in order to specifically address the high rising problems posed by deepfake technology[4] and this law has been in force since 10 January 2023. In addition to the above, China has also introduced the law on Interim Measures for the Management of Generative Artificial Intelligence Service Management Services[5] on 15 August 2023 and the Trial Guidelines on the Review of Science and Technology Ethics on 8 October 2023.[6]

The Deep Synthesis Law defines 'deep synthesis technology' to mean *"generative and synthesizing algorithms, with deep learning and virtual reality as representative examples, to produce text, images, sound, video, virtual settings, and other such information, including but not limited to generation of textual content, voice content, non-voice sound content, content with human faces and characteristics, and content involving other virtual settings."*[7] *The law defines 'deep synthesis service providers' as "organizations providing deep synthesis services or providing technical support to deep synthesis service;"* and most of the obligations are primarily enforceable on such organizations which fall under the ambit of deep synthesis service providers.

1 Please see: https://www.voanews.com/a/as-deepfake-fraud-permeates-china-authorities-target-political-challenges-posed-by-ai-/7137321.html , (last accessed October 10, 2024).

2 Please see: http://english.www.gov.cn/statecouncil/ministries/201911/30/content_WS5dec5772c6d0bcf8c4c1886c.html, (last accessed October 10, 2024).

3 Please see: https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/, (last accessed October 10, 2024).

4 Please see: https://digichina.stanford.edu/work/translation-internet-information-service-deep-synthesis-management-provisions-draft-for-comment-jan-2022/ for the translated version, (last accessed October 10, 2024).

5 Please see: https://www.airuniversity.af.edu/Portals/10/CASI/documents/Translations/2023-08-07%20ITOW%20Interim%20Measures%20for%20the%20Management%20of%20Generative%20Artificial%20Intelligence%20Services.pdf for the translated version, (last accessed October 10, 2024).

6 Please see: http://english.scio.gov.cn/pressroom/2023-10/08/content_116731459.htm., (last accessed October 10, 2024).

7 Article 2 of the Deep Synthesis Law.

These obligations include the requirement of individuals' consent before deployment of deep synthesis technology on them,[8] prohibition on fake news dissemination through the usage of deepfake technology,[9] the need for identification of the individuals' real identities,[10] and a requirement for labels to inform the users that the content they are viewing has been created and/or altered through the usage of deepfake technology.[11]

European Union

Recognizing the technological and societal impacts of the deepfake technology, the research wing at the European Parliament released a policy report titled "Tackling deepfakes in European policy" in July 2021 **("Deepfakes Report")**.[12] The Deepfakes Report provides an assessment of the risks associated with deepfakes along with the policy measures that could be incorporated under the Digital Services Act, 2023 **("DSA")** and the Artificial Intelligence Act **("AI Act")**.[13] In June 2022, the Strengthened Code of Practice on Disinformation **("2022 Code")**[14] was released by the European Union **("EU")**. The 2022 Code mandates transparency requirements for platforms that operate on AI systems and publish content that has been generated or modified by AI, including deepfakes.[15] A transparency taskforce has also been established by the EU, which seeks to review and adapt the transparency commitments with respect to disinformation including the implications of using deepfake technology to spread the same.[16] However, the 2022 Code appears to be voluntary, which implies that the liability for non-compliance would only extend to the signatories, who may have to pay up to six percent of their total global revenue in an event of non-compliance.[17] The DSA as enacted in 2022,[18] governs internet safety and platform accountability. It provides legal backing to the 2022 Code through provisions on online disinformation.[19] The DSA also contemplates for the users to identify deepfakes, by requiring online platforms to give users a tool to indicate that their content is synthetically manipulated with a malicious intent to deceive other users.[20]

The AI Act as originally proposed in 2021[21] did not explicitly define the term 'deepfakes' in the list of key definitions.[22]

8 Article 12 of the Deep Synthesis Law.

9 Article 6 of the Deep Synthesis Law.

10 Article 9 of the Deep Synthesis Law.

11 Article 14 of the Deep Synthesis Law.

12 Please see: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf, (last accessed October 10, 2024).

13 Please see: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689, (last accessed October 10, 2024).

14 Please see: https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation, (last accessed October 10, 2024).

15 Commitment 15 of Strengthened Code of Practice on Disinformation.

16 Please see: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_3665, (last accessed October 10, 2024).

17 Please see: https://www.cnet.com/news/politics/eu-strengthens-disinformation-rules-to-target-deepfakes-bots-fake-accounts/, (last accessed October 10, 2024).

18 Please see: https://www.eu-digital-services-act.com/Digital_Services_Act_Articles.html, (last accessed October 10, 2024).

19 Please see: https://www.reuters.com/technology/google-facebook-twitter-will-have-tackle-deepfakes-or-risk-eu-fines-sources-2022-06-13/, (last accessed October 10, 2024).

20 Article 35 (1) k of the Digital Services Act.

21 Please see: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, (last accessed October 10, 2024).

22 Article 3 of the draft AI Act.

However, through amendments made in 2023 to the proposed text of the AI Act, a definition was introduced to define 'deepfakes' as *"manipulated or synthetic audio, image or video content that would falsely appear to be authentic or truthful, and which features depictions of persons appearing to say or do things they did not say or do, produced using AI techniques, including machine learning and deep learning."*[23] The final text of the AI Act, while not having an express definition of deepfakes, appears to describe a deepfake as a generated or manipulated image, audio, or video content that appreciably resembles existing persons, objects, places, or other entities or events and would falsely appear to a person to be authentic or truthful. The AI Act provides that users generating any such content must disclose that the content has been artificially generated or manipulated.[24] The AI Act also appears to classify AI systems deployed by law enforcement agencies to detect deepfake content as 'High-risk AI systems'.[25]

United States of America

Due to the growing impact of unregulated uses of artificial intelligence and the risks associated with it, an executive order was passed in the United States of America **("USA")** on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence in October 2023 **("2023 AI Order")**.[26] While the term deepfakes has not been used or defined in the 2023 AI Order, deepfakes are one of the many AI-related issues that the 2023 AI Order aims to govern[27] as it provides guidelines on the content generated or synthesized by AI. In furtherance to mitigate the risks posed by synthetic content generated by AI, various US state actors have been tasked to device tools, methods, and standards to detect synthetic content, and label such content using watermarking, authenticating, auditing, and maintaining such synthetically generated contents.[28] Further, it also prevents generative AI from producing non-consensual sexual content of any real person and content on child sexual abuse.[29] The US Government has proceeded to implement some of the milestones laid down in the 2023 AI Order.[30]

Additionally, certain state laws govern aspects of deepfakes. For instance, in 2019, the states of Texas[31] and California[32] enacted a law that prohibits the production and dissemination of deceptive deepfake videos intended to harm political candidates or influence elections. Another law in Texas, which was introduced in 2023, bans the generation and dissemination of pornographic deepfakes and makes it a criminal offence punishable by up to one year of imprisonment and a fine of up to USD 4000.[33]

---

23 Amendment 203 to the draft AI Act as adopted by the European Parliament on 14 June 2023, see: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html, (last accessed October 10, 2024).

24 Article 52(3) of the draft AI Act provides transparency requirements for AI systems on deepfakes as "Users of an AI system that generates or manipulates text, audio or visual content that would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do, without their consent ('deep fake') [emphasis added], shall disclose in an appropriate, timely, clear and visible manner that the content has been artificially generated or manipulated, as well as, whenever possible, the name of the natural or legal person that generated or manipulated it".

25 Annex III read with Article 6(2) of the AI Act.

26 Please see: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/, (last accessed October 10, 2024).

27 Please see: https://www.bloomberg.com/news/newsletters/2023-11-06/biden-ai-executive-order-shows-urgency-of-deepfakes, (last accessed October 10, 2024).

28 Section 4.5 (a) of the Order on AI.

29 Section 4.5 (a) (iv) of the Order on AI.

30 Please see: https://www.whitehouse.gov/briefing-room/statements-releases/2024/04/29/biden-harris-administration-announces-key-ai-actions-180-days-following-president-bidens-landmark-executive-order/, (last accessed October 10, 2024).

31 Please see: https://legiscan.com/TX/text/SB751/id/1902830, (last accessed October 10, 2024).

32 Please see: https://www.theverge.com/2019/10/7/20902884/california-deepfake-political-ban-election-2020, (last accessed October 10, 2024).

33 Please see: https://statutes.capitol.texas.gov/Docs/PE/htm/PE.21.htm#:~:text=September%201%2C%202019.-,Sec.%2021.165,-.%20%20UNLAWFUL%20PRODUCTION%20OR, (last accessed October 10, 2024).

As a recent development in the USA, the Deepfakes Accountability Bill was proposed in 2023 **("Deepfakes Bill")**[34] and it contemplates a detailed framework governing deepfake technology as well as the harm posed by deepfakes. The Deepfakes Bill defines deepfakes to also include their derivatives.[35] It provides for the establishment of a Deepfakes Task Force[36] obliges online platforms to maintain minimum technical capabilities to identify the origin of the content being hosted,[37] make disclosures, and have an active system in place to detect deepfakes.[38] The Deepfakes Bill also provides for criminal as well as civil penalties in case of non-compliance by online platforms.[39]

Furthermore, in January 2024, the No Artificial Intelligence Fake Replicas and Unauthorized Duplications (No AI FRAUD) Bill, 2024 **("No AI Fraud Bill")** specifically seeks to protect the rights of persons against the misuse of AI-led technologies, including the deepfake technology. The lawmakers have quoted several use as well as misuse cases that contributed towards the introduction of the No AI Fraud Bill and have delved into the depths of explaining the meaning and relevance of 'likeness' in relation to deepfakes.

