ATMOSPHERICS

07.07.23

DEEPFAKE IDENTITIES

Bottom Line Up Front:

- Just a few years ago, deepfakes were relatively easy to spot. However, as the technology progresses, it's becoming significantly more difficult to determine what is real and what is 'fake'.
- The techniques used across the industry, as well as by governments, to detect deepfakes are also progressing. However, they are not entirely reliable and can be severely limited by both talent and budget.
- The businesses within the US that have been most affected by deepfake-related fraud are IT services, online media, and FinTech. The methods commonly employed include forged ID cards and liveness bypass.
- Going forward, it will be important to recognize the impact that deepfakes can have on society beyond the 'literal' level, such as a video of someone saying something they've never said. It extends into the 'psychological' realm, involving the erosion of trust due to the sheer volume of fakes.

INFORMATION

A view of the information space related to the topic of the week, based on on headline frequency.

Deepfake Identities defined: For the uninitiated, a "deepfake" is essentially a type of digital imitation. It's a video, photo, or audio that has been altered or created using artificial intelligence, in such a way that it appears to show something that didn't really happen. Over the years there have been several ‘versions’ of deepfakes used for a wide range of purposes, some less serious (e.g., novelty, entertainment), and some with negative implications (e.g., misinformation). This week’s topic focuses on the intersection of this relatively new (and largely unexplored) technological advancement and how hackers are leveraging it to both steal and imitate one of the most personal aspects of people’s lives – their identities.

Why this topic is important right now: Let’s start with a typical hacker’s MO. Ever since personal information started becoming digitized, hackers have seized the opportunity to pursue an ever-growing honeypot, and with wide range of risk vs reward scenarios. In most cases the intent behind hacking large amount of personal data is extorsion (e.g., threaten to release the data publicly if not paid), however with advances in AI hackers can instead use data to ‘become’(through imitation) the person they’ve stolen data on. The ability to do this is possible because deepfakes have become considerably more accurate in creating a realistic digital likeness of people, allowing traditional security checks to be largely bypassed. In fact, figures as recent as last week show the number of deepfakes used in scams in Q1 2023 alone outstripped all of 2022. As more tools become available this trend in ‘synthetic fraud’ is only expected to increase.

TECHNOLOGY

AI Generated Image

" Picture a deepfake, deftly casting a politician in a fallacious scandal. Such a digital specter could swing elections and warp public discourse - all with a click and a whisper. "

- ChatGPT 2023

Deepfake identities are part of the larger, more pervasive universe of synthetic media, deepfakes employ artificial intelligence and machine learning (AI/ML) to craft convincing videos, pictures, audio, and text of events that never occurred.

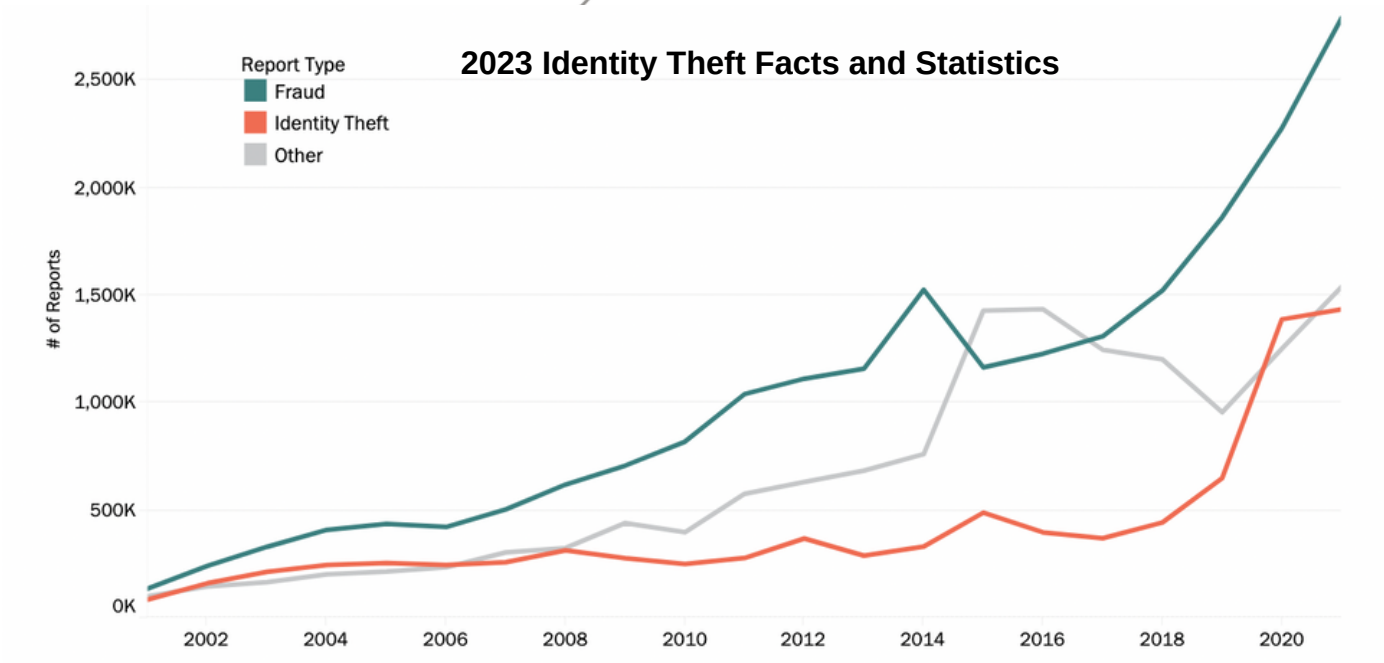
The proliferation of this technology poses a formidable challenge to the modern world. Leveraging our innate inclination to trust our senses, these deceptive creations need not be flawless to effectively spread misinformation. As deepfake technology becomes increasingly indistinguishable from reality, it presents a tool that could be manipulated to distort public opinion, jeopardize reputations, and even facilitate financial fraud.

