Private Industry Notification

FEDERAL BUREAU OF INVESTIGATION, CYBER DIVISION

10 March 2021

PIN Number
210310-001

Please contact the FBI with any questions related to this Private Industry Notification at either your local **Cyber Task Force** or **FBI CyWatch**.

Local Field Offices:
www.fbi.gov/contact-us/field-offices

E-mail:
cywatch@fbi.gov

Phone:
1-855-292-3937

The following information is being provided by the FBI, with no guarantees or warranties, for potential use at the sole discretion of recipients to protect against cyber threats. This data is provided to help cyber security professionals and system administrators guard against the persistent malicious actions of cyber actors. This PIN was coordinated with DHS-CISA.

This PIN has been released **TLP:WHITE**: Subject to standard copyright rules, **TLP:WHITE** information may be distributed without restriction.

Malicious Actors Almost Certainly Will Leverage Synthetic Content for Cyber and Foreign Influence Operations

Summary

Malicious actors almost certainly will leverage synthetic content for cyber and foreign influence operations in the next 12-18 months. Foreign actors are currently using synthetic content in their influence campaigns, and the FBI anticipates it will be increasingly used by foreign and criminal cyber actors for spearphishing and social engineering in an evolution of cyber operational tradecraft.

Explaining Synthetic Content

The FBI defines synthetic content as the broad spectrum of generated or manipulated digital content, which includes images, video, audio, and text. While traditional techniques like Photoshop can be used to create synthetic content, this report highlights techniques based on artificial intelligence (AI) or machine learning (ML) technologies. These techniques are known popularly as deepfakes or GANs (generative adversarial networks). Generally, synthetic content is considered protected speech under the First Amendment. The FBI, however, may investigate malicious synthetic content which is attributed to foreign actors or is otherwise associated with criminal activities.

Recent and Anticipated Uses of Synthetic Content

Since late 2019, private sector researchers have identified multiple campaigns which have leveraged synthetic content in the form of ML-generated social media profile images. Additionally, advances in AI- and ML- based content generation and manipulation technologies likely could be used by malicious cyber actors to advance tradecraft and increase the impact of their activities. ML-generated profile images may help malicious actors spread their narratives, increasing the likelihood they will be more widely shared, making the message and messenger appear more authentic to consumers.

- Russian,[1,2] Chinese,[3] and Chinese-language[4,5] actors are using synthetic profile images derived from GANs, according to multiple private sector research reports. These profile images are associated with foreign influence campaigns, according to the same sources.

- Since 2017, unknown actors have created fictitious "journalists" who generated articles which were unwittingly published and amplified by a variety of online and print media outlets, according to press reports.[6,7] These falsified personas often have a seemingly robust online presence, including the use of GANs profile images, however, basic fact-checks can quickly reveal that the profiles are fraudulent.

Currently, individuals are more likely to encounter information online whose context has been altered by malicious actors versus fraudulent, synthesized content. This trend, however, will likely change as AL and ML technologies continue to advance.

[1] Report | Graphika | "IRA Again: Unlucky Thirteen" | 1 September 2020 | https://graphika.com/reports/ira-again-unlucky-thirteen/ | accessed on 2 September 2020.

[2] Report | Graphika | "Step into My Parler" | 1 October 2020 | https://graphika.com/reports/step-into-my-parler/ | accessed on 3 October 2020.

[3] Report | Graphika | "Operation Naval Gazing" | 22 September 2020 | https://graphika.com/reports/operation-naval-gazing/ | accessed on 23 September 2020.

[4] Report | Graphika | "Spamouflage Goes to America" | 12 August 2020 | https://graphika.com/reports/spamouflage-dragon-goes-to-america/ | accessed on 13 August 2020.

[5] Report | Graphika and DFRLab | "#OperationFFS: Fake Face Swarm" | 20 December 2019 | https://graphika.com/reports/operationffs-fake-face-swarm/ | accessed on 23 December 2019.

[6] News Article | Buzzfeed | "The Independent Used A Journalist Who Doesn't Exist On A Football Report from Cyprus." | 16 October 2017 | https://www.buzzfeed.com/markdistefano/the-independent-used-a-journalist-who-doesnt-exist-on-a | accessed on 17 February 2020.

[7] News Article | Reuters | "Deepfake used to attack activity couple shows new disinformation frontier" | 15 July 2020 | https://www.reuters.com/article/us-cyber-deepfake-activist/deepfake-used-to-attack-activist-couple-shows-new-disinformation-frontier-idUSKCN24G15E | accessed 17 February 2020.

We anticipate malicious cyber actors will use these techniques broadly across their cyber operations—likely as an extension of existing spearphishing and social engineering campaigns, but with more severe and widespread impact due to the sophistication level of the synthetic media used.

- Malicious cyber actors may use synthetic content to create highly believable spearphishing messages or engage in sophisticated social engineering attacks, according to a late 2020 joint research report.[8]

Synthetic content may also be used in a newly defined cyber attack vector referred to as Business Identity Compromise (BIC). BIC will represent an evolution in Business Email Compromise (BEC) tradecraft by leveraging advanced techniques and new tools. Whereas BEC primarily includes the compromise of corporate email accounts to conduct fraudulent financial activities, BIC will involve the use of content generation and manipulation tools to develop synthetic corporate personas or to create a sophisticated emulation of an existing employee. This emerging attack vector will likely have very significant financial and reputational impacts to victim businesses and organizations.

