Sang Jun Kim
12/6/2025

Key Judgments:

- **KJ1:** We assess with moderate confidence that AI-generated deepfakes will very likely increase the success rate of cyber fraud and impersonation attempts targeting financial institutions and the general public between 2025-2028,, therefore eroding public trust.
- **KJ2:** We have moderate confidence that AI-generated deepfakes will likely amplify attackers' social manipulation, enabling threat actors to distort public perception and exploit societal vulnerabilities.
- **KJ3:** Given data assess with moderate confidence that current policy against AI-generated deepfakes are not kept in pace with rapidly advancing AI tools. The gap between regulation and AI will very likely impede efforts to mitigate fraud and repair public trust due to AI-generated deepfake enabled incidents.

Scope Notes: This Intelligence report focuses on AI-generated deepfakes used in cyber operations affecting financial organizations, political communication, and public trust. It provides in-depth analysis of how AI-generated deepfakes enables impersonation, and influencing public perception through manipulated media.. This Intelligence report covers trends and incidents between 2020-2025 to evaluate policy implementation, strategic risks . This report does not address the entertainment and non-malicious purpose of AI-generated deepfakes and the technical side of deepfake creation.

Substantiation:

- **KJ1:** FBI and financial organization data shows a steady increase in fraud cases and financial losses involving AI-generated audio and video, where the attacker impersonates an authority, or a customer to gain access to sensitive information. As AI tools become more advanced and widely accessible to anyone, the number of victims and scale of financial losses continue to grow everyday. The data trends support the key judgement that deepfakes will very likely to increase the success rate of cyber enabled fraud and provide extensive erosion of public trust, perception, and authentication in digital communication.
- **KJ2:** Data from intelligence agencies, and research organizations provides that AI-generated deepfakes have already been used in political communication to influence public perception during political events or an armed conflict. Due to the AI-generated deepfakes appearing visually and audibly “real”, people are easily convinced compared to traditional fraud methods. This forces AI-generated deepfakes to be a highly effective method for attackers who target emotion and perception to intensify chaos.
- **KJ3:** According to global reports, industry assessments, and academic studies, policy development against AI-generated deepfakes has not kept pace with the advancement of AI. While many countries have implemented regulation against the rising problem of AI enabled fraud, the data shows that the effort remains limited, focusing more towards consumer protection. Outdated policies that are not as effective leaves an undeniable gap in areas of media verification, response and mitigation process against AI-generated deepfake frauds.

Perspective:

- Deepfakes has a long history of defiance in spreading misinformation, impersonation, and social engineering. Compared to traditional text scam methods the impact of deepfake creating higher risk became more natural due to societies trust in visual and audible content. In the past, the threat actors depended on fabricated news and edited images to spread misinformation, which were less challenging to detect and verify. Nonetheless, rapid advancements in Artificial Intelligence allow threat actors to create highly realistic media that manipulate consumers' perception with minimal technical skills. AI-generated deepfakes are much harder to detect and verify compared to traditional methods. AI enabled media attacks society vulnerability by exploiting consumers' emotion, environment, and biases, which becomes more effective for adversaries to use.

According to ABA, “Since 2020, deepfakes of realistic image, audio, videos that are generated by Artificial Intelligence (AI) have been rapidly evolving threat to cyber security, “According to the FBI, more than 4.2 million fraud reports have been filed since 2020, resulting in over $50.5 billion in losses, with growing portion stemming from deepfake scams” (American Bankers Association, 2025, pg.1). This provides a need for implementation of new policies around AI-generated deepfakes. Even with current policies and regulations, the risk and the losses are increasing every year from 2020 to 2025. In addition, “Deepfakes have already disrupted political landscapes by showing political figures saying or doing things they never did. One famous example of this involves Ukrainian President Volodymyr Zelenskiy in a deep fake video asking his army to cease fighting. This not only undermines the trust in political leaders but also has the potential to change public opinion and influence elections (Jacobson, 2024, pg.1)” This quote highlights that AI-generated deepfakes undeniably influence our perception of information. Even though the deepfake of President Zelenky got detected and removed, due to the nature of the fast digital world of social media proves how fast fake information can be spread and harm national security by causing psychological fear.

Outlook:

- We assess with moderate confidence that in the future, AI-generated deepfakes will be very close to being perfect, faster and easier to produce than now. For threat actors it will require little to no technique to create fake media that is very cost effective and realistic, which could widen the pool of potential new threat actors. The advancement of AI is becoming more accessible to the public, which will make it more difficult for financial organizations, government institutions, social media and policy makers to implement new regulations to detect, and verify the media. Without the implementation of advanced training and better authentication standards, many organizations will continue to struggle against deepfake incidents. Over the next several years, AI-generated deepfake will influence higher strategic risk, psychological damage, and manipulation of public trust by creating confusion.