Beyond their explicit misuse, the mere existence of deepfakes poses an insidious threat, a concept dubbed the "Liar's Dividend" by scholars Danielle K Citron and Robert Chesney. The potential for deepfake deployment can erode public trust in traditional institutions like the press, government, and academia.

Should public suspicion swing to an extreme, cultivating a default posture of disbelief and mistrust, opportunistic actors could exploit this atmosphere to further blur the lines of reality. The omnipresent risk of deepfakes can breed pervasive skepticism, undermining faith in reliable institutions and prompting questioning of the legitimacy of genuine content and media.

In our politically polarized, media-saturated world, deepfakes provide an evasive shield. Politicians caught in scandals could dismiss damaging evidence as deepfakes, thereby deflecting accountability. Moreover, savvy actors could synthetically recreate genuine events with detectable "flaws", sparking doubt about the authenticity of the original content, and, in extreme cases, the credibility of history itself. As we grapple with the implications of deepfakes, the challenge lies in retaining faith in our societal pillars amidst this potent new tool of deception.

SENTIMENT



Source: https://identitytheft.org/statistics/

Deepfakes, or synthetic media created using deep learning algorithms, have generated widespread concerns and discussions around topics like privacy, misinformation and potential misuse. The sentiment regarding deepfakes in the US can vary among different groups and individuals. Here are a few perspectives commonly discussed:

1. Concerns About Misinformation: Deepfakes have the potential to disseminate false or misleading information, prompting concerns over its impact on public trust, journalism and elections. Some view deepfakes as threats to media authenticity while also fearing their spread as misinformation.

2. Privacy and Consent: Concerns have been expressed over the use of deepfakes to create non-consensual explicit content or alter an individual's voice without their knowledge, without permission from that individual. Many individuals advocate for stricter laws and regulations to safeguard people against this kind of misuse.

3. Entertainment and Creative Potential: For some individuals, deepfakes provide both entertainment and creative expression - for instance by creating realistic special effects in movies or enriching virtual reality experiences with them. They appreciate the artistic possibilities deepfakes offer.

4. Security and Fraud: Deepfakes have raised concern among security analysts because of their potential use in fraudulent activities, including impersonating individuals for financial scams or altering video evidence in legal proceedings. This has spurred discussions around the necessity for detection and authentication technologies that can reduce risks.

INFORMATION:

1. Department of Homeland Security. (2023). Increasing threats of deepfake identities. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf
2. Srivastava, M. (2023). Fears grow of deepfake ID scams following Progress hack. Financial Times. https://www.ft.com/content/167befa0-123f-4384-a37e-c8a5b78604b2
3. Bond, S. (2023, April 27). AI-generated deepfakes are moving fast. Policymakers can't keep up National Public Radio. https://www.npr.org/2023/04/27/1172387911/how-can-people-spot-fake-images-created-by-artificial-intelligence
4. Artificial Intelligence, Deepfakes, and Disinformation A Primer. RAND Corporation. https://www.rand.org/pubs/perspectives/PEA1043-1.html
5. Deloitte. (2023). Deepfakes and AI Questioning artificial intelligence ethics and the dangers of AI. https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/deepfakes-artificial-intelligence-ethics.html
6. Congressional Research Service. (2023). Deep Fakes and National Security. https://crsreports.congress.gov/product/pdf/IF/IF11333
7. Daniel L. Byman, Chongyang Gao, Chris Meserole, and V.S. Subrahmanian. (January 2023). DEEPFAKES AND INTERNATIONAL CONFLICT. Brookings Institution. https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf
8. Neptune. (2023). Atmospherics. Retrieved from Portal.

TECHNOLOGY:

1. Midjourney. [Artwork and Images] (2023).. https://www.midjourney.com/
2. OpenAI. (2023). [ChatGPT response to prompts about deepfake identities]. https://chat.openai.com/?model=gpt-4
3. Google Bard. (2023). [Bard response to prompts about deepfake identities]. https://bard.google.com/

SENTIMENT:

1. Multiple social platforms and proprietary listening tools.

NEPTUNE

 / 415 8TH STREET SE, 2ND FLOOR, WASHINGTON, DC 20003
info@neptuneasc.com