Deepfake – How real is it?

KPMG. Make the Difference.

December 2024

kpmg.com/in

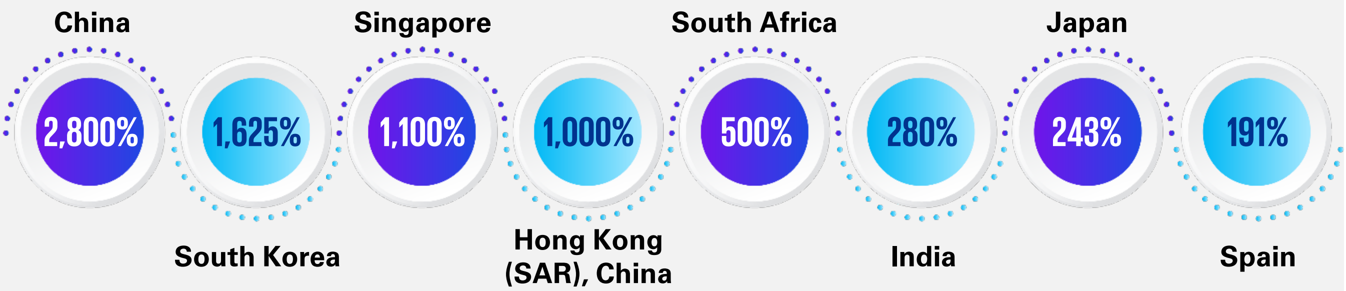
Deepfake based cyber incidents are gaining more global recognition, with a 245 per cent year-on-year in detections globally from Q1 2023 to Q1 2024.[1] Deepfake has become prevalent with the increased adoption of artificial intelligence (AI), considering deepfake refers to using AI and machine learning (ML) techniques to create hyper-realistic but entirely fabricated videos, images or audio recordings that appear authentic. The term deepfake is a combination of deep learning, a subset of AI that enables computers to learn and replicate patterns and fake, emphasising the fraudulent nature of the content. This technology uses deep learning techniques, such as generative adversarial networks (GANs), to generate synthetic data that looks like real data, by training algorithms on extensive datasets of a particular individual's images or recordings, thus producing convincing imitations of that individual's voice, facial expressions and movements.

While this technology has the potential to transform various sectors, it also poses considerable risks, particularly in terms of misinformation, privacy violations and cybercrime.

Deepfake – Reality check

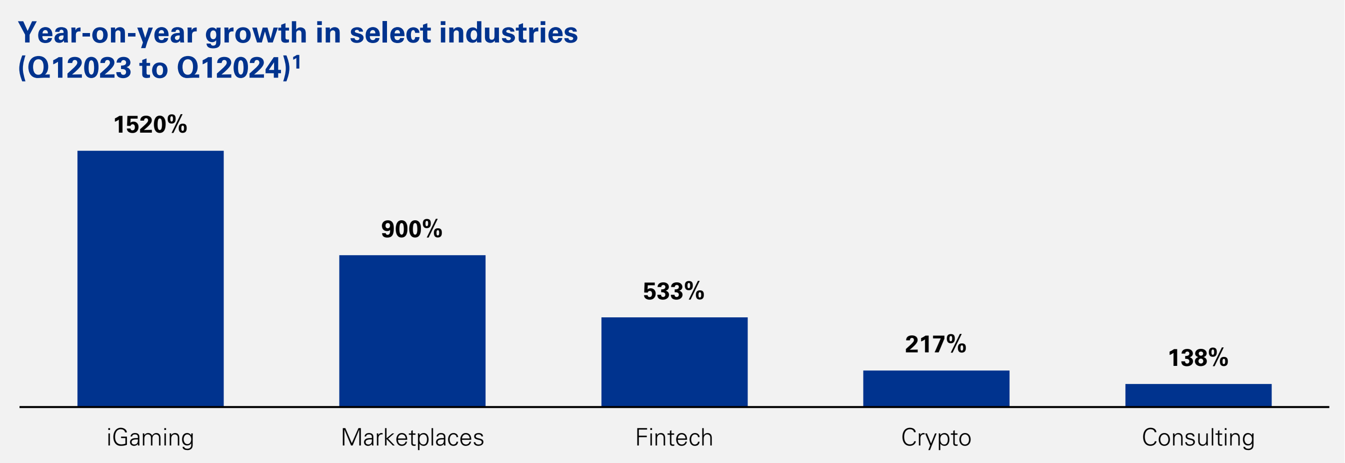
The term 'deepfake' was initially associated with the alteration of content. However, this technology has now expanded to encompass a wide range of applications, with the widespread availability of advanced computing and AI. This has led to the growth in the prevalence of deepfake, including the development of deepfake-as-a-service. Deepfake has made it increasingly feasible for individuals and entities to generate highly convincing synthetic media, often leading to the rise of trends such as political manipulation, financial fraud and privacy breaches.