Implication:

- We assess with moderate confidence that AI-generated deepfakes have a very serious implication for public trust, national security, and safety of political and financial organizations. As AI-generated deepfake empower fraud to be more realistic and convincing, consumers may lose confidence in traditional communication sources. This erosion of trust can affect response timing, disruption of organization, and break down the credibility of organizations that depend on public communication.
- For policymakers, the gap between advancing AI tools and regulations creates higher risk of delayed response and mitigation. Without new and updated authentication, and accountability, many organizations will face harder challenges to fight against impacts of synthetic media. To avoid an uncontrollable environment of media deception, addressing AI-generated deepfake must be prioritized and implemented through clear regulation standards with constant awareness training and verification requirements.

Source Summary Statement:

- This Intelligent Assessment is based on a mixture of government reports, industry analysis, academic sources, and case studies. These sources are from 2020-2025 and related to social manipulation, and the effect of AI-generated deepfake fraud in different fields. Most sources are reliable for identifying trends and intel on cyber fraud, impersonation, policy development, and influence on organizations. We assess with high confidence in analysis from Microsoft threat analysis, FBI, NCSL, ABA, UNESCO, SEC, and OpenAI analysis, which are reputable government and industry sources. We assess with moderate confidence in The Hacker News, CPI

Openfox, and AJG Cybersecurity article, these sources provide good context, however not academic or official sources.

Recommendations:

- Implement and expand nationwide public awareness training, policymakers should create a plan to implement public education training to school and teachers. When students from a young age can learn about the danger of manipulated intent from AI-generated deepfakes, it can help reduce the risk of successful deception from synthetic media.
- Incorporating AI-enabled verification systems to social media, before consumers can upload audio or a visual, there has to be a label that clearly establishes that the content is AI-generated. In addition, adding a hold or review process before the video can be posted can help reduce the risk of spreading misinformation at a short period of time.
- Integrating deepfake preparedness and response books into national crisis plans, emergency responders and crisis management organizations should develop an AI-generated deepfake crisis plan in case of communication disruption due to synthetic media spreading. This can reduce loss of confidence of consumers and prevent public chaos due to unreliable synthetic communication.

Sources:

Marler, J. (2025, April 14). Deepfake cybersecurity: Impacts and solutions. Cybersecurity & Compliance Solutions.
https://www.vikingcloud.com/blog/deepfake-cybersecurity#:~:text=Deepfake%20as%20a%20Defense,strategically%20to%20neutralize%20bad%20actors.

Haber, M. J. (2025, August 18). Defending against adversarial AI and Deepfake attacks. The Hacker News.
https://thehackernews.com/expert-insights/2025/08/defending-against-adversarial-ai-and. Html

OpenAI. (2024, October 1). Influence and cyber operations: An update. OpenAI.
https://cdn.openai.com/threat-intelligence-reports/influence-and-cyber-operations-an-update_October-2024.pdf

Watts, C. (2024, October 23). As the U.S. election nears, Russia, Iran and China step up influence efforts. Microsoft.
https://blogs.microsoft.com/on-the-issues/2024/10/23/as-the-u-s-election-nears-russia-iran-and-china-step-up-influence-efforts/

Federal Bureau of Investigation. (2024, December 3). Criminals use generative artificial intelligence to facilitate fraud (IC3 PSA241203). Internet Crime Complaint Center.
https://www.ic3.gov/PSA/2024/PSA241203

Jacobson, N. (2024, February 26). Deepfakes and their impact on society. CPI OpenFox.
https://www.openfox.com/deepfakes-and-their-impact-on-society/

American Bankers Association. (2025, September 3). ABA Foundation and FBI release New Infographic to help Americans spot and avoid deepfake scams. ABA Foundation and FBI Release New Infographic to Help Americans Spot and Avoid Deepfake Scams | American Bankers Association.
https://www.aba.com/about-us/press-room/press-releases/aba-foundation-and-fbi-joint-infographic-on-dee pfake-scams
NCSL. (2024, November 22). Deceptive audio or Visual Media (“deepfakes”) 2024 legislation.
https://www.ncsl.org/technology-and-communication/deceptive-audio-or-visual-media-deepfakes-2

Sang Jun Kim
12/6/2025

024-legislation

Farley, J. (2025). Deepfake Technology: The Frightening Evolution of Social Engineering. Ajg.com. https://www.ajg.com/news-and-insights/deep-fake-technology-the-frightening-evolution-of-socialengineering/

Carpenter, P. (2025, March 6). *AI, Deepfakes, and the Future of Financial Deception*. SEC.gov. https://www.sec.gov/files/carpenter-sec-statements-march2025.pdf

UNESCO. (2025, October 1). *Deepfakes and the crisis of knowing*. UNESCO.org. https://www.unesco.org/en/articles/deepfakes-and-crisis-knowing