United Kingdom

The United Kingdom **("UK")** government published its National Artificial Intelligence Strategy in September 2021 addressing the risks associated with deepfakes, and the targeted misinformation in the section on 'Governing AI effectively'.[40] Currently, the Online Safety Act, 2023 **("Online Safety Act")**[41] imposes certain obligations on online platforms regarding illegal pornographic content which also appears to include content created through the use of deepfake technology. Moreover, the UK government is planning to bring a law governing AI with specific regulations on deepfake content including labelling requirements for AI-generated content to counter online misinformation.[42] It would also regulate the use of deepfake technology within political campaigns and require that any public use should be clearly labelled.[43]

Under the Online Safety Act, the definition of deepfakes or AI-generated content has not been provided,[44] although, it provides a holistic definition of 'content' to mean any communication within online space, including those generated using automated tools.[45] The Online Safety Act prohibits non-consensual pornographic deepfakes by obligating online platforms that generate and host pornographic content[46] to comply with authentication and age verification of users to access such content.[47]

---

34 Please see: https://www.congress.gov/bill/118th-congress/house-bill/5586/text, (last accessed October 10, 2024).

35 Section 2 (n) (3) of the Deepfakes Bill defines deepfake as "The term 'deepfake' means any video recording, motion-picture film, sound recording, electronic image, or photograph, or any technological representation of speech or conduct substantially derivative thereof— (A) which appears to authentically depict any speech or conduct of a person who did not in fact engage in such speech or conduct; and (B) the production of which was substantially dependent upon technical means, rather than the ability of another person to physically or verbally impersonate such person."

36 Section 7 of the Deepfakes Bill.

37 Section 10 (a) of the Deepfakes Bill.

38 Section 7 (b) of the Deepfakes Bill.

39 Please see: https://abcnews.go.com/Politics/bill-criminalize-extremely-harmful-online-deepfakes/story?id=103286802, (last accessed October 10, 2024).

40 Please see: https://www.gov.uk/government/publications/national-ai-strategy, (last accessed October 10, 2024).

41 Please see: https://www.legislation.gov.uk/ukpga/2023/50/enacted/data.pdf, (last accessed October 10, 2024).

42 Please see: https://www.openaccessgovernment.org/uk-considers-clear-labeling-law-combat-ai-deepfakes/161861/, (last accessed October 10, 2024).

43 Please see: https://lyon.tech/public/news-detail/deepfake-regulation-what-is-next-for-ai-laws-in-the-uk, (last accessed O ctober 10, 2024).

44 Part 2, Key Definitions of the Online Safety Act.

45 Section 236(1) defines content as "anything communicated by means of an internet service, whether publicly or privately, including written material or messages, oral communications, photographs, videos, visual images, music and data of any description".

46 Section 79 (2) of the Online Safety Act.

47 Sections 81, 81 of the Online Safety Act.

Further, the Online safety Act has tasked the Office of Communications in the UK to device strategies and tools to establish the reliability, accuracy and authenticity of content,[48] as well as mitigate online misinformation.[49] UK parliamentarians have also tabled amendments to the Online Safety Act through the Criminal Justice Bill effectively making the creation of deepfake intimate images an offence.[50]

Other jurisdictions

Jurisdictions like Australia and Taiwan have also made efforts to curb the impact of widespread misuse through deepfake technology. In Australia, while the Online Safety Act, 2021 **("OS Act")**[51] penalizes offences such as revenge porn (i.e., non-consensual sharing of intimate images)[52] which applies irrespective of whether an image has been altered, it does not provide any definition for deepfakes or expressly govern other diverse issues associated with it. The Australian government is planning to review the same in order to govern the deepfakes-related harms.[53] The OS Act further authorizes the eSafety Commissioner to provide a removal notice to providers of social media services, relevant electronic services, or designated internet services requiring them to remove the imager or content within a time period of 24 hours[54] and penalize them on failure.[55]

Similarly, in May 2023, the Taiwan government passed amendments to their criminal laws to prohibit pornographic deepfakes.[56] As per news reports, the amendments also state that online platforms should have an active mechanism to immediately remove illegal images and videos, and monitor technology offenders.[57] The amendments have been made to the Criminal Code of the Republic of China **("Criminal code of Taiwan")**, which lay down provisions for offences against sexual privacy and synthetic sexual videos.[58] As per the amendments, any person disseminating or broadcasting publicly sexual content of any real person synthetically generated using technological methods[59] may be penalized with imprisonment of up to seven years and a fine up to 7,000 dollars.[60] Further, these amendments also seem to require internet service providers to maintain online databases with information on crimes and criminals, personal data, and user records of suspected criminals for up to 180 days.[61]

48 Section 165 (3) of the Online safety Act.

49 Ibid.

50 Please see: https://publications.parliament.uk/pa/bills/cbill/58-04/0155/amend/criminal_rm_rep_0513.pdf at pages 146 and 147, (last accessed October 10, 2024).

51 Please see: https://www.legislation.gov.au/C2021A00076/latest/text, (last accessed October 10, 2024).

52 Section 75 (a) of the OS Act.

53 Please see: https://www.infrastructure.gov.au/sites/default/files/documents/online-safety-act-2021-review-issues-paper-26-april-2024.pdf, (last accessed October 10, 2024).

54 Section 77 (1) (f) (i) of the OS Act.

55 Section 80 of the OS Act.

56 Please see: https://focustaiwan.tw/politics/202305160024, (last accessed October 10, 2024).

57 Please see: https://opengovasia.com/taiwan-initiates-laws-to-curtail-ai-deepfake-abuse/, (last accessed October 10, 2024).

58 Chapter 28 (1) of the Criminal code of Taiwan.

59 Article 319 (4) of the Criminal code of Taiwan.

60 Ibid.

61 Please see: https://ai.taiwan.gov.tw/news/cabinet-approves-draft-amendments-to-curb-ai-powered-deepfakes/, (last accessed October 10, 2024).

B. Indian Legal and Regulatory Implications

In view of the growing use of deepfakes and associated harms, Indian lawmakers have been reportedly calling for specific laws to curb the threats posed by deepfakes.[62] Currently, there are a number of laws and regulations that regulate the use and issues associated with deepfakes. Such laws include information technology laws, penal code, privacy laws, and intellectual property laws amongst others. We have discussed and analyzed the high-level implications of each of these laws on deepfake technology. We have also discussed the various issues arising from the use of deepfakes such as privacy concerns, associated cyber-crimes, intermediary liability, impact on politics and elections, fake news dissemination, and advertising.

Information Technology Law

Certain provisions under the Information Technology Act, 2000 **("IT Act")**, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 **("IT Rules")**, and other rules notified under the IT Act regulate deepfakes through various provisions regarding intermediary liability, content blocking, penalties, etc. We have discussed the impact of such provisions on deepfakes below. Furthermore, it has been reported that the proposed Digital India Act will govern, and address issues linked with AI[63] hence having an impact on deepfakes.

Criminal Penalties

The IT Act addresses the offence of cheating by personation through the means of any computer resource[64] and imposes a punishment of imprisonment of a maximum of three years and a fine of up to INR 1,00,000. The IT act also criminalizes the transmission of obscene[65] and sexually explicit[66] content in electronic form with imprisonment of a maximum of seven years and a maximum fine of up to INR 10,00,000. Further, the IT Act also entails provisions for punishment for violating the privacy of any person, including dissemination of any intimate images of any person without their consent.[67] Such provisions and penalties thereunder may be invoked to penalize individuals involved in the creation and/or intentional circulation of deepfakes including pornographic deepfakes, such as revenge porn.

In addition to the above-discussed provisions, there is a penalty provision under the IT Act for unauthorized access to others' computer system, network, and resources.[68] Computer systems have a wide definition under the IT Act to include any device such as mobile phones and mobile applications.[69]

62 Please see: https://www.ndtv.com/india-news/deepfakes-must-be-controlled-need-law-for-this-union-minister-ashwini-vaishnaw-to-ndtv-at-indian-of-the-year-award-5297208 and https://www.hindustantimes.com/india-news/govt-may-introduce-law-against-deepfakes-misinformation-mos-101700594198089.html, (last accessed October 10, 2024).

63 Please see: https://www.business-standard.com/india-news/digital-india-act-will-deal-with-ill-effects-of-ai-mos-chandrasekhar-123112401136_1.html, (last accessed October 10, 2024).

64 Section 66D of IT Act.

65 Section 67 of IT Act.

66 Section 67 A of IT Act.

67 Section 66E of the IT Act.

68 Section 43 of IT Act.

69 Section 2(1)(i) of IT Act.

If a person intends to create deepfakes of another person by securing unauthorized access to the computer system,[70] or collecting any data including pictures or videos[71] without permission, such a person may be in violation of the IT Act. These could also be online platform providers who may create deepfakes by illegally crawling input data from the internet to create deepfakes.

Intermediary Liability

The IT Act provides a safe harbor provision[72] for online intermediaries from liability arising out of user-generated content as long as they comply with the requirements under section 79. These requirements obligate the intermediaries to (i) not initiate any transmission of online content,[73] (ii) not select or alter any information of the content transmitted by a third-party,[74] and (iii) observe mandatory due diligence[75] as prescribed by the government. It also provides exceptions to the safe harbor provisions when the intermediaries can be held liable.[76]

The IT Rules lay down a framework for due diligence and grievance redressal by the intermediaries.[77] The meaning of "intermediary" is connected to electronic records, which mean *"data, record or data generated, image or sound stored, received or sent in an electronic form or micro film or computer generated micro fiche"*.[78] Thus, content created through deepfake technology may likely fall under the ambit of electronic records. This, in turn, may result in platforms that enable content creation through deepfake technology to fall under the scope of "intermediary."

The term intermediary includes social media intermediaries **("SMI")**,[79] significant social media intermediaries **("SSMI")**,[80] and online gaming intermediaries **("OGIs")**[81] in addition to other intermediaries that merely receive, store or transmit electronic records or provide any service with respect to that record on behalf of another party. IT Rules require all intermediaries to comply with certain due diligence[82] requirements such as publishing privacy policy and user agreement on their website or application, notifying the users of any modification of such policy or agreement, taking reasonable efforts not to publish or transmit any information that belongs to other persons on which the user has no right, publishing pornographic content, and publishing content that impersonates another person among others. Further, intermediary platforms are required to ensure that no content hosted, displayed, uploaded, modified, published, transmitted, stored, updated or shared on such platform infringes any patent, trademark, copyright, or other proprietary rights.

