Abuse of AI in the Entertainment industry

Alicia Gonzalez Cruz

ITAI 2372

Prof. Anna Devarakonda

AI Application: Deepfakes

The groundbreaking power of artificial intelligence is a double-edged sword. That's nowhere more evident than in AI's capacity to generate realistic-looking images, audio and video.

What is a deepfake?

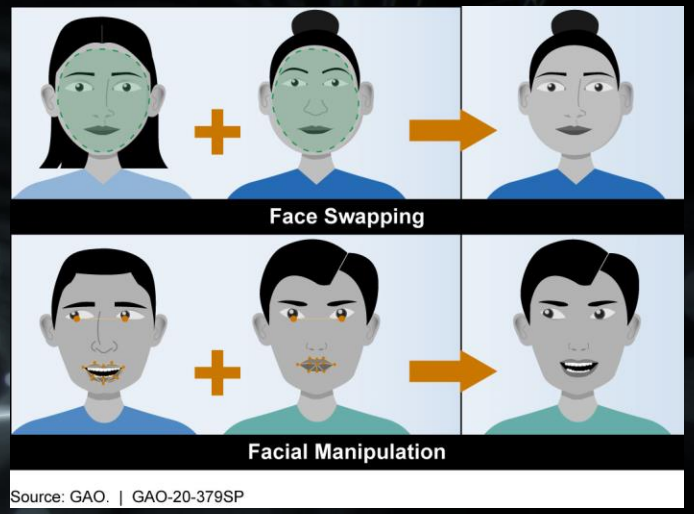
- A deepfake is an artificial image or video (a series of images) generated by a special kind of *machine learning* called "deep" learning (hence the name).

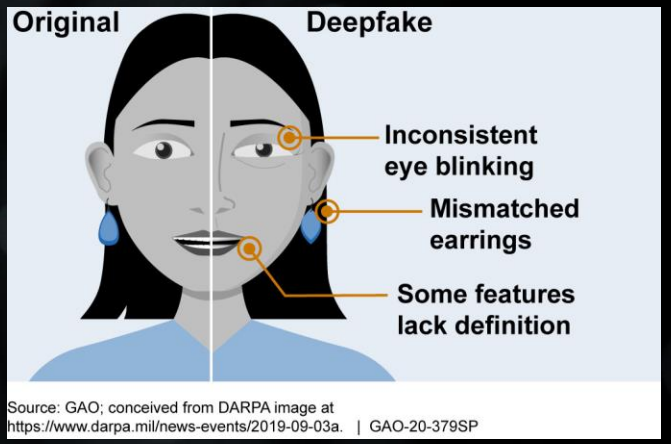
How is it made?

- A deepfake is made with machine learning. "The generator builds a training data set based on the desired output, creating the initial fake digital content, while the discriminator analyzes how realistic or fake the initial version of the content is. This process is repeated, enabling the generator to improve at creating realistic content and the discriminator to become more skilled at spotting flaws for the generator to correct."
- A deepfake can be created using a video, voice recording, image, or photo.

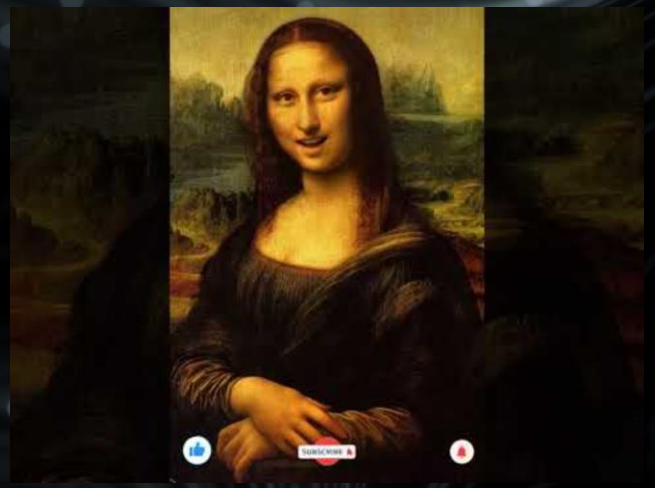
How does it work?

- Current cutting-edge deepfake AI is powered by two machine learning models working against each other. The "generator" algorithm is trained using sample imagery, audio, and/or video to create a new piece of media – or manipulate an existing one – that collectively resembles the samples as closely as possible.





