Combating Deep Fakes

This article is based on "Countering deep fakes, the most serious AI threat" which was published in The Hindu on 20/10/2020. It talks about the dangers associated with the Deep Fakes.

Technological advancement that has fueled media creation today, has provided opportunities for all people, regardless of their demography, ethnicity, religion etc. It can give people a voice, purpose, and ability to make an impact at scale and with speed.

However, as access to synthetic media technology increases, so does the risk of exploitation. One such aspect of media creation is deep fakes. Deep Fakes are the digital media (video, audio, and images) manipulated using Artificial Intelligence.

Deep fakes have evolved from mere annoyance to high stake warfare for creating social discord, increasing polarisation, and in some cases, influencing an election outcome. It allows hyper-realistic digital falsification that can inflict damage to individuals, institutions, businesses and democracy.

Dangers Associated With Deep Fake

Deep Fake makes it possible to fabricate media — swap faces, lip-syncing, and puppeteers — mostly without consent and bring threat to psychology, security, political stability, and business disruption.

- **New Front of Warfare:** A deepfake could act as a powerful tool by a nation-state to undermine public safety and create uncertainty and chaos in the target country.
    - Nation-state actors with geopolitical aspirations, ideological believers, violent extremists, and economically motivated enterprises can manipulate media narratives using deepfakes.
    - It can be used by insurgent groups and terrorist organisations, to represent their adversaries as making inflammatory speeches or engaging in provocative actions to stir up anti-state sentiments among people.
- **Targeting Women:** The malicious use of a deepfake can be seen in pornography, inflicting emotional, reputational, and in some cases, violence towards the individual.
    - Pornographic deep fakes can threaten, intimidate, and inflict psychological harm and reduce women to sexual objects. Deepfake pornography majorly targets women.
- **Damage to Personal Reputation**: Deepfake can depict a person indulging in antisocial behaviours and saying vile things.
    - These can have severe implications on their reputation, sabotaging their professional and personal life.
    - Even if the victim could debunk the deep fake, it may come too late to remedy the initial harm.
    - Further, Deepfakes can be deployed to extract money, confidential information, or exact favours from individuals.
- **Undermining Democracy:** A deepfake can also aid in altering the democratic discourse and undermine trust in institutions and impair diplomacy.
    - False information about institutions, public policy, and politicians powered by a deepfake can be exploited to spin the story and manipulate belief.