70 Section 43(a) of IT Act.

71 Section 43(b) of IT Act.

72 Section 79(2) of the IT Act.

73 Section 79(2) (b) of the IT Act.

74 Ibid.

75 Section 79(2) (c) of the IT Act.

76 Section 79(3) of the IT Act.

77 Section 2(w) of the IT Act defines "intermediary", with respect to any particular electronic records, to mean any person who on behalf of another person receives, stores or transmits that record or provides any service with respect to that record and includes telecom service providers, network service providers, internet service providers, web-hosting service providers, search engines, online payment sites, online-auction sites, online-market places and cyber cafes.

78 Section 2(t) of the IT Act.

79 Rule 2 (1)(w) of the IT Rules.

80 Rule 2 (1)(v) of the IT Rules.

81 Rule 2 (1)(qb) of the IT Rules.

82 Rule 3(1) of the IT Rules.

The same may be ensured through reasonable efforts undertaken by the intermediary platform as well as its users.[83] Intermediaries are also required to remove or disable access to any information that is stored, hosted, or published, within 36 hours of receiving actual knowledge in the form of a court order or notification of the appropriate government (e.g., the Ministry of Home Affairs) or its agency (e.g., the Indian Cyber Crime Coordination Centre i.e., I4C).[84]

IT Rules also mandate the intermediaries to have an active grievance redressal mechanism[85] in place with contact details of the grievance officer[86] prominently published on the mobile application and/or the website of the intermediary platform. The grievance redressal mechanism is mandated with the intention to facilitate the reporting of incidents or filing of complaints by the victims. For complaints against sexually explicit content published on any intermediary platform, the IT Rules prescribe a timeline of seventy-two (72) hours within which an action must be taken regarding the resolution of such a complaint. For any other kinds of complaints, the IT Rules mandate acknowledging the complaint within twenty-four (24) hours and resolving the same within fifteen (15) days of the receipt of such complaint.

Further, the IT Rules provide additional due diligence requirements for SSMIs and OGIs[87] and require SSMIs to enable identification of the first originator of any online content.[88] Further, in case of any non-compliance under the provisions of IT Rules, the intermediaries may lose the safe harbor protection under Section 79(1) of the IT Act and in turn, may be held liable under other applicable penal provisions apart from the penal provisions under the IT Act and rules thereunder.[89]

In addition to the above, the Indian Ministry of Electronics and Information Technology **("MeitY")** had published a press release on November 7, 2023 stating that an advisory was issued to SSMIs to remove any misinformation and deepfakes within 36 hours. The Union Minister of State for Electronics & IT was also quoted to have said that this 36-hour timeline must be met upon receiving a report from any user or government authority.[90] Although, as per the IT Rules detailed above, such a timeline must be met only upon receiving a court order or notification from an appropriate government or its designated agency. MeitY privately issued an advisory to many intermediary platforms on 26 December 2023 specifically requiring them to identify and remove deepfake content. A further public advisory was issued on March 1, 2024[91] that directed the intermediary platforms to ensure that the content generated through the usage of AI technology (which may include deepfake technology) does not violate the content-related requirements laid down under the IT Rules. On March 15, 2024, MeitY issued another advisory under the IT Rules **("Advisory")**,[92] which superseded the previous advisory issued on March 1, 2024. The Advisory broadly directed the intermediary platforms to ensure that the use of AI technology does not result in violation of the requirements under IT Rules and publicly referred to the privately issued advisory dated 26 December 2023.

83 Rule 3(1)(a)(iv) of the IT Rules.

84 Rule 3(1)(d) of the IT Rules.

85 Rule 3(2) of the IT Rules.

86 Rule 2(k) of the IT Rules defines 'Grievance Officer' to mean an officer appointed by the intermediary or the publisher, as the case may be, for the purposes of the IT Rules.

87 Rule 4 of the IT Rules.

88 Rule 4(2) of the IT Rules.

89 Rule 7 of the IT Rules.

90 See: https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1975445, (last accessed September 15, 2024).

91 See: https://www.businesstoday.in/technology/news/story/government-advises-platforms-to-comply-with-it-rules-requests-submission-of-action-taken-report-to-the-ministry-within-15-days-419741-2024-03-01, (last accessed October 10, 2024).

92 See: https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf, (last accessed October 10, 2024).

In addition, the Advisory directed the intermediary platforms to ensure that under-tested or unreliable AI models must be made available to the users after appropriate labelling and implementation of consent pop-ups. The same must be done to ensure that the possibility of fallibility or unreliability of the output is communicated to the user.

More significantly, the Advisory lays down that the intermediary platforms that *"permit or facilitate synthetic creation, generation or modification of a text, audio, visual or audio-visual information, in such a manner that information may be used potentially as misinformation or deepfake,"* the content created must be labelled or embedded with a permanent unique metadata or identifier to enable identification of content created, generated or modified using computer resource of the said intermediary. Further, if any modification is made in the said content by a user, the metadata should be so configured to enable identification of such user or computer resource that has effected such change.

On September 3, 2024, a further advisory was issued by MeitY[93] following an order of the High Court of Bombay[94] requiring intermediaries to delete or disable deepfake videos relating to the National Stock Exchange of India **("NSE")**. The NSE has approached the High Court to obtain interim relief seeking directions against social media platforms to takedown and remove 2 deepfake videos carrying NSE's proprietary marks. These videos contained deepfakes of the managing direction & chief executive officer of NSE himself misleading investors to pick stocks communicated through social media platforms.

Content Blocking and Takedown

The IT Act lays down certain provisions on blocking any content on online platforms under Section 69A[95] that survived a constitution challenge where the Supreme Court upheld its constitutional validity.[96] This provision of take down of content may be applicable to deepfake content as well. Section 69A empowers the government to block the public from accessing *"any information"* on the internet when the Government believes it is *"necessary or expedient;"* and it is in the interests of the defence, sovereignty, integrity, or security of India or its relations with foreign states, public order, or the incitement of a cognizable offence relating to these categories. The Information Technology (Procedure and Safeguards for Blocking for Access of Information by Public) Rules, 2009 **("Blocking Rules")** have also been enforced which lay down the procedure of content blocking.

Under the Blocking Rules, any aggrieved person can forward their complaint for content blocking[97] that is further examined by a designated committee.[98] On examination of the complaint, the designated committee can issue notice to the person or intermediary who disseminated or published such content[99] and allow them to respond within 48 hours. Moreover, the Blocking Rules also contain an emergency provision[100] under which the hearing is not required, and the government can directly pass an order directing the content to be blocked.

93 See: https://www.meity.gov.in/writereaddata/files/Merged%20for%20MeitY%20website3.pdf, (last accessed September 15, 2024).

94 Order dated July 16, 2024 of the High Court of Bombay in National Stock Exchange of India Ltd. vs. Meta Platforms, Inc. & Ors.

95 Section 69A of the IT Act.

96 Shreya Singhal v the Union of India, (2015) 5 SCC 1.

97 Rule 6 of Blocking Rules.

98 Rule 7 of Blocking Rules.

99 Rule 8 of Blocking Rules.

100 Rule 9 of Blocking Rules.

In May 2024, shortly before the general elections in India, the Delhi High Court disposed of a PIL seeking directions for the Election Commission of India and the Union of India to establish guidelines for the use of deepfake technologies in political campaigns for the 2024 general elections to the Lok Sabha and the State Legislative Assemblies of Andhra Pradesh, Arunachal Pradesh, Odisha, and Sikkim. The Court observed that the Election Commission of India would be best placed to take action and refused to intervene.[101] The Delhi High Court is also seized with two public interest litigation **("PILs")**, one filed in December 2023[102] and another in May 2024[103] seeking directions to be issued to the government to pass regulations/ directions that regulate issues in relation to deepfakes. Hearing both PILs together, the Delhi High Court remarked that deepfakes are a 'serious menace' and only technology may offer an 'antidote'. It directed the petitioners to file affidavits containing their suggestions on how to tackle issues surrounding deepfakes.[104]

Fake news and Disinformation

The IT Rules 2021 attempt to restrict the dissemination of fake news published on intermediaries platforms. IT Rules 2021 provide for a code of ethics[105] to be followed by publishers of news and current affairs content and any online curated content that includes a self-regulatory body[106] and oversight mechanism by the central government.[107] It also authorizes the Ministry of Broadcasting to issue emergency orders to publishers or intermediaries to block any content as per the provisions of section 69A of the IT Act.[108] Intermediaries are also permitted to voluntarily remove non-compliant information either on their own or on the basis of complaints received, without jeopardizing their safe harbor under Section 79 of the IT Act, in line with its terms and conditions.[109] Further, through an amendment to the IT Rules, Rule 3(1)(b)(v) of IT Rules gave the central government the power to establish a fact check unit **("FCU")** to identify information in respect of any business of the central government as "fake or false or misleading". As per the amendment, intermediaries were required to take "reasonable efforts" to not host such information identified by the FCU.[110] Kunal Kamra, an Indian stand-up comedian and political satirist, along with several other petitioners, had challenged the constitutionality of Rule 3(1)(b)(v) of IT Rules before the Bombay High Court. While the case is pending before the Court, the Central Government proceeded with notifying the FCU under Rule 3(1)(b)(v) as the petitioners failed to make a case for the FCU notification to be stayed. The Union of India argued that "reasonable efforts" under Rule 3(1)(b) did not mandate takedown, and intermediaries could opt for a disclaimer.

Separately, the Minister of Communications and Electronics and Information technology also held consultations with industry stakeholders on issues arising out of deepfake.[111]

101 Please see: https://www.livelaw.in/high-court/delhi-high-court/deepfake-videos-during-lok-sabha-elections-delhi-high-court-256762, (last accessed October 10, 2024).

102 Chaitanya Rohilla v. Union of India, WP(C) No. 15596/2023.

103 Rajat Sharma v. Union of India, WP(C) No. 6560/2024.

104 Please see: https://www.livelaw.in/high-court/delhi-high-court/delhi-high-court-deepfake-ai-technology-267918, (last accessed September 15, 2024).