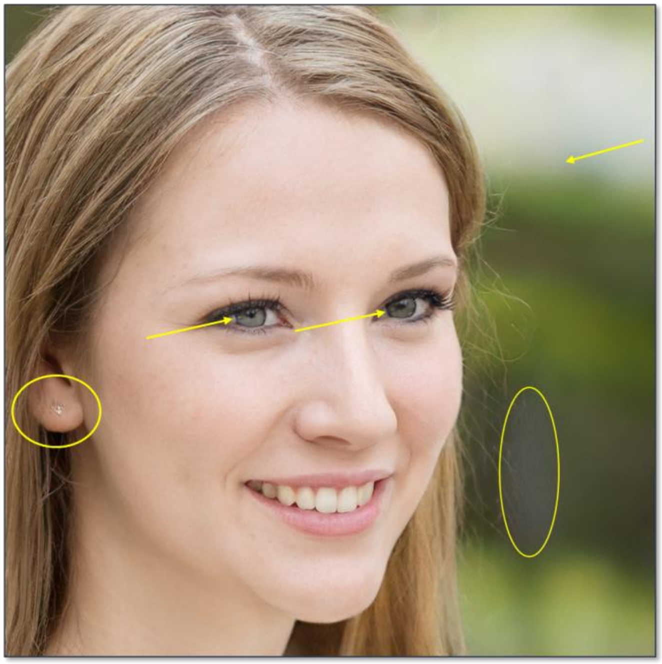
How to Identify and Mitigate Synthetic Content

Visual indicators such as distortions, warping, or inconsistencies in images and video may be an indicator of synthetic images, particularly in social media profile avatars. For example, distinct, consistent eye spacing and placement across a wide sample of synthetic images provides one indicator of synthetic content. Similar visual inconsistencies are typically present in synthetic video, often demonstrated by noticeable head and torso movements as well as syncing issues between face and lip movement, and any associated audio. Third-party research and forensic organizations, as well as some reputable cyber security companies, can aid in the identification and evaluation of suspected synthetic content. Finally, familiarity with media resiliency frameworks like the SIFT methodology can help mitigate the impact of cyber and influence operations.

- The SIFT methodology encourages individuals to **S**top, **I**nvestigate the source, **F**ind trusted coverage, and **T**race the original content when consuming information online.[9]

[8] Report | Trend Micro Research | "Malicious Uses and Abuses of Artificial Intelligence" | 19 November 2020 | https://europol.europa.eu/publications-documents/malicious-uses-and-abuses-of-artificial-intelligence/ | accessed on 20 November 2020.

[9] Blog Article | Infodemic.blog | "Sifting through the Pandemic" | 2021 | https://infodemic.blog | accessed on 17 February 2021.



Distinct visual distortions are good indicators, typically around pupils and earlobes

Indistinct and blurry backgrounds can be an indicator

Random distortions or visual artifacts are good indicators

Image source: *thispersondoesnotexist.com*

Individuals and organizations can lower the risk of becoming victim to malicious actors using synthetic content by adopting good cyber hygiene and other security measures to include the following tips.

- Be aware of the potential for cyber or foreign influence activities using synthetic content. Be alert when consuming information online, particularly when topics are especially divisive or inflammatory;
- Seek multiple, independent sources of information;
- Do not assume an online persona or individual is legitimate based on the existence of video, photographs, or audio on their profile;
- Seek media literacy or media resiliency resources like SIFT, as well as training to harden individuals and corporate interests from the potential effects of influence campaigns;
- Use multi-factor authentication on all systems, especially on shared corporate social media accounts;
- Train users to identify and report attempts at social engineering and spearphishing which may compromise personal and corporate accounts;
- Establish and exercise communications continuity plans in the event social media accounts are compromised and used to spread synthetic content;

- Do not open attachments or click links within emails received from senders you do not recognize;
- Do not provide personal information, including usernames, passwords, birth dates, social security numbers, financial data, or other information in response to unsolicited inquiries;
- Be cautious when providing sensitive personal or corporate information electronically or over the phone, particularly if unsolicited or anomalous. Confirm, if possible, requests for sensitive information through secondary channels;
- Always verify the web address of legitimate websites and manually type them into your browser.

The FBI's Protected Voices initiative provides additional tools and resources to companies, individuals, and political campaigns to protect against online foreign influence operations and cybersecurity threats.

Reporting Notice

The FBI encourages recipients of this document to report information concerning suspicious or criminal cyber activity to their local FBI field office or the FBI's 24/7 Cyber Watch (CyWatch). Additionally, corporations, individuals, or other entities who believe they are the target of foreign influence actors or other malign foreign entities are encouraged to contact their local FBI Field Office. Field office contacts can be identified at www.fbi.gov/contact-us/field-offices. CyWatch can be contacted by phone at (855) 292-3937 or by e-mail at CyWatch@fbi.gov. When available, each report submitted should include the date, time, location, type of activity, number of people, and type of equipment used for the activity, the name of the submitting company or organization, and a designated point of contact. Press inquiries should be directed to the FBI's National Press Office at NPO@fbi.gov or (202) 324-3691.

Administrative Note

This product is marked TLP:WHITE. Subject to standard copyright rules, TLP:WHITE information may be distributed without restriction.

Your Feedback Regarding this Product is Critical

Please take a few minutes to send us your feedback. Your feedback submission may be anonymous. We read each submission carefully, and your feedback will be extremely valuable to the FBI. Feedback should be specific to your experience with our written products to enable the FBI to make quick and continuous improvements to these products. Feedback may be submitted online here: https://www.ic3.gov/PIFSurvey