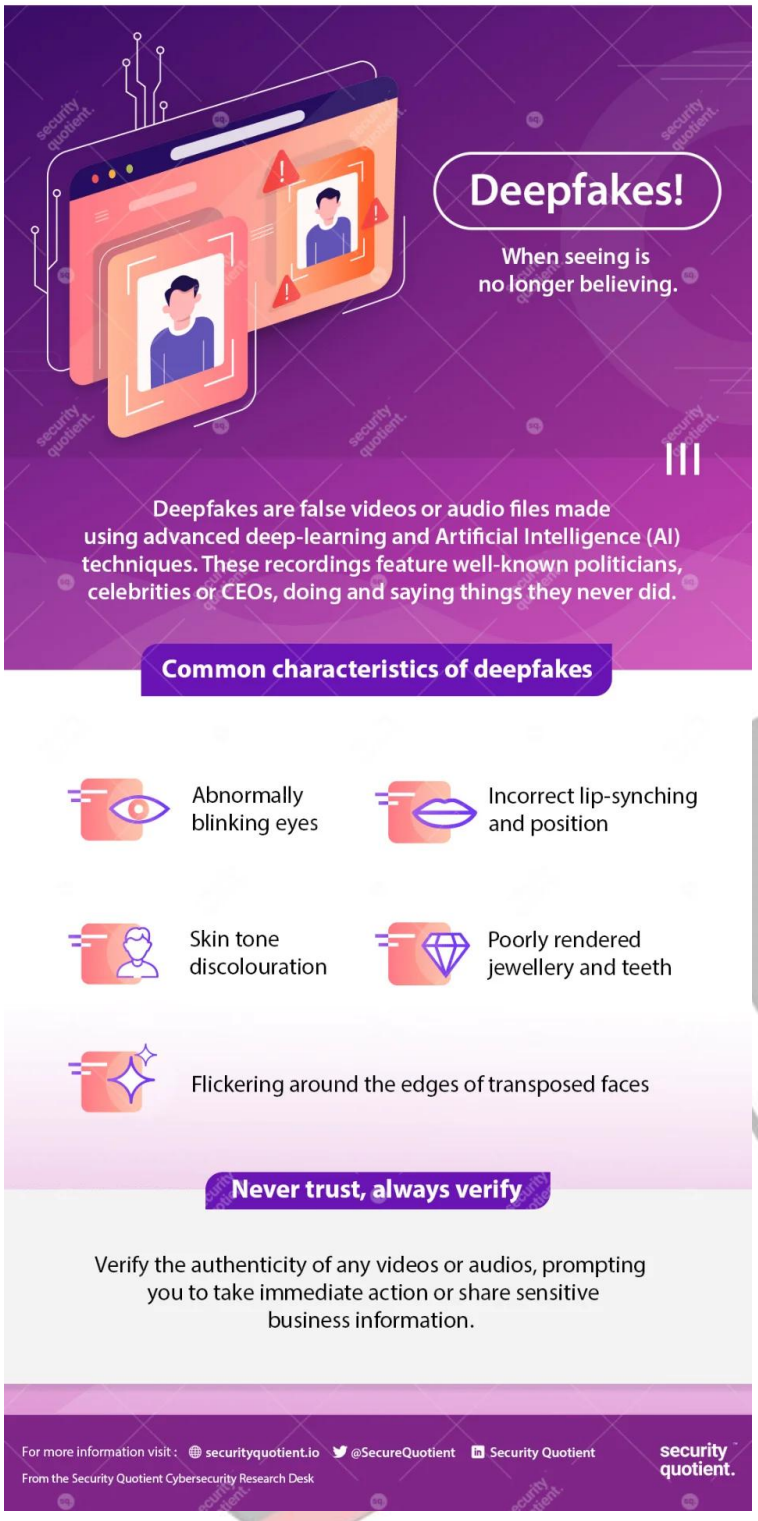
- **Disrupting Electioneering:** A deepfake of a political candidate can sabotage their image and reputation. A well-executed one, a few days before polling, of a political candidate spewing out racial epithets or indulging in an unethical act can damage their campaign.
    - A high-quality deepfake can inject compelling false information that can cast a shadow of illegitimacy over the voting process and election results.
    - Leaders can also use them to increase populism and consolidate power.
    - Deepfakes can become a very effective tool to sow the seeds of polarisation, amplifying division in society, and suppressing dissent.

Way Forward

- **Enhancing Media Literacy:** Media literacy for consumers and journalists is the most effective tool to combat disinformation and deep fakes.
    - Improving media literacy is a precursor to addressing the challenges presented by deepfakes.
    - Media literacy efforts must be enhanced to cultivate a discerning public.
    - As consumers of media, they must have the ability to decipher, understand, translate, and use the information.
    - Even a short intervention with media understanding, learning the motivations and context, can lessen the damage.
- **Need for Regulation:** Meaningful regulations with a collaborative discussion with the technology industry, civil society, and policymakers can facilitate disincentivizing the creation and distribution of malicious deep fakes.
- **Technological Interventions:** There is also a need for easy-to-use and accessible technology solutions to detect deep fakes, authenticate media, and amplify authoritative sources.
- **Behavioural Change:** On the part of society, to counter the menace of deep fakes, there is a need to take the responsibility to be a critical consumer of media on the Internet, think and pause before sharing on social media, and be part of the solution to this infodemic.

Conclusion

To defend the truth and secure freedom of expression, there is a need for a multi-stakeholder and multi-modal approach. Collaborative actions and collective techniques across legislative regulations, platform policies, technology intervention, and media literacy can provide effective and ethical countermeasures to mitigate the threat of malicious deep fakes.



Drishti Mains Question

Deep fakes allow hyper-realistic digital falsification that can inflict damage to individuals, institutions, businesse

This editorial is based on "Indo-US Ties: Beyond specific leaders" which was published in The Economic Times on October 28th,2020. Now watch this on our Youtube channel.