105 Rule 9 of IT Rules.

106 Rule 9(3)(a), (b) of IT Rules.

107 Rule 9(3)(c) of IT Rules.

108 Rule 16 of IT Rules.

109 Rule 3(1)(d) of IT Rules.

110 Rule 3(1)(v) of IT Rules.

111 Please see: https://pib.gov.in/PressReleasePage.aspx?PRID=1979042, (last accessed October 10, 2024).

In the discussion, it was concluded that the government, academia, social media companies, and NASSCOM will jointly work towards responding to deepfake especially addressing the issues of detection, prevention, reporting, and awareness regarding deepfake technology.

Privacy Laws

One of the major concerns linked with deepfakes is the violation of individual privacy. A notable example of such violations is in cases of revenge porn. Using publicly available pictures, audio, and videos to create deepfakes without an individual's consent also poses significant privacy-related issues. These altered media can be used to spread misinformation, damaging reputations and causing emotional distress. The dissemination of deepfakes over online platforms exacerbates these issues, as they can quickly reach a wide audience, making it challenging for victims to control or remove the content.

Emerging Privacy issues with deepfakes

There are privacy-centric emerging issues with deepfakes, such as its effect on facial recognition technology.[112] Facial recognition technology identifies the unique features of any person's face and matches it with the information stored in the database. This technology allows anyone to instantly create accurate deepfake images and videos of any person and hence pose threat to individual privacy and cybersecurity.[113] Research reports suggest that facial recognition technologies that operate on a specific user-detection technique are highly vulnerable to deepfake-based attacks that could lead to significant privacy and security concerns for the users.[114] Furthermore, for a long time, there has been a privacy and free speech debate[115] so it becomes important to address as to what extent can publicly available data be used including deepfake creation.

Right to Privacy

In a landmark judgement,[116] the Supreme Court affirmed the right to privacy as a fundamental right of right to life and dignity under Article 21 of the Indian constitution. The ruling emphasizes the concept of 'informational privacy' which states that a person must possess control over the dissemination of personal content and any unauthorized use of such information would be regarded as an infringement of informational privacy.[117] Under the existing regulatory regime, privacy is governed by certain provisions of the IT Act and the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011 **("SPDI Rules").**

112 Please see: https://blog.acer.com/en/discussion/619/the-threat-of-deepfakes-to-facial-recognition-security, (last accessed October 10, 2024).

113 Please see: https://www.sciencedirect.com/science/article/pii/S2405844023022971#:~:text=Finally%2C%20deepfakes%20might%20be%20used,be%20used%20against%20biometrics%20systems., (last accessed October 10, 2024).

114 Please see: https://www.psu.edu/news/information-sciences-and-technology/story/deepfakes-expose-vulnerabilities-certain-facial/, (last accessed October 10, 2024).

115 Please see: https://globalfreedomofexpression.columbia.edu/wp-content/uploads/2023/01/Privacy-and-Freedom-of-Expression-4.pdf , (last accessed October 10, 2024).

116 Justice K.S. Puttaswamy vs Union of India and Others, 10 SCC 1 (2017).

117 Ibid.

The IT Act prohibits any unauthorized use of the sensitive personal data of any individual and also implements reasonable security practices[118] to safeguard such information.[119] Further, the SPDI rules define sensitive personal data to include biometric data.[120] Biometrics of an individual includes face, voice patterns, fingerprints, eye retinas, and irises. As AI-generated deepfakes make use of such sensitive personal data like biometric[121] details including facial patterns and voice patterns, they would be liable for infringement of privacy under the above-stated provisions.

New Data Protection Law

In a privacy legislative overhaul, the Digital Personal Data Protection Act, 2023 **("DPDP Act")**[122] has been brought forward by the government which contemplates that the personal data of any person may be processed only when such person has given their explicit consent[123] and in case of certain legitimate uses.[124] Further, an exhaustive list of uses has been indicated that qualify as legitimate use.[125] Such legitimate uses include employment-related processing of personal data and processing of personal data when the data principal has voluntarily provided their data to the data fiduciary, among other uses. Hence, for instance, in cases where deepfake technology is deployed to create an internal-facing video by the human resources department of an organization, wherein when a data principal feeds in their personal data for the output in the video to contain personalized messages and content, such purpose may fall under legitimate uses as per the DPDP Act.

The DPDP Act also specifies hefty penalties for non-compliance with any of the provisions.[126] Since publicly available online data can be vulnerable to unauthorized creation of deepfakes, there are research reports[127] that suggest the detection of deepfakes generated using public data through advanced technological tools.

---

118 Rule 8 of the SPDI Rules.

119 Section 43A of the Information Technology Act.

120 Rule 3(vi) of the SPDI Rules.

121 Rule 2(1)(b) of the SPDI Rules define Biometrics as "the technologies that measure and analyse human body characteristics, such as 'fingerprints', 'eye retinas and irises', 'voice patterns', 'facial patterns', 'hand measurements' and 'DNA' for authentication purposes."

122 As of October 10, 2024, the DPDP Act is yet to be enacted as law.

123 Section 4(1)(a) of the DPDP Act.

124 Section 4(1)(b) of the DPDP Act.

125 Section 7 of the DPDP Act states that a Data Fiduciary may process personal data of a Data Principal for any of following uses, namely:— (a) for the specified purpose for which the Data Principal has voluntarily provided her personal data to the Data Fiduciary, and in respect of which she has not indicated to the Data Fiduciary that she does not consent to the use of her personal data. (b) for the State and any of its instrumentalities to provide or issue to the Data Principal such subsidy, benefit, service, certificate, licence or permit as may be prescribed, where–(i) she has previously consented to the processing of her personal data by the State or any of its instrumentalities for any subsidy, benefit, service, certificate, licence or permit; or (ii) such personal data is available in digital form in, or in non-digital form and digitised subsequently from, any database, register, book or other document which is maintained by the State or any of its instrumentalities and is notified by the Central Government, subject to standards followed for processing being in accordance with the policy issued by the Central Government or any law for the time being in force for governance of personal data. (c) for the performance by the State or any of its instrumentalities of any function under any law for the time being in force in India or in the interest of sovereignty and integrity of India or security of the State; (d) for fulfilling any obligation under any law for the time being in force in India on any person to disclose any information to the State or any of its instrumentalities, subject to such processing being in accordance with the provisions regarding disclosure of such information in any other law for the time being in force; (e) for compliance with any judgment or decree or order issued under any law for the time being in force in India, or any judgment or order relating to claims of a contractual or civil nature under any law for the time being in force outside India; (f) for responding to a medical emergency involving a threat to the life or immediate threat to the health of the Data Principal or any other individual; (g) for taking measures to provide medical treatment or health services to any individual during an epidemic, outbreak of disease, or any other threat to public health; (h) for taking measures to ensure safety of, or provide assistance or services to, any individual during any disaster, or any breakdown of public order. Explanation.—For the purposes of this clause, the expression "disaster" shall have the same meaning as assigned to it in clause (d) of section 2 of the Disaster Management Act, 2005; or (i) for the purposes of employment or those related to safeguarding the employer from loss or liability, such as prevention of corporate espionage, maintenance of confidentiality of trade secrets, intellectual property, classified information or provision of any service or benefit sought by a Data Principal who is an employee.

126 Section 33 of the DPDP Act.

127 Please see: https://ieeexplore.ieee.org/document/9308428, (last accessed October 10, 2024).

However, the DPDP Act does not protect publicly available information and hence, may not be applicable to deepfakes created using such data.[128]

Criminal Laws

The Bharatiya Nyaya Sanhita **("BNS")** does not expressly consider intentional and unauthorized creation and dissemination of deepfakes as a crime. However, there are certain provisions that may be applicable to deepfakes and impose criminal liability on the persons accused under the said provisions.

Defamation and Forgery

Defamation is one such offence where any visual representation intended to harm the reputation of any person may attract a penalty of two-year imprisonment or a fine or community service.[129] An instance where deepfakes are created to bring defamation to a person may come within the purview of these provisions. Further, the offence of forgery[130] includes false electronic records made to cause injury to any person and there is a separate offence for forgery with an intention to cheat.[131] The criminal law jurisprudence on this is unclear whether an AI-generated deepfake of a person can be considered as forgery or not.

Blackmailing, extortion, and cheating

Other offences associated with deepfakes are blackmailing, phishing scams, and cheating by personation. If a person intends to blackmail the victim by creating their deepfakes, especially of a sexually explicit nature, may be liable for offences of criminal intimidation[132] and extortion[133] under the BNS. Furthermore, personation becomes an important issue to be addressed since deepfakes can be created of a person which can include images, video as well as voice of that person which may be intended to deceive others. In this regard, section 319 of the BNS provides punishment for the offence of cheating by personation.

Prohibition on pornographic content

The BNS also prohibits the sale of obscene material[134] and such sales intended to minors[135] attract a higher penalty. Hence, unauthorized transmission and sale of pornographic deepfakes may attract the aforesaid provisions of the IT Act and BNS. However, it is a complex and emerging issue as to whom may the liability be imposed, whether the AI system generates such content or the person giving commands or employing tools to create such deepfake using the AI system.

128 Section 1(3)(c)(ii) states that the DPDP Act shall not apply to— (i) personal data processed by an individual for any personal or domestic purpose; and (ii) personal data that is made or caused to be made publicly available by— (A) the Data Principal to whom such personal data relates; or (B) any other person who is under an obligation under any law for the time being in force in India to make such personal data publicly available.