Global deception: Mapping the rise of deepfakes across nations and territories year-on-year (Q12023 to Q12024)[1]:

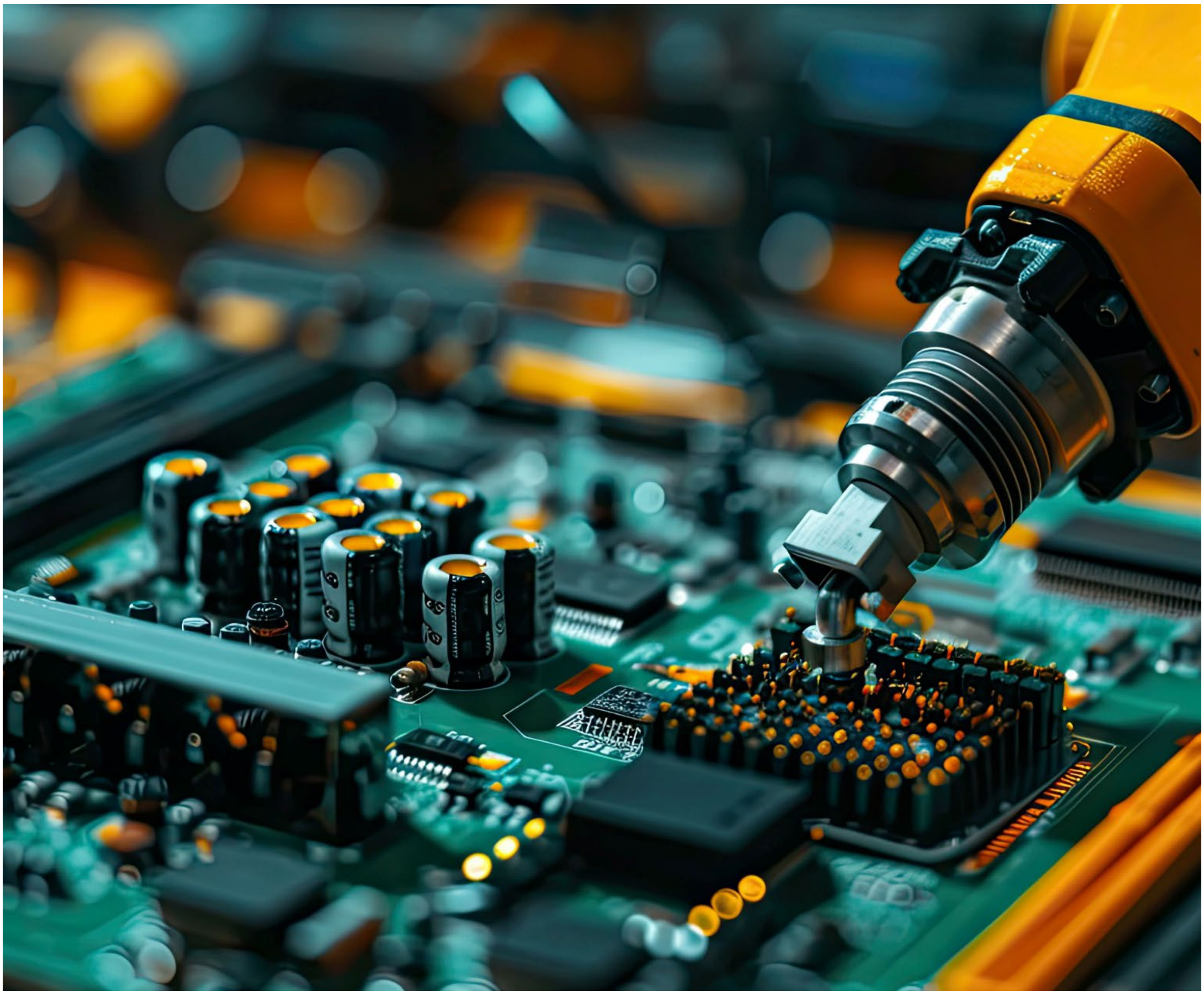


Deepfakes are growing rapidly in large and emerging economies due to several factors. Their popularity stems from the increasing availability of powerful computing resources, advancements in AI technology and the proliferation of smartphones capable of producing high-quality video footage. In India specifically, the large online population[2], heavy reliance on digital infrastructure and political events accelerates the use of deepfakes. This trend highlights the intersection of technological progress, societal engagement with digital media and the potential misuse of AI. The growth of deepfake is also observed in industries that are digitally immersive including iGaming, marketplaces, fintech and crypto, among others.

[1] Detecting Deepfakes: Fighting AI With AI, Forbes, 06 August 2024
[2] Internet usage in India - statistics & facts, Statista, 18 September 2024



Deepfake has been emerging considerably in India, over 75 per cent of Indians online have encountered some form of deepfake content in the past year, with at least 38 per cent experiencing a deepfake-related scam.[3] This highlights the widespread exposure and vulnerability of the Indian population to the misuse of this technology.



[3] 75% Indians have viewed some deepfake content in last 12 months, says McAfee survey, Economic Times, 25 April 2024

Impact of deepfake