Benefits



Education:

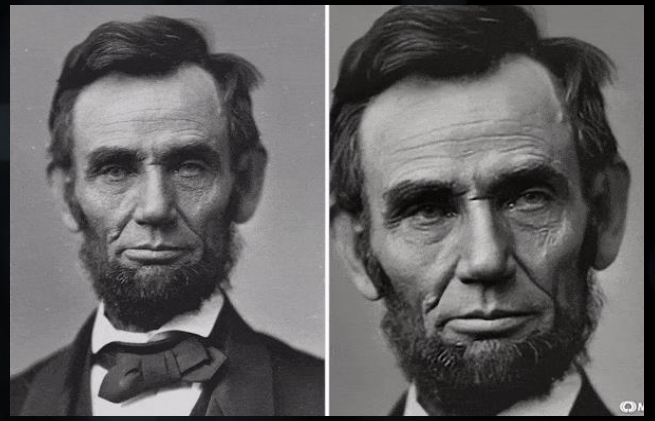
- Deepfake algorithms can animate historical photos and footage, allowing influential figures to give speeches and presentations as if they were in the classroom
- Education platforms are harnessing deepfake technology to create AI tutors that provide customized support to students

Costumer Service:

- Rather than relying solely on text-based chatbots, businesses could offer customer service through customized deepfake avatars tailored to each client.

Art:

- The examples above show how deepfakes can serve to help bring history and art 'alive' for a wider audience
- AI-generated graphics and imagery can speed up game development in the video gaming industry.



Public Safety:

- Artificial intelligence-generated synthetic media can aid in the reconstruction of a crime scene.

Challenges

Even though deepfakes bring some benefits to the community, one will argue it brings more harm than good, because of the way this type of AI technology is mostly used. "Not only has this technology created confusion, skepticism, and the spread of misinformation, deepfakes also pose a threat to privacy and security."

Threats of Deepfake

- Deepfakes blur the line between reality and fiction by creating extremely convincing fake media.
- Deepfakes are being specifically designed to exploit vulnerabilities in individuals or organizations. The increasing sophistication of deepfake technology is posing challenges for detection and debunking. As deepfakes are becoming more realistic, distinguishing between genuine and manipulated content is becoming more difficult

Use Cases:

1) A few weeks ago, AI-generated pornographic images of female students at a New Jersey high school were circulated by male classmates. A company that studies deep fakes found ninety percent of deep fake images are pornographic.

2) 2) Criminals also use this technology for deceptive practices. They collect information from emails, magazines, as well as social media posts to produce deepfakes. Sometimes, fake images and real sounds are modified and produced into deepfakes to appear more realistic to the general public.

3) 3) Russia is using AI to create deepfakes and is developing the capability to fool experts. Individuals in warzones and unstable political environments may serve as some of the highest-value targets for such deepfake malign influence.

4) 4) finance worker at a multinational firm who was tricked into paying $25 million to fraudsters after he was duped into attending a video conferencing call with whom he thought were colleagues. Everyone in the meeting apart from him was a deepfake, including the Chief Financial Officer who ordered the transfer.

