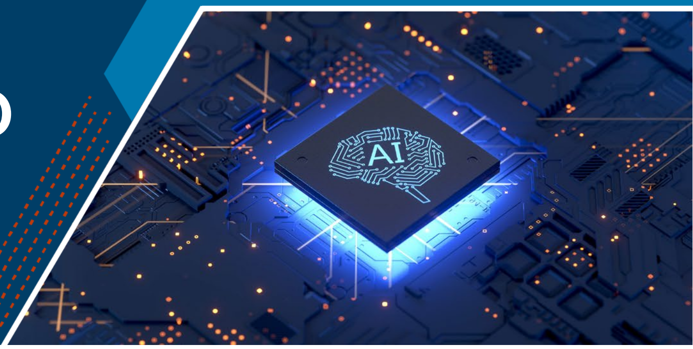
RISK IN FOCUS: GENERATIVE A.I. AND THE 2024 ELECTION CYCLE



OVERVIEW

As generative Artificial Intelligence (AI)-enabled capabilities become more widely adopted, it is important for election officials to understand how these capabilities could impact the security and integrity of election infrastructure. AI capabilities present opportunities for increased productivity, potentially enhancing both election security and election administration. However, these capabilities also carry the increased potential for greater harm, as malicious actors, including foreign nation state actors and cybercriminals, could leverage these same capabilities for nefarious purposes. For the 2024 election cycle, generative AI capabilities will likely not introduce new risks, but they may amplify existing risks to election infrastructure. Election officials have the power to mitigate against these risks heading into 2024, and many of these mitigations are the same security best practices experts have recommended taking for years. This fact sheet provides an overview of relevant generative AI-enabled capabilities, how these capabilities can be used by malicious actors to target the security and integrity of election infrastructure, and basic mitigations to counter the enhanced risks from generative AI-enabled capabilities.

GENERATIVE AI TAXONOMY

Generative AI is a type of software. It uses statistical models that generalize the patterns and structures of existing data to either reorganize existing data or create new content. This content can range from writing computer code to authoring new text to developing synthetic media such as video, image, and audio files. Examples of how malicious actors have used generative AI capabilities include:

Synthetic Content Type and Examples		Known Tactics
Video	- Text-to-Video - Deepfake[i] Video	- A foreign nation state actor uses text-to-video software to generate fake videos of real news anchors reporting on fake stories to spread disinformation as part of a foreign influence operation.[ii] - Cybercriminals use deepfake videos of famous individuals to convince the public to fall for scams.[iii]
Image	- Text-to-Image - AI-Altered Image	- Foreign nation state actors use text-to-image generators to create false and misleading images to alter public perception of facts during a crisis.[iv] - Foreign nation state actors create synthetic images for fake account profiles used in influence operations.[v] - Foreign nation state actors alter authentic images or videos to support influence operation narratives.[vi]
Audio	- Text-to-Speech - Voice Cloning	- Cybercriminals use AI-generated audio to impersonate employees and gain access to sensitive information or convince organizations to take specific actions.[vii] - Cybercriminals use generative AI tools to clone the voice of unsuspecting victims as part of AI voice scams or disinformation campaigns.[viii]
Text	- Text-to-text (Large Language Models)	- Foreign nation state actors use AI-generated text to enhance covert foreign influence operations with grammatically correct English-language content and lower marginal costs.[ix] - Cybercriminals use generative AI-enabled chatbots for sophisticated social engineering and phishing campaigns.[x]

POTENTIAL ELECTION-RELATED TARGETS OF MALICIOUS AI USE

Malicious actors can use generative AI tools to reduce the costs and increase the scale of cyber incidents and foreign influence operations. For cyber incidents, malicious actors may use generative AI to help create strains of malware that can evade cybersecurity defenses. It may also enable more effective distributed denial of service (DDoS) attacks, which can take down websites, including election-related websites, by flooding them with massive amounts of data. Malicious actors may also use generative AI to assist in deploying phishing and social engineering techniques, generating lifelike audio in any individual's voice, creating highly realistic fake images, powering counterfeit social media profiles, and producing deepfakes to support influence operation narratives. These tactics and types of attacks are not new, but generative AI allows malicious actors to employ them with greater speed and sophistication and for a much lower cost. Potential examples of malicious AI-enabled targeting of elections include:

Election Processes	▪ Chatbots, AI-generated voice, or videos could be used to spread false information about time, manner, or place of voting via text, email, social media channels, or print. ▪ Use of AI-generated content and tools could increase scale and persuasiveness of foreign influence operations and disinformation campaigns targeting election processes. ▪ AI capabilities could be used to generate convincing fake election records.
Election Offices	▪ Voice cloning tools could be used to impersonate election office staff to gain access to sensitive election administration or security information. ▪ Use of AI tools could enable higher quality spearphishing attacks against election officials or staff to gain access to sensitive information. ▪ AI coding tools could be used to develop malware and potentially even enhanced malware that could more readily evade detection systems. ▪ AI-generated scripts and voice cloning could be used to generate fake voter calls to overwhelm call centers.
Election Officials	▪ AI-generated content, such as compromising deepfake videos, could be used to harass, impersonate, or delegitimize election officials. ▪ AI tools could be used to make audio or video files impersonating election officials that spread incorrect information to the public about the security or integrity of the elections process. ▪ AI capabilities could be used to enhance public information data aggregation to enable doxing attacks against election officials.
Election Vendors	▪ AI-generated technology enables sophisticated use of phishing and social engineering techniques. ▪ AI-generated tools could be used to create a fake video of an election vendor making a false statement that calls the security of election technologies into question.