The advent of this technology has raised global concern as it is being consistently exploited for harmful purposes, posing a risk to not only personal reputations but severely impacting societal trust and corporate integrity, emphasising the urgent need for robust safeguards.

Deepfakes pose significant threats to enterprises, society and data privacy in India. They can compromise brand integrity, damage reputations and spread misinformation. In the corporate realm, they may lead to direct financial impact and legal non-compliance. Societal impacts include erosion of trust in media and institutions. Privacy concerns arise as deepfakes can be used to manipulate personal images without consent.

The Indian government has taken steps to regulate deepfakes, but challenges remain in balancing innovation with ethical considerations. As technology advances, organisations must adapt strategies to mitigate risks associated with deepfake technologies.

A Understanding deepfake's effect on Indian corporates: The Real vs. fake dilemma

Deepfakes have several implications for Indian organisations as they pose challenges to corporate security, reputation management and operational integrity.[4]

Impact of deepfake on corporates

Reputational damages through financial frauds

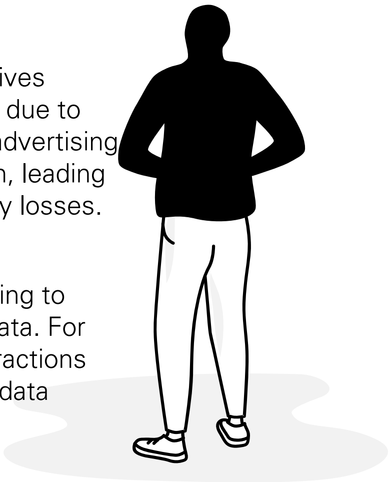
Misinformation leading to financial implications is one of the significant concerns for organisations. In the insurance sector, for instance, fraudsters create realistic incidents or damages using deepfake to trick companies into disbursing unwarranted claims.