Ethical and Societal Implications

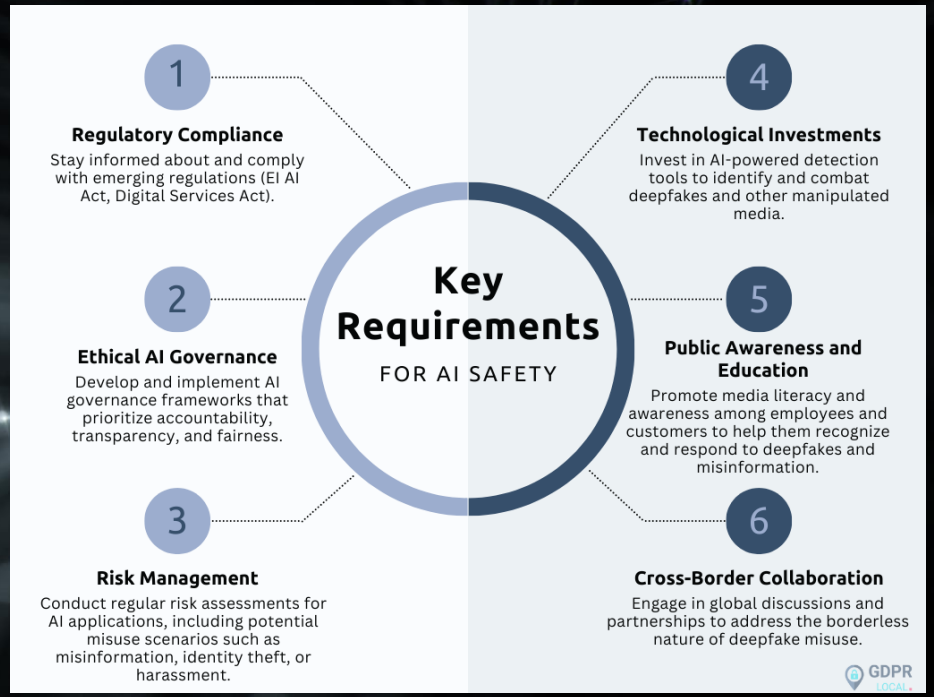
The ethical obligation of the social and technology platforms is to prevent harm. While users on these platforms have a responsibility towards sharing and consuming content, structural and informational asymmetries make it hard to expect users to play a primary role in effectively responding to malicious deepfakes. Still, platforms must do the right thing and bear the primary responsibility of identifying and preventing the spread of misleading and manipulated media.

Impact on Trust and Public Discourse

- These platforms should act to add dissemination controls or differential promotional tactics like limited sharing or downranking to stop the spread of deepfakes on their networks. Labelling content is another effective tool, which should be deployed objectively and transparently, without any political bias or business model considerations. Platforms bear ethical obligations to create and maintain the dissemination norms of their user community.

Deception and Misinformation

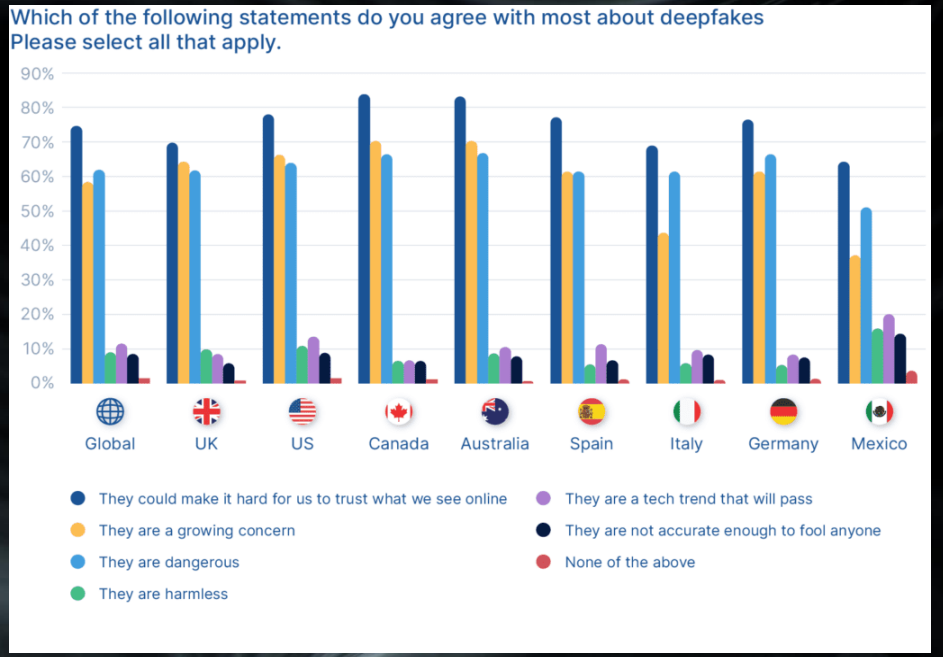
- Platforms must empower users with knowledge and critical media literacy skills to build resiliency and engage intelligently to consume, process, and share information. Practical media knowledge can enable users to think critically about the context of media and become more engaged citizens, while still appreciating satire and parody.