129 Section 356 of BNS.

130 Section 336 of BNS.

131 Section 336(3) of BNS.

132 Section 351 of BNS.

133 Section 308 of BNS.

134 Section 294 of BNS

135 Section 295 of BNS.

Deepfakes endangering the sovereignty, unity, and integrity of India

Additionally, deepfakes pertaining to specific issues or individuals of national importance may also attract liability as offences deemed to endanger the sovereignty, unity, and integrity of India under the BNS.[136] In April 2024 under the erstwhile Indian Penal Code which the BNS replaced, A case was registered by the Mumbai police against a political party's social media handle and 16 individuals for allegedly sharing a deepfake video of Union Home Minister Amit Shah. The video falsely depicted Shah announcing the reduction of reservation rights for Scheduled Castes, Scheduled Tribes, and Other Backward Classes.[137] Subsequently, in May 2024, Delhi Police suspected the national social media coordinator of a political party of having a connection with the Home Minister's deepfakes being circulated online. The FIR against such coordinator included allegations of offences under the erstwhile Indian Penal Code for criminal conspiracy, alongside initial charges of promoting enmity and forgery. The accused was in police custody for interrogation following his arrest based on a complaint from the Indian Cyber Crime Coordination Centre.[138]

Spreading of false information

The BNS penalizes any information that promotes enmity between different groups on grounds of religion, race, place of birth, residence, language, etc. or that is prejudicial to the maintenance of harmony,[139] and information that is deliberate, malicious, and is intended to outrage religious feelings of any class, by insulting its religion or religious belief[140] are also criminal offences. Further, if any person makes imputations or assertions prejudicial to national integration[141] and any statement made, rumor, or report causing public mischief and enmity, hatred or ill-will between classes[142] may also be criminally liable. However, these provisions do not address the multiplicity of other evolving categories of fake news that may be disseminated using deepfake technology for instance when accurate deepfakes of news anchors[143] of prominent news channels may be created to disseminate targeted fake news.

Deepfakes that may propagate fake news and disinformation on Cable Television

Fake news can also be disseminated using deepfake technology,[144] especially when deepfakes of prominent government officials and industry figures[145] are used. The Cable Television Network Regulation Act, 1995 **("CTNR Act")** regulates the operation of television networks and channels in India. Deepfakes may also be used to create fake news telecasted on television channels to disinform the viewers.

The CTNR Act regulates broadcasting content, by prohibiting the transmission or re-transmission of television programs and advertisements, that do not comply with the Program Code and Code for Self-Regulation of Advertising Content in India **("ASCI Code")** prescribed by the government under the Cable Television Network

136 Section 152 of BNS.

137 Please see: https://www.thehindu.com/news/national/maharashtra/amit-shah-deepfake-video-case-registered-against-maharashtra-youth-congresss-social-media-handle/article68124147.ece, (last accessed October 10, 2024).

138 Please see: https://timesofindia.indiatimes.com/city/delhi/police-add-criminal-conspiracy-charges-in-shah-fake-video-case/articleshow/109847633.cms, (last accessed October 10, 2024).

139 Section 196 of BNS.

140 Section 299 and 302 of BNS.

141 Section 197 of BNS.

142 Section 353 of the BNS.

143 Please see: https://www.forbes.com/sites/alexandralevine/2023/10/12/in-a-new-era-of-deepfakes-ai-makes-real-news-anchors-report-fake-stories/, (last accessed October 10, 2024).

144 Please see: https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html, (last accessed October 10, 2024).

145 Please see: https://www.business-standard.com/india-news/sachin-tendulkar-s-deepfake-video-sparks-concern-how-to-protect-yourself-124011500687_1.html, (last accessed October 10, 2024).

Regulation, Act, 1994 **("CTNR Rules")**.[146] The CTNR Rules prohibit programs that contain anything obscene, defamatory, deliberate, false, and suggestive innuendos and half-truths[147] or contents that criticizes, maligns, or slanders any individual in person or certain groups, segments of social, public, and moral life of the country.[148]

Copyright Laws

In September 2023, the Court in the case of *Anil Kapoor vs Simply Life India and Ors.* **("Anil Kapoor Case")** granted an ex-parte injunction in favor of Anil Kapoor, a reputed Bollywood actor, restraining the defendants, or anyone acting on their behalf from inter alia misusing the name, likeness, image, voice, and other attributes of Anil Kapoor's persona to create any merchandise, videos, photographs or other commercial purposes. In this case, Anil Kapoor had sought relief for inter alia violation of his personality rights and common laws rights including passing off, dilution, and unfair competition. He had alleged that his name and persona have immense commercial value and are liable to be protected against misuse and tarnishing. The Court, in a novel move, clarified that misuse includes technological tools such as artificial intelligence, machine learning, deep fakes, etc. to create photos, videos, etc., either for monetary gain or otherwise. This is the first instance of an Indian court granting relief for breach of personality rights specifically through the use of technology such as deepfakes.[149]

The deepfake technology has several implications under the Indian Copyright Act, 1957 **("Copyright Act")**, including protection of work created using the deepfake technology, application of fair use doctrine, and issues regarding moral and personality rights. Copyright Act content may have a twofold impact on deepfake content, wherein liability may arise at the stage (i) of inputting content for the creation of deepfakes, where usage of such content to create deepfakes may be deemed as copyright infringement, and (ii) generation of output through prompts where the content generated through deepfake technology may be infringing in nature. In February 2024, the Government of India published a notification stating that the existing intellectual property law regime including the Copyright Act is well equipped to protect AI-generated work and that there is no requirement for a new regime to regulate such works.[150]

Treatment under the Copyright Act

Under the Copyright Act, any work of visual recording and/or any sound recording qualifies for protection.[151] The Copyright Act provides that the owner of an artistic work, cinematograph film, and sound recording has the exclusive right over such work and to license the same, which also includes images forming part thereof or making any other sound recording embodying it, respectively.[152] Further, Section 51 of the Copyright Act lays down the acts that lead to copyright infringement of the protected work.[153] Therefore, any person attempting to make a deepfake of photographs or video or audio sound recordings

146 Section 5 and 6 of CTNR Act.

147 Rule 6(1)(d) of CTNR Rules.

148 Rule 6(1)(i) of CTNR Rules.

149 Anil Kapoor v Simply Life India and Others, 2023 SCC OnLine Del 6914.

150 See: https://pib.gov.in/PressReleasePage.aspx?PRID=2004715, (last accessed October 10, 2024).

151 Section 2(f) of the Copyright Act.

152 Section 14 of the Copyright Act.

153 Section 51 of the Copyright Act.

and publish any material without the proper permission and authorization of the author can be held liable for copyright infringement.

Furthermore, under the Copyright Act, the doctrine of fair use has also been provided that stipulates an exhaustive list of works that are exempt from copyright infringement.[154] Section 52(1)(a) of the Copyright Act stipulates that fair dealing with any work for private or personal use, including research, does not constitute copyright infringement. Indian courts have evaluated whether a use is protected under the fair use exception using a four-factor test: (i) the purpose and character of the use, especially if it is commercial or non-profit; (ii) the nature of the copyrighted work; (iii) the amount and substantiality of the portion used relative to the entire work; and (iv) the effect of the use on the potential market or value of the original work. The Indian Courts have also adopted the concept of transformative use by interpreting the term 'review' under Section 52(1)(a)(ii) of the Copyright Act[155] and by allowing exemption to certain specific works owing to their beneficial nature to society at large.[156] For instance, a deepfake audio/video created and published with the intention of criticizing or reviewing another work under the Copyright Act, or a deepfake audio/video created and published as a parody of an existing work with the intention of criticizing or reviewing such work may enjoy protection under Section 52 of the Copyright Act. Similarly, a deepfake audio/video that is created and published to report a current event may enjoy such protection too.

Moral Rights and Personality Rights

The Copyright Act provides that a copyrighted work must be protected from distortion, mutilation, and modification.[157] It may be noted that where deepfakes are generally modifications or alterations of any work or part thereof, the provisions on special rights of an author[158] may be applicable. Personality rights may also be infringed with the emerging issues of deepfakes, as was seen in the Anil Kapoor Case (discussed above). Earlier in a similar case, actor Amitabh Bachchan filed a case before the Delhi High Court seeking protection of his personality rights against fake calls as well as other online scams targeting individuals that used his voice from the reality show "Kaun Banega Crorepati." The actor contended that his publicity rights, as recognized previously in *Titan Industries Ltd. vs Ramkumar Jewellers*[159] were being infringed by the defendant's unauthorized use of his personality rights to promote their own goods and services. A single judge bench of the Court held that a prima facie case existed in the actor's favour. Further, the Court restrained the defendants' unauthorized use of the actor's personality rights (name, voice, image, or any other attribute exclusive to him) by granting an ex parte ad interim injunction against the same. The Court also directed the MeitY and DoT to undertake steps and pull down all the links provided by the plaintiff in his plaint, while ordering the telecom service providers to block access to the contact numbers involved in the above.

154 Section 52 of the Copyright Act.

155 University of Oxford and Ors. v Narendra Publishing and Ors., ILR (2009) 2 Del 221.

156 Super Cassettes Industries Ltd v Mr. Chintamani Rao and Ors (2011) SCC OnLine Del 4712.

157 Section 57(1)(b) of the Copyright Act.

158 Section 57 of the Copyright Act.

159 2012 SCC OnLine Del 2382.

Ownership of AI-generated content

The current copyright laws in India do not explicitly recognize AI system or tool as an author.[160] Moreover, the Parliamentary Standing Committee in its 161st Report, 2021 reviewing the intellectual property rights regime in India has recommended creating a separate category of rights for AI and related innovations.[161] Further, under Section 2(d) of the Copyright Act, 1957, an "author" includes "the person who causes the work to be created" in relation to any computer-generated literary, dramatic, musical, or artistic work. This definition raises the issue of whether artificial entities, such as AI tools, can be recognized as authors under the Copyright Act and if authorship can be attributed to entities or corporations. Identifying who "causes" the creation of a work depends on the individual's proximity and direct involvement in creating the "expression" in the content. The closer and more direct their involvement, the more likely they are to qualify as the person who caused the work to be created.

Advertising Laws

The far-reaching consequences of deepfake technology are impacting the advertising industry as well.[162] One such issue is the use of unauthorized deepfakes of celebrities in advertising, which has the potential to mislead consumers.[163] In India, Cadbury had deployed deepfake technology to create and use the deepfakes of Shahrukh Khan in their advertisement under a licensed agreement.[164] Using the deepfake technology, the actor's voice and face were modified to create personalized ads for multiple brands and stores.[165] In cases where brands use deepfakes of celebrities or other prominent industry figures in the absence of required licensed agreements or without their permission, such brand endorsements may be considered unfair trade practices and misleading advertisements. Further, in April 2024, a survey by McAfee revealed that 75% of Indians have encountered deepfake content in the past year. The study highlights increasing exposure to manipulated media, which poses significant concerns regarding misinformation and digital security. The report suggested a growing need for awareness and technological solutions to combat the spread of deepfakes.[166] There is a comprehensive set of laws that may be applicable to advertisements created through the use of deepfake technology.