Loss of trust

Deepfake has made spear-phishing campaigns more dangerous by enabling the creation of highly credible-looking phishing emails. India ranked as the third-largest country globally for phishing attacks after the U.S.A. and U.K., with a total of 79 million attacks in 2023.[4]

Disruptions in company operations

Impersonation of the company's high profile executives leads to potential fraud or organisational disruptions due to false instructions. For instance, the CEO of a large advertising group was the target of an elaborate deepfake scam, leading to exploitation of sensitive information and monetary losses.



Potential misuse of sensitive data

Deepfake significantly threatens cybersecurity, leading to identity theft and unauthorised access to financial data. For instance, voice-cloning technology in customer interactions can override voice-authentication systems, causing data breaches.

Employing professionals trained in AI algorithms to spot deepfakes is a crucial aspect in this direction. Cybersecurity professionals examine digital traces left during content creation or alteration to reveal signs of tampering and identify potential fraud. Further, companies must also invest in skilling initiatives through partnering with educational institutions and industry certifications and conducting internal training.

[4] India ranks third globally for phishing attacks after US, UK: Report, Business Standard, 30 April 2024

B The societal impact: Artificial or authentic

As corporations grapple with the serious repercussions of deepfake, they also face numerous challenges with significant societal impacts. Despite the World Economic Forum (WEF) identifying it as a top risk, the accessibility of user-friendly interfaces for large-scale AI models has already sparked a surge in counterfeit information and synthetic content, ranging from advanced voice cloning to fake websites.[5] The potential misuse of this technology for identity impersonation and information misrepresentation could significantly influence public opinion, thereby elevating data privacy concerns.