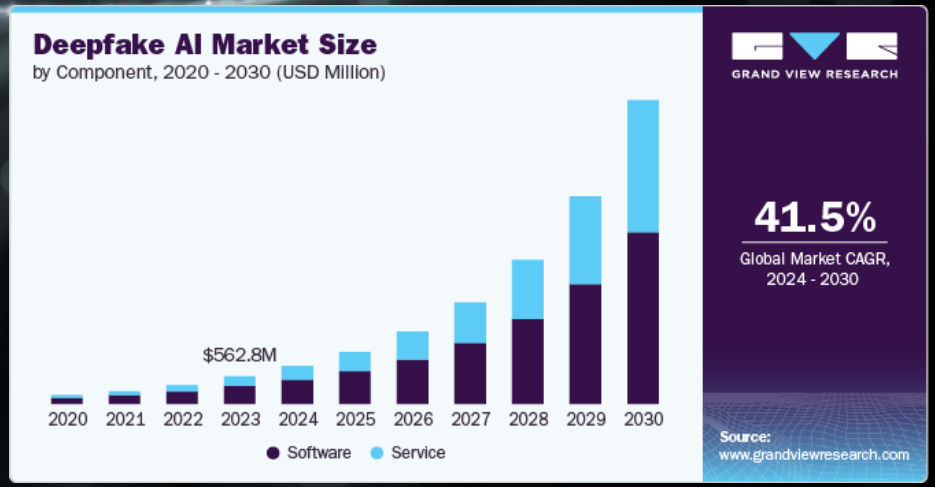


Future Directions

Recommendations

- They are requiring online platforms to detect and label content generated by AI, and for developers to **build safeguards** that prevent malicious actors from using the technology to create deepfakes.
- **Media literacy programmer:** prioritize critical thinking and equip people with the tools to verify the information they consume.
- **Biometric Implementation**
- **Enforce Consent Requirements**
- **Enhance Public Awareness and Media Literacy**
- **ransparency and Explainability**





Citations

- *Applications of Deepfake Technology: Its Benefits and Threats*. (2023a, November 3). Knowledgenile.com. https://www.knowledgenile.com/blogs/applications-of-deepfake-technology-positives-and-dangers#DigitalReconstruction&PublicSafety
- Barney, N. (2020a, October). *What is deepfake AI? A definition from WhatIs.com.* WhatIs.com. https://www.techtarget.com/whatis/definition/deepfake
- Collard, A. M. (2024a, February 12). *4 ways to future-proof against deepfakes in 2024 and beyond*. World Economic Forum. https://www.weforum.org/stories/2024/02/4-ways-to-future-proof-against-deepfakes-in-2024-and-beyond/
- *Deepfake: How the Technology Works & How to Prevent Fraud*. (n.d.-a). Www.unit21.Ai. https://www.unit21.ai/fraud-aml-dictionary/deepfake
- Diakopoulos, N., & Johnson, D. (2020a). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, *23*(7), 146144482092581. https://doi.org/10.1177/1461444820925811
- European Innovation Council and SMEs Executive Agency. (2024a, August 28). *Deepfake- A Global Crisis*. IP Helpdesk. https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/deepfake-global-crisis-2024-08-28_en
- Jaiman, A. (2020a, August 27). *Debating the ethics of deepfakes*. ORF. https://www.orfonline.org/expert-speak/debating-the-ethics-of-deepfakes
- MacDonald, A. (2022a, February). *The Uses and Abuses of Deepfake Technology*. Canadian Global Affairs Institute. https://www.cgai.ca/the_uses_and_abuses_of_deepfake_technology
- *Mace Announces Second Hearing on Deepfakes - United States House Committee on Oversight and Accountability*. (2024a, March 6). United States House Committee on Oversight and Accountability. https://oversight.house.gov/release/mace-announces-second-hearing-on-deepfakes%EF%BF%BC/
- Patterson, D. (2023b, October 5). *The positive aspect of deep fakes*. TechInformed. https://techinformed.com/deepfakes-for-good-how-synthetic-media-is-transforming-business/
- Schorr, V. (2024a, June 13). *Deepfake Dangers on Social Media*. Daon. https://www.daon.com/resource/deepfake-dangers-on-social-media/
- Stanford University. (2024a, February 22). *Dangers of Deepfake: What to Watch For | University IT.* Uit.stanford.edu. https://uit.stanford.edu/news/dangers-deepfake-what-watch
- University of Virginia. (2023a). *What the heck is a deepfake? | Information Security at UVA, U.Va.* Security.virginia.edu. https://security.virginia.edu/deepfakes