160 Section 2(d) of the Copyright Act.

161 Please see: https://www.indiatoday.in/law/story/chatgpt-ai-generated-content-copyright-ownership-complexities-india-2439165-2023-09-22, (last accessed October 10, 2024).

162 Please see: https://economictimes.indiatimes.com/tech/technology/reality-check-how-deepfake-tech-is-reshaping-advertising-norms/articleshow/105207631.cms?from=mdr, (last accessed October 10, 2024).

163 Please see: https://www.wsj.com/articles/deepfakes-of-celebrities-have-begun-appearing-in-ads-with-or-without-their-permission-11666692003, (last accessed October 10, 2024).

164 Please see: https://yourstory.com/2023/10/celebrity-deepfake-indian-ads, (last accessed October 10, 2024).

165 Ibid.

166 Please see: 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, (last accessed October 10, 2024).

In addition to the same, various platforms like Meta,[167] Google,[168] Microsoft,[169] OpenAI,[170] TikTok,[171] Adobe,[172] X[173] and IBM[174] have adopted policies and implemented measures to regulate the dissemination of content created through deepfake technology.

Consumer Protection Act, 2019

The primary legislation regulating advertisements, the Consumer Protection Act, 2019 **("CPA")** lays down provisions regarding the protection of consumer rights,[175] the prevention of unfair trade practices.[176]

---

167 Please see: https://economictimes.indiatimes.com/tech/technology/meta-to-label-ai-generated-deepfake-content-as-altered-as-election-approaches/articleshow/108620343.cms?from=mdr, (last accessed October 10, 2024).

168 Please see: https://indiaai.gov.in/news/google-to-work-with-the-government-of-india-to-fight-deep-fakes, (last accessed October 10, 2024).

169 Please see: https://news.microsoft.com/en-in/microsofts-efforts-to-enhance-the-security-of-indian-elections/, (last accessed October 10, 2024).

170 Please see: https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/, (last accessed October 10, 2024).

171 Please see: https://www.nbcnews.com/tech/tech-news/tiktok-bans-deepfakes-young-people-updates-guidelines-rcna75949, (last accessed October 10, 2024).

172 Please see: https://blog.adobe.com/en/publish/2021/08/23/deepfake-task-force-danger-of-disinformation-needs-new-collaboration, (last accessed October 10, 2024).

173 Please see: https://help.x.com/en/rules-and-policies/manipulated-media, (last accessed October 10, 2024).

174 Please see: https://newsroom.ibm.com/Blog-Heres-What-Policymakers-Can-Do-About-Deepfakes#:~:text=IBM%20encourages%20policy-makers%20to%20seize,the%20global%20economy%20and%20society., (last accessed October 10, 2024).

175 Section 2(9) of CPA: "consumer rights" includes,— (i) the right to be protected against the marketing of goods, products or services which are hazardous to life and property; (ii) the right to be informed about the quality, quantity, potency, purity, standard and price of goods, products or services, as the case may be, so as to protect the consumer against unfair trade practices; (iii) the right to be assured, wherever possible, access to a variety of goods, products or services at competitive prices; (iv) the right to be heard and to be assured that consumer's interests will receive due consideration at appropriate fora; (v) the right to seek redressal against unfair trade practice or restrictive trade practices or unscrupulous exploitation of consumers; and (vi) the right to consumer awareness.

176 Section 2(46) of CPA: "unfair trade practice" means a trade practice which, for the purpose of promoting the sale, use or supply of any goods or for the provision of any service, adopts any unfair method or unfair or deceptive practice including any of the following practices, namely:—

(i) making any statement, whether orally or in writing or by visible representation including by means of electronic record, which—
  (a) falsely represents that the goods are of a particular standard, quality, quantity, grade, composition, style or model;
  (b) falsely represents that the services are of a particular standard, quality or grade;
  (c) falsely represents any re-built, second-hand, renovated, reconditioned or old goods as new goods;
  (d) represents that the goods or services have sponsorship, approval, performance, characteristics, accessories, uses or benefits which such goods or services do not have;
  (e) represents that the seller or the supplier has a sponsorship or approval or affiliation which such seller or supplier does not have;
  (f) makes a false or misleading representation concerning the need for, or the usefulness of, any goods or services;
  (g) gives to the public any warranty or guarantee of the performance, efficacy or length of life of a product or of any goods that is not based on an adequate or proper test thereof:
  Provided that where a defence is raised to the effect that such warranty or guarantee is based on adequate or proper test, the burden of proof of such defence shall lie on the person raising such defence;
  (h) makes to the public a representation in a form that purports to be—
    (A) a warranty or guarantee of a product or of any goods or services; or
    (B) a promise to replace, maintain or repair an article or any part thereof or to repeat or continue a service until it has achieved a specified result, if such purported warranty or guarantee or promise is materially misleading or if there is no reasonable prospect that such warranty, guarantee or promise will be carried out;
  (i) materially misleads the public concerning the price at which a product or like products or goods or services, have been or are, ordinarily sold or provided, and, for this purpose, a representation as to price shall be deemed to refer to the price at which the product or goods or services has or have been sold by sellers or provided by suppliers generally in the relevant market unless it is clearly specified to be the price at which the product has been sold or services have been provided by the person by whom or on whose behalf the representation is made;
  (j) gives false or misleading facts disparaging the goods, services or trade of another person.
(ii) permitting the publication of any advertisement, whether in any newspaper or otherwise, including by way of electronic record, for the sale or supply at a bargain price of goods or services that are not intended to be offered for sale or supply at the bargain price, or for a period that is, and in quantities that are, reasonable, having regard to the nature of the market in which the business is carried on, the nature and size of business,and the nature of the advertisement.
(iii) permitting—
  (a) the offering of gifts, prizes or other items with the intention of not providing them as offered or creating impression that something is being given or offered free of charge when it is fully or partly covered by the amount charged, in the transaction as a whole;
  (b) the conduct of any contest, lottery, game of chance or skill, for the purpose of promoting, directly or indirectly, the sale, use or supply of any product or any business interest, except such contest, lottery, game of chance or skill as may be prescribed;
  (c) withholding from the participants of any scheme offering gifts, prizes or other items free of charge on its closure, the information about final results of the scheme.
(iv) permitting the sale or supply of goods intended to be used, or are of a kind likely to be used by consumers, knowing or having reason to believe that the goods do not comply with the standards prescribed by the competent authority relating to performance, composition, contents, design, constructions, finishing or packaging as are necessary to prevent or reduce the risk of injury to the person using the goods;
(v) permitting the hoarding or destruction of goods, or refusal to sell the goods or to make them available for sale or to provide any service, if such hoarding or destruction or refusal raises or tends to raise or is intended to raise, the cost of those or other similar goods or services;
(vi) manufacturing of spurious goods or offering such goods for sale or adopting deceptive practices in the provision of services;
(vii) not issuing bill or cash memo or receipt for the goods sold or services rendered in such manner as may be prescribed;
(viii) refusing, after selling goods or rendering services, to take back or withdraw defective goods or to withdraw or discontinue deficient services

Further, complaints can be filed to the relevant authorities against violation of consumer rights, unfair trade practices, and false or misleading advertisements that are prejudicial to the interests of consumers.[177] Unauthorized use of deepfake technology used in advertisements may come within the scope of "misleading advertisements"[178] and "unfair trade practices" under the CPA. The CPA also makes the publication of false or misleading advertisements a criminal offence with a penalty up to 2 years of imprisonment and a fine.[179]

Guidelines on Misleading Ads and the E-commerce Rules

The Guidelines for Prevention of Misleading Advertisements and Endorsements for Misleading Advertisements, 2022 **("Guidelines on Misleading Ads")** are applicable to advertisements of all forms and on all mediums including online platforms.[180] The Guidelines on Misleading Ads primarily apply to (i) manufacturers, (ii) service providers or traders whose goods, products or services are the subject of an advertisement, (iii) advertising agencies, and (iv) endorsers who provide services for the advertisement of such goods, product or service. They provide an exhaustive list of conditions[181] for an advertisement to be categorized as misleading. Further, in the case of endorsement of advertisements, the Guidelines on Misleading Ads provide for the requirement to conduct due diligence on the authenticity and adequacy of information related to the endorsed advertisement.[182] It should be noted that any endorsement using unauthorized deepfakes can amount to contravention of this provision.

Further, the Consumer Protection (E-Commerce) Rules, 2020 **("E-commerce Rules")** provide a regulatory framework for goods and services offered over digital or electronic networks including all e-commerce models.[183] It is also applicable to all forms of unfair trade practices across all models of e-commerce[184] and prohibits e-commerce entities from employing any unfair trade practice on their platforms.[185]

Regulation of Dark Patterns, 2023

The Guidelines for Prevention and Regulation of Dark Patterns, 2023 **("Dark Patterns Guidelines")** issued under the CPA apply to platforms, advertisers, and sellers and prohibit the use of dark patterns[186] as well as specified dark patterns.[187]

---

and to refund the consideration thereof, if paid, within the period stipulated in the bill or cash memo or receipt or in the absence of such stipulation, within a period of thirty days;

(ix) disclosing to other person any personal information given in confidence by the consumer unless such disclosure is made in accordance with the provisions of any law for the time being in force

177 Section 17 of CPA.

178 Section 2(28) of CPA: "misleading advertisement" in relation to any product or service, means an advertisement, which— (i) falsely describes such product or service; or (ii) gives a false guarantee to, or is likely to mislead the consumers as to the nature, substance, quantity or quality of such product or service; or (iii) conveys an express or implied representation which, if made by the manufacturer or seller or service provider thereof, would constitute an unfair trade practice; or (iv) deliberately conceals important information;

179 Section 89 of CPA.

180 Guideline 3(a) of the Guidelines on Misleading Ads.

181 Guideline 4 of the Guidelines on Misleading Ads.

182 Guideline 13 of the Guidelines on Misleading Ads.

183 Rule 2 of E-commerce Rules.

184 Rule 2(1)(d) of E-commerce Rules.

185 Rule 4(3) of E-commerce Rules.

186 Guideline 4, Dark Patterns Guidelines.

187 Guideline 2(i) defines specified dark patterns to mean dark patterns as listed and defined in Annexure 1 and any other dark pattern that CCPA may specify from time to time or otherwise. The Dark Pattern Guidelines can be found at https://consumeraffairs.nic.in/sites/default/files/file-uploads/latestnews/Draft%20Guidelines%20for%20Prevention%20and%20Regulation%20of%20Dark%20Patterns%202023.pdf , (last accessed on June 28, 2024) wherein Annexure A provides a description of the captioned dark patterns.