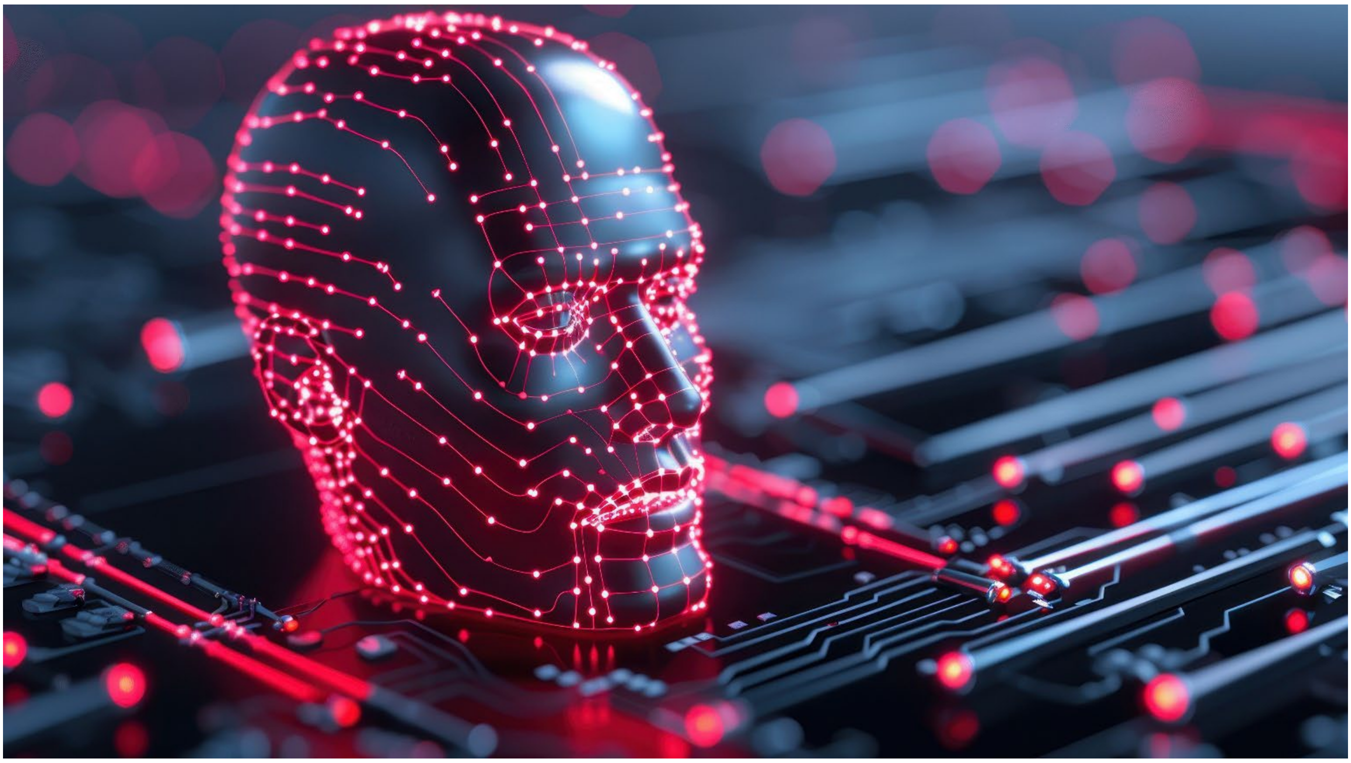
Non-consensual usage

The accessibility of personal data has facilitated the creation of non-consensual deepfakes. For instance, non-consensual deepfake pornography is a growing issue in India, where explicit videos are created using the likeness of unsuspecting individuals, predominantly women, leading to severe emotional and psychological distress.

Privacy intrusion and impersonation

Deepfake technology can also be used to impersonate individuals, leading to impersonation that may result in the breach of privacy. This was exemplified when a deepfake video of a political leader in India went viral on social media, prompting government warnings on cybersecurity.

Therefore, to combat such situations there's a need for robust digital privacy laws to protect individuals' data and prevent its unauthorised use. Additionally, awareness campaigns and necessary precautions must be taken to ensure cybersecurity.

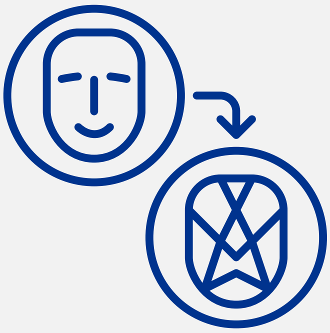


[5] The global risks report 2024, WEF, January 2024

Safeguards for this real threat

Addressing the challenges posed by deepfakes requires a strategy involving legislation, corporate governance, awareness and technological solutions for detection and prevention.

How to detect deepfake content?



Visual irregularities

Look for inconsistencies in video quality, blurred edges, unnatural blinking patterns, unmatched lip movements, or unnatural lighting.

Audio discrepancies

Unusual or inconsistent background noise may suggest manipulation.

Utilisation of technology

Employ online platforms and software that analyse videos for signs of manipulation.

Cross-verification

Validate the content with other credible sources to ensure its authenticity.

At present, India's legislation does not directly address deepfakes. However, the existing legislation against misrepresentation and spreading false information does offer some form of legal remedy. Globally, several countries are focusing on proactive and innovative solutions beyond traditional legal frameworks to combat deepfakes. Efforts in India include partnerships between tech companies and educational institutions to develop responsible AI practices and enhance public awareness about the risks associated with deepfakes and how they can impact organisations and society.

The interplay of legislation and regulatory frameworks

As the adoption and sophistication of deepfake continue to grow, appropriate frameworks and regulations must be established to mitigate potential misuse. Therefore, a comprehensive regulatory framework is needed to encourage innovation and safeguard against potential harm. This necessitates active collaboration among policymakers, industry leaders and legal experts to balance technological advancement and ethical responsibility.

The Indian government is pivotal in safeguarding the digital landscape, acting as the primary custodian of the nation's cybersecurity and digital integrity. With over 830 million[6] internet users, India is the world's largest digitally connected democracy.[2] Thus, recognising the potential threats posed by technologies such as deepfakes, the government has undertaken numerous initiatives to combat potential risks and protect the privacy and security of its citizens.