The Dark Patterns Guidelines define dark patterns to mean *"any practices or deceptive design patterns using UI/UX (user interface/user experience) interactions on any platform; designed to mislead or trick users to do something they originally did not intend or want to do; by subverting or impairing the consumer autonomy, decision making or choice; amounting to misleading advertisement or unfair trade practice or violation of consumer rights."*[188]

Annexure 1 of the Dark Patterns Guidelines lists the following specified dark patterns: false urgency, basket sneaking, confirm shaming, forced action, subscription trap, interface interference, bait and switch, drip pricing, disguised advertisement, and nagging. Deepfake technology may be used to deploy such dark patterns. For instance, a deepfake video of the country's prime minister may be circulated with the intention to create false urgency within the public (e.g., regarding a widespread medical situation like a pandemic) wherein the prime minister is insisting on a purchase of a certain item. Similarly, deepfake technology may be used to create an in-app advertisement wherein a celebrity insists the users to download additional apps in order to continue using the present app, which may amount to forced action.

ASCI Code and provisions on Television Ads

The CTNR Act provides[189] that no person should transmit or re-transmit through a cable service any advertisement unless such advertisement is in conformity with the advertisement code prescribed under the CTNR Rules. The advertisement code[190] under CTNR Rules provides a detailed list of advertisements that are prohibited including advertisements in contravention of CPA and the Code for self-regulation in advertising of the Advertising Standard Council of India **("ASCI Code")**. The ASCI Code provides for safeguards against false and misleading advertisements[191] and lays down provisions to ensure fair competition in the course of advertising.[192] Compliance with the ASCI Code is mandatory[193] for cable television[194] advertisements and is not mandatory for non-members. However, ASCI suo moto investigates violations of the ASCI Code by all entities including non-member companies, and notifies the relevant sectoral regulator which subsequently initiates action.

The above-discussed framework on advertisements effectively addresses major issues in the advertising industry and may be interpreted to address certain deepfakes-related impacts on the advertising industry as well.

Regulation of Politics and Election related issues

Deepfakes can have serious consequences and compromise the integrity of the democratic process when used in elections.[195]

188 Guideline 2(e), Dark Patterns Guidelines.

189 Section 6 of CTNR Act.

190 Rule 7 of CTNR Rules.

191 Chapter I of ASCI Advertising Code.

192 Chapter IV of the ASCI Advertising Code.

193 Compliance with the ASCI Code is mandatory for advertisements on cable television under the Cable Television Network Rules, 1994. It is not legally enforceable vis-à-vis advertisements on other media. However, in the case of Common Cause (A Regd. Society) v Union of India & Ors, the Supreme Court recognized the role of ASCI in acting as an effective complaint redressal mechanism for television and radio programs.

194 Section 2(c) of the Cable Television Networks (Regulation) Act, 1995 defines "cable television network" as any system consisting of a set of closed transmission paths and associated signal generation, control and distribution equipment, designed to provide cable service for reception by multiple subscribers.

195 Please see: https://www.indiatoday.in/elections/story/how-deepfakes-could-impact-indian-elections-2464241-2023-11-17, (last accessed October 10, 2024).

Deepfakes have the potential to influence public opinion, impact election results, and disseminate inaccurate or misleading information about political candidates.[196] In April 2024, Adobe released a report that revealed that eighty-six percent of Indians believed that misinformation and harmful deepfakes would impact future elections. This concern highlighted the potential threat of manipulated media on the integrity of the electoral process.[197] Various deepfakes of politicians, as well as celebrities supporting various political parties, were widely circulated throughout India before the general elections.[198] For instance, videos created using deepfake technology featured Aamir Khan and Ranveer Singh, wherein the celebrities were seen criticizing the current prime minister and endorsing the opposition political party. Both celebrities had filed complaints regarding such deepfake videos.[199]

The Representation of People Act, 1951 prohibits certain acts which amount to "corrupt practices" with respect to elections.[200] It prohibits any information that is directed at promoting enmity on grounds of religion, race, caste, community, or language in connection with the election,[201] and deliberate and intentional statements of misinformation directed in relation to the personal character or conduct of any candidate, or in relation to the candidature, or withdrawal, of any candidate, being a statement reasonably calculated to prejudice the prospects of that candidate's election.[202]

The Election Commission of India **("ECI")** requires registered political parties and candidates must get pre-approval for all political advertisements on electronic media, including TV and social media sites, to help ensure their accuracy and fairness. It has been clarified by the ECI in its 'FAQs on Social Media'[203] that ads on websites and social media platforms come within the purview of pre-certification requirements by the respective media certification and monitoring committees. It has been stated in the FAQs that political ads issued in the e-paper of any online newspaper would require pre-certification as well. Google already has a detailed policy for political advertisements, where the advertisers are required to submit a valid pre-certificate issued by the ECI-authorized agency[204] that is verified by the Google team before making the advertisement live.[205] Further, the model code of conduct[206] released by the ECI before elections is also applicable to the content on the internet including social media. Recently, a deepfake video of a political candidate from the ruling party in India went viral during election campaigning.[207]

Further, platforms that are members of the Internet and Mobile Association of India **("IAMAI")**, a voluntary industry association, have committed to a voluntary code of ethics.

196 Please see: https://business.outlookindia.com/technology/deepfake-elections-how-indian-politicians-are-using-ai-manipulated-media-to-malign-opponents#:~:text=Election%20propaganda%20in%20India%20has,constituency%20hundreds%20of%20miles%20away, (last accessed October 10, 2024).

197 Please see: https://indianexpress.com/article/technology/tech-news-technology/misinformation-and-harmful-deepfakes-will-affect-future-elections-in-india-adobe-9357027/, (last accessed October 10, 2024).

198 Please see: https://www.bbc.com/news/world-asia-india-68918330, (last accessed October 10, 2024).

199 Please see: https://www.livemint.com/news/india/aamir-khan-ranveer-singh-deepfakes-videos-mumbai-dcp-says-investigation-on-request-people-to-not-forward-them-11713950924394.html, (last accessed October 10, 2024).

200 Section 123 of Representation of People Act, 1951.

201 Section 123(3A) of Representation of People Act, 1951.

202 Section 123(4) of Representation of People Act, 1951.

203 Please see: https://www.eci.gov.in/files/file/13753-faq-on-social-media/, (last accessed October 10, 2024).

204 Please see: https://support.google.com/adspolicy/answer/6014595?hl=en#710_def&zippy=%2Cadvertiser-verification-requirement-for-india-election-ads%2Cad-pre-certificate-requirement-for-election-ads-in-india%2Ctroubleshooter%2Cindia-election-ads, (last accessed October 10, 2024).

205 Please see: https://support.google.com/displayvideo/contact/precertificate_in?sjid=8466073435337254961-AP, (last accessed October 10, 2024).

206 https://economictimes.indiatimes.com/news/elections/lok-sabha/india/model-code-political-ad-rules-will-apply-to-social-media-too/articleshow/68350634.cms?from=mdr.

207 Please see: https://www.ndtv.com/india-news/in-bjps-deepfake-video-shared-on-whatsapp-manoj-tiwari-speaks-in-2-languages-2182923, (last accessed October 10, 2024).

This code includes adopting a system for political advertisers to submit media certification and monitoring committee certificates.
Additionally, IAMAI has pledged to promptly address paid political advertisements lawfully notified by the ECI that lack such certification. Consequently, if a platform has incorporated these mechanisms into its internal policies or industry association commitments, obtaining media certification and monitoring committee certification may be mandatory for advertising political content on that platform.

In addition to the above, BNS regulates political advertisements and penalizes anyone who, with the intent to influence election results, makes or publishes a false statement about a candidate's personal character or conduct. The punishment for this offense is a fine.[208] As discussed in the section on "Criminal Law" above, BNS also deals with the creation or circulation of statements, rumors, or reports that promote enmity, hatred, or ill will between different classes. The penalty for violating this section includes imprisonment for up to three years, a fine, or both.[209] These sections collectively aim to maintain the integrity of personal reputations, especially in the context of elections, and to prevent the spread of harmful misinformation and divisive content in society.

In May 2024, the ECI issued a letter dated May 6, 2024 to all political parties **("ECI Letter")** with the subject *"Responsible and ethical use of social media platforms and strict avoidance of any wrongful use by political parties and their representatives during MCC period in General Elections and byelections."*[210] The ECI letter stated that, in light of various legal provisions, political parties were required to adhere to certain guidelines to maintain the integrity of the electoral process. Further, it stated that political parties must not use social media platforms to disseminate any misinformation or information that is patently false, untrue, or misleading. Such information included any synthetically created, generated, or modified content that appears to be authentic but is intended to deceive. Furthermore, parties were prohibited from impersonating others, including other political parties or their representatives, and must not post or promote derogatory content towards women or anything that violates the dignity of women.

Additionally, political parties were restricted from using social media to circulate content involving violence, harm, or harassment of animals, and must not include children in any political campaigning contrary to the ECI's advisories. Deepfake audio/videos that violated existing rules and regulations were also explicitly banned. Parties were required to remove such content within a maximum of three hours of it coming to their attention and identify and warn the responsible individuals. They were also instructed to report any unlawful information and fake user accounts resembling their official handles to the relevant social media platforms immediately. If such unlawful information or fake accounts persisted after reporting, parties were to approach the Grievance Appellate Committee (GAC) under the IT Rules.

Through the ECI Letter, ECI also emphasized the need for political parties and their leaders to maintain decorum and exercise utmost restraint during public campaigning. It further reiterated that political parties and their leaders must avoid using any technological or AI-based tools that distort information or spread misinformation, as these practices undermine the standards of electioneering.

208 Section 175 of BNS.

209 Section 353 of BNS.

210 Please see: https://www.eci.gov.in/eci-backend/public/api/download?url=LMAhAK6sOPBp%2FNFF0iRfXbEB1EVSLT41NNLRjYNJJP1KivrUxbfqkDatmHy12e%2FztfbUTpXSxLP8g7dpVrk7%2FeVrNt%2BDLH%2BfDYj3Vx2GKWdqTwI8TJ87gdJ3xZOaDBMndOFtn933icz0MOeiesxvsQ%3D%3D, (last accessed October 10, 2024).

Way Forward

Deepfake technology has been evolving for nearly a decade, and its accessibility and ease of use have significantly increased over time. Today, creating convincing audio and video deepfakes has become relatively simple, raising a wide range of concerns about its potential misuse. As discussed in the paper, while there are legitimate and useful applications for deepfake technology, its misuse has garnered significant attention from governments and corporations worldwide. In response, various jurisdictions have introduced detailed legal frameworks to address the challenges posed by deepfakes, aiming to regulate their use and ensure they are employed only for legitimate purposes. Additionally, organizations across the globe have implemented measures to create a more controlled ecosystem where deepfake technology is used responsibly and ethically.

Despite these initiatives, the unique nature of deepfake technology and the rapid dissemination of digital information demand more stringent and effective measures. Establishing a clear legal definition of deepfakes, providing appropriate disclosure norms, and imposing severe penalties for their misuse may prove to be effective steps to deter potential offenders. The judiciary must also play a proactive role in interpreting these laws within the context of new and emerging technologies, setting precedents that can provide guidance in the future. Furthermore, creating a dedicated co-regulatory mechanism through government authorities and corporate entities, or a set of self-regulatory bodies could ensure the smooth implementation of these laws. Self-regulatory bodies, in particular, could leverage the initiatives already taken by globally active organizations to curb deepfake misuse, fostering a collaborative approach to building an effective regulatory mechanism.

Investing in technological solutions to detect and counteract deepfakes may prove to be crucial as well. Supporting research and development in AI and machine learning to create advanced detection tools is essential. Public-private partnerships can facilitate the development of these advanced technologies, ensuring they are accessible not just to law enforcement agencies and other stakeholders, but also to common people to equip them to not fall prey to scams. Establishing a strong centralized database of known deepfake content can aid in quicker identification and removal. Encouraging or mandating tech companies to incorporate watermarking and other verification methods for digital content can also help distinguish authentic media from manipulated ones.

Further, given the global nature of digital technologies, international collaboration may prove to be crucial to effectively combating the challenges posed by deepfakes. Establishing bilateral and multilateral agreements for information sharing and joint actions against cross-border digital crimes will strengthen the overall response to deepfake-related issues. This collaborative approach will ensure a more cohesive and effective strategy in dealing with the complexities of deepfake technology.

About NDA

At Nishith Desai Associates, we have earned the reputation of being Asia's most Innovative Law Firm — and the go-to specialists for companies around the world, looking to conduct businesses in India and for Indian companies considering business expansion abroad. In fact, we have conceptualized and created a state-of-the-art Blue Sky Thinking and Research Campus, Imaginarium Aligunjan, an international institution dedicated to designing a premeditated future with an embedded strategic foresight capability.

We are a research and strategy driven international firm with offices in Mumbai, Palo Alto (Silicon Valley), Bengaluru, Singapore, New Delhi, Munich, and New York. Our team comprises of specialists who provide strategic advice on legal, regulatory, and tax related matters in an integrated manner basis key insights carefully culled from the allied industries.

As an active participant in shaping India's regulatory environment, we at NDA, have the expertise and more importantly — the VISION — to navigate its complexities. Our ongoing endeavors in conducting and facilitating original research in emerging areas of law has helped us develop unparalleled proficiency to anticipate legal obstacles, mitigate potential risks and identify new opportunities for our clients on a global scale. Simply put, for conglomerates looking to conduct business in the subcontinent, NDA takes the uncertainty out of new frontiers.

As a firm of doyens, we pride ourselves in working with select clients within select verticals on complex matters. Our forte lies in providing innovative and strategic advice in futuristic areas of law such as those relating to Blockchain and virtual currencies, Internet of Things (IOT), Aviation, Artificial Intelligence, Privatization of Outer Space, Drones, Robotics, Virtual Reality, Ed-Tech, Med-Tech and Medical Devices and Nanotechnology with our key clientele comprising of marquee Fortune 500 corporations.

The firm has been consistently ranked as one of the Most Innovative Law Firms, across the globe. In fact, NDA has been the proud recipient of the Financial Times–RSG award 4 times in a row, (2014-2017) as the Most Innovative Indian Law Firm.

We are a trust based, non-hierarchical, democratic organization that leverages research and knowledge to deliver extraordinary value to our clients. Datum, our unique employer proposition has been developed into a global case study, aptly titled 'Management by Trust in a Democratic Enterprise,' published by John Wiley & Sons, USA.

Research@NDA

Research is the DNA of NDA. In early 1980s, our firm emerged from an extensive, and then pioneering, research by Nishith M. Desai on the taxation of cross-border transactions. The research book written by him provided the foundation for our international tax practice. Since then, we have relied upon research to be the cornerstone of our practice development. Today, research is fully ingrained in the firm's culture.

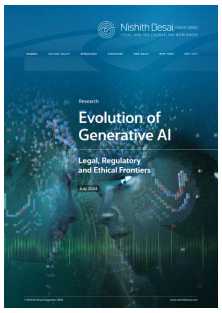
Over the years, we have produced some outstanding research papers, reports and articles. Almost on a daily basis, we analyze and offer our perspective on latest legal developments through our "Hotlines". These Hotlines provide immediate awareness and quick reference, and have been eagerly received. We also provide expanded commentary on issues through detailed articles for publication in newspapers and periodicals for dissemination to wider audience. Our NDA Labs dissect and analyze a published, distinctive legal transaction using multiple lenses and offer various perspectives, including some even overlooked by the executors of the transaction. We regularly write extensive research papers and disseminate them through our website. Our ThinkTank discourses on Taxation of eCommerce, Arbitration, and Direct Tax Code have been widely acknowledged.

As we continue to grow through our research-based approach, we now have established an exclusive four-acre, state-of-the-art research center, just a 45-minute ferry ride from Mumbai but in the middle of verdant hills of reclusive Alibaug-Raigadh district. Imaginarium AliGunjan is a platform for creative thinking; an apolitical ecosystem that connects multi-disciplinary threads of ideas, innovation and imagination. Designed to inspire 'blue sky' thinking, research, exploration and synthesis, reflections and communication, it aims to bring in wholeness — that leads to answers to the biggest challenges of our time and beyond. It seeks to be a bridge that connects the futuristic advancements of diverse disciplines. It offers a space, both virtually and literally, for integration and synthesis of knowhow and innovation from various streams and serves as a dais to internationally renowned professionals to share their expertise and experience with our associates and select clients.

We would love to hear from you about any suggestions you may have on our research publications. Please feel free to contact us at **research@nishithdesai.com**.

Other Research Papers

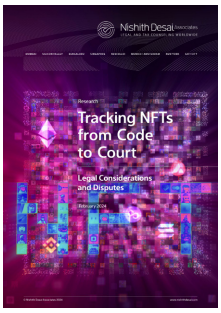
Extensive knowledge gained through our original research is a source of our expertise.



July 2024

Evolution of Generative AI

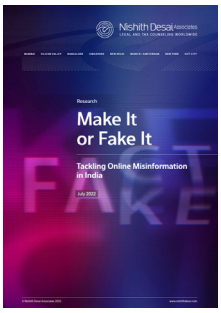
Legal, Regulatory and Ethical Frontiers



February 2024

Tracking NFTs from Code to Court

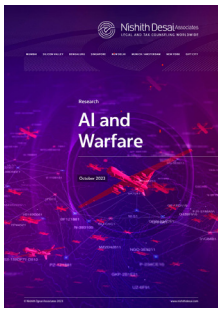
Legal Considerations and Disputes



July 2022

Make It or Fake It

Tackling Online Misinformation in India



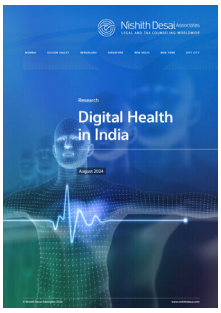
October 2023

AI and Warfare

July 2023

Cybersecurity Law and Policy

Present Scenario and the Way Forward



August 2024

Digital Health in India

For more research papers **click here**.

Nishith Desai Associates
LEGAL AND TAX COUNSELING WORLDWIDE

MUMBAI

93 B, Mittal Court, Nariman Point
Mumbai 400 021, India

Tel +91 22 6669 5000

SILICON VALLEY

220 S California Ave., Suite 201
Palo Alto, California 94306, USA

Tel +1 650 325 7100

BENGALURU

Prestige Loka, G01, 7/1 Brunton Rd
Bengaluru 560 025, India

Tel +91 80 6693 5000

SINGAPORE

Level 24, CapitaGreen
138 Market St
Singapore 048 946

Tel +65 6550 9855

MUMBAI BKC

3, North Avenue, Maker Maxity
Bandra–Kurla Complex
Mumbai 400 051, India

Tel +91 22 6159 5000

NEW DELHI

13-H, Hansalaya Building, 15
Barakhamba Road, Connaught Place
New Delhi 110 001, India

Tel +91 11 4906 5000

NEW YORK

1185 6th Avenue, Suite 326
New York, NY 10036, USA

Tel +1 212 464 7050

GIFT CITY

408, 4th Floor, Pragya Towers
GIFT City, Gandhinagar
Gujarat 382 355, India

Unmasking Deepfakes
Legal, Regulatory and Ethical Considerations

concierge@nishithdesai.com