[6] MoS Rajeev Chandrasekhar to hold a Digital India Dialogues' session tomorrow in Mumbai on principles of Digital India Act, PIB, 22 May 2023

The Digital India Bill 2023[7]

- The bill is expected to positively impact corporates in several ways as it aims to create a comprehensive framework for regulating digital activities, enhancing cybersecurity measures and protecting consumer rights
- This legislation will provide organisations with clearer guidelines and reduced regulatory uncertainty, allowing them to operate more efficiently and confidently in the digital space
- Additionally, the bill's provisions on data localisation and enhanced consumer protections may open new opportunities for organisations to engage with local customers while adhering to stricter data handling standards.

Overall, the bill is likely to create a more stable and secure digital environment, benefiting organisations operating in various sectors of the Indian economy.

Digital Personal Data Protection Act (DPDPA) 2023[8]

- The act is set to significantly impact both corporates and society in India
- It will primarily govern the management of personal data collected by organisations and seeks to protect an individual's privacy by bestowing upon them rights over the processing of their data
- The act is expected to significantly affect organisations and society in India, enhancing data protection standards and granting consumers increased control over their personal information
- Organisations are set to benefit as well, from the clear guidelines on data collection and usage, which will help reduce compliance costs and legal risks.

Industry and government wide partnerships and developments

In addition to other policies including the National Data Governance Policy and amendments to the Indian Penal Code related to cybercrime, the government is considering public-private partnerships to regulate the digital landscape:

Removal of non-consensual fake content

An American multinational corporation and technology company announced new safety features designed to tackle explicit unauthorised deepfakes globally.

Addressing unethical issues surrounding AI

A U.S.A. tech major collaborated with governments around the world to combat deepfakes, fraudulent activities and other unethical issues surrounding AI.

Incorporating effective cybercrime measures

In an effort to address the rise in fake news, Karnataka government announced to form an Information Disorder Tackling Unit (IDTU) in collaboration with the Department of IT and BT.

Increased focus on customer support

An American multinational technology conglomerate dedicated a fact-checking helpline on their instant messaging app, in partnership with the Misinformation Combat Alliance (MCA) to help fight 'fake news' spread through the increasing use of AI-generated content and deepfakes in India.

[7] Proposed Digital India Act, 2023, MEITY, 09 March 2023

[8] Digital Personal Data Protection Act, 2023, KPMG India, as accessed on 3rd October 2024

Global learnings: Tackling the deepfake

EU AI Act

Taking a proactive stance, the European Union has implemented stringent measures aimed at regulating AI systems, including the realm of deepfakes, through its recent enactment of the AI Act. The main aspects of the EU AI Act regarding deepfakes are:

Transparency obligations

- Deepfake technology users must clearly state that their content is AI-generated to prevent misinformation and ensure viewers recognise its artificial origin
- Marking AI content, including deepfakes, with labels or watermarks is recommended/required.

High-risk classification

- Deepfakes used in contexts that can significantly impact individuals' rights or society (e.g., political manipulation, defamation) may be classified as high-risk and thus subject to stricter regulatory requirements.

Accountability and traceability

- Traceability and accountability in the creation and dissemination of deepfakes is to be ensured
  - This includes keeping logs of the methods and data used in creating deepfakes, which would allow authorities to trace their source if required.

Prohibited Uses

- Unacceptable risk category forbids certain harmful applications of deepfakes, such as those used for social grading or unlawful monitoring.

Hiroshima AI Process Friends Group

This initiative by Japan, supported by 54 countries and regions, including India, aims to advance cooperation for global access to safe, secure and trustworthy generative AI.[9]

The Hiroshima Process International Code of Conduct for Advanced AI Systems (HCOC) was established to align with the G7 nations' existing policies.

Enhances interoperability contribution by addressing security risks and common advanced AI risks such as inaccurate outputs, deepfakes and IP infringement.

Some other measures undertaken by developed nations include:

U.S.A.

The Disrupt Explicit Forged Images and Non-Consensual Edits (Defiance) Act of 2024, which allows victims of non-consensual sexually explicit deepfakes to sue the creators, sharers and receivers, has passed the Senate.[10]

U.K.

In April 2024, the U.K. government introduced a new law that will criminalise the creation of non-consensual sexually explicit deepfake images, punishable by an unlimited fine and potential jail time if shared.[11]

Australia

In June 2024, Australian legislation introduced the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024, which will impose serious criminal penalties for sharing non-consensual deepfake sexually explicit material to tackle gender-based violence and online abuse.[12]

Japan

In March 2024, Japan announced plans to implement laws for large-scale AI developers aimed at curbing disinformation by planning to form an AI expert council.[13]

[9] Hiroshima AI Process, SOUMU, as accessed on 28th November 2024

[10] DEFIANCE Act of 2024, Congress.gov, 23 July 2024

[11] Government cracks down on 'deepfakes' creation, GOV.UK, 16 April 2024

[12] New criminal laws to combat sexually explicit deepfakes, Attorney general's portfolio, 5 June 2024

[13] AI Guidelines for Business Ver, METI, 19 April 2024

Paving the way

The threat from deepfake is real and is envisaged to become more personalised and it is expected to manifest in multiple ways directly impacting citizens, organisations and society at large. In response to the increasing risks related to access, identity and fabricated content, as well as a broadening regulatory landscape, it is recommended that corporates adopt novel procedures, tools and tactics to enhance the security of their systems, data and infrastructure. In organisations, 73 per cent of employees expressed apprehension about the possible dangers of AI.[14] These hazards encompass misuse and harmful applications, system malfunctions, biased or incorrect results, as well as cybersecurity and privacy violations.

Countermeasures that can be undertaken by organisations

Establishing a strong cybersecurity culture and practices

- Prioritising fostering a **strong cybersecurity culture in all corporate activities** and decision-making processes, including remote work scenarios
- This involves engaging employees and emphasising the necessary behaviours to develop their cybersecurity practices. This is likely to enable them to act as human firewalls.

Policy efforts

Counteractive steps such as **mandating gen AI and large language model (LLM) providers** to embed traceability and watermarks into the deepfake creation processes before distribution.

Enhance identity verification to combat advanced forgeries

- Remove any inherent trust in network systems and constantly **verify throughout each digital interaction**
- Incorporate a zero-trust model, multi-factor authentication, behavioural biometrics, single sign-on, password management and privileged access management to fortify the identity confirmation process.

Implement security from the outset

Determine **critical risk indicators and calculate the financial implications of security threats** and risks associated with altered content.

There isn't a single, ideal solution to efficiently counter the risk of deepfake technology. Instead, a diverse strategy including investments in research and development of counter-technologies, establishment of norms and regulatory methods and increased public consciousness is needed to harness the potential of this technology responsibly.

[14] Accelerate value with confidence, KPMG India, October 2023

Acknowledgement

We are sincerely grateful to the following team members who have helped in the preparation of this report.

KPMG in India research team

- Reshma Pai (Lead - Research)
- Gagandeep Singh (Assistant Manager)
- Damini Sharda (Consultant)
- Riya Malhotra (Executive)

KPMG in India compliance and design team

- Pooja Patel (Assistant Manager)
- Angeeta Baweja (Manager)

0
10
03
17
10
0

KPMG in India contacts:

Akhilesh Tuteja
Head – Clients & Markets
Global Head – Cyber Security
E: Atuteja@kpmg com

Atul Gupta
Partner and Head – Digital Trust and Cyber
E: ATULGUPTA@kpmg.com

kpmg.com/in

Access our latest insights
on KPMG Insights Edge

Follow us on:
kpmg.com/in/socialmedia

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

KPMG Assurance and Consulting Services LLP, Lodha Excelus, Apollo Mills Compound, NM Joshi Marg, Mahalaxmi, Mumbai – 400 011 Phone: +91 22 3989 6000, Fax: +91 22 3983 6000.

This document is for e-communication only.