MITIGATIONS

While developments in and nefarious use of generative AI capabilities impact the risk landscape, election officials are well positioned to effectively mitigate these potential threats. Election officials are already familiar with risks like phishing and foreign influence operations and disinformation that can be amplified by generative AI. Many of the best mitigation measures for generative AI-enhanced threats are the same cybersecurity best practices that have been recommended for years and election officials may already have in place. Basic mitigation measures election subsector stakeholders can take to reduce risk from generative AI-enabled threats include:

Defend against use of sophisticated AI-enabled phishing/ social engineering techniques

Implement security controls:

- Establish and enforce strong cybersecurity protocols, like Multifactor Authentication (MFA), especially phishing-resistant MFA like Fast Identity Online (FIDO authentication), and end point detection and response software.
- Adopt email authentication security protocols such as Domain-based Message Authentication, Reporting and Conformance (DMARC), Sender Policy Framework (SPF), and DomainKeys Identified Mail (DKIM) to better guard against email spoofing.[xi]
- Harden personal and organizational social media accounts by implementing changes such as applying the strongest security and privacy controls possible, deactivating, or deleting profiles no longer in use, and removing any personally identifying information (PII) from social media profiles.
- Move toward zero trust security principles to prevent unauthorized access to data and services, and make access control enforcement as specific and detailed as possible. CISA's Zero Trust Maturity Model helps with the transition by providing a scale of recommendations for implementation.

Limit opportunities for AI-enabled impersonation and harassment

Protect yourself:

- Consider making personal social media accounts private so malicious actors have less access to your image and voice.
- Regularly request that personal information be removed from public records websites.
- Harden personal and organizational social media accounts by implementing changes such as applying the strongest security and privacy controls possible, deactivating or deleting profiles no longer in use, and removing any PII from social media profiles.
- Report instances of harassment to the appropriate authorities.

Protect sensitive information:

- Before releasing sensitive information, even internally among election staff, confirm requests through secondary channels and consider implementing identity verification for real-time communications.
- Protect against virtual impersonation attempts by adopting a rolling passphrase that only authorized personnel know, especially during active voting periods.
- Educate employees on the potential for impersonation.

Plan for AI-generated content to exceed bandwidth to respond

Proactively communicate:

- If eligible, sign up for.gov website domain to easily signal your status as a government organization.
- Ensure all official websites and social media accounts are visible and accessible to the public.
- Establish relationships with local media and community leaders; build a team of trusted voices to amplify accurate information in the event of an incident.
- Provide answers to common questions in advance through multiple information channels (e.g., online reading rooms, websites, social media, pre-recorded messages and traditional media).

Implement technical controls to limit inauthentic requests:

- Human authentication tools, such as zero trust identities, CAPTCHAs, and physical verification, serve to differentiate human users from automated processes. Implementing such tools on forms and open records requests, especially website-based submissions, can reduce the volume of inauthentic requests an office receives. Continuously review authentication tools to ensure that they are resilient to evolving capabilities, to include AI-enabled capabilities, such as by using tools that employ "AI-hardened" tasks or using hardware linked software actions such as rotating a phone.

Prepare for AI-enabled foreign influence operations & disinformation

Proactively communicate secure election practices:

- If eligible, sign up for.gov website domain to easily signal your status as a government organization.
- Establish relationships with local media and community leaders; build a team of trusted voices to amplify accurate information in the event of an incident.
- Consider utilizing active authentication techniques, such as watermarks, to mark your content as verifiably originating from you and be able to identify when files were altered after credentials were applied.
- Talk to vendors about adopting provenance and authentication measures for election-related records.
- Continue to build trust in your secure practices with proactive and responsive public communications, such as developing talking points conveying fact-based evidence of why your voters should have confidence in the security of the elections process.

Prepare for response:

- Train staff on standard procedures for responding to suspected manipulated media and understand the mechanisms for reporting this activity within your organization.
- Consider ways to authenticate election information released outside your organization (signatures, hashing, watermarks, etc.).

ADDITIONAL RESOURCES:

- Avoiding Social Engineering and Phishing Attacks | CISA
- AI Risk Management Framework | NIST
- EI-ISAC membership
- Disinformation Stops with You
- Tactics of Disinformation
- Election Disinformation Toolkit
- Election Cybersecurity Toolkit
- Cyber Incident Detection and Notification Planning Guide for Election Security
- Election Security Services
- Building Trust Through Secure Practices
- Contextualizing Deepfake Threats to Organizations
- No Downtime in Elections: A Guide to Mitigating Risks of Denial-of-Service

[i] Deepfakes are a type of synthetic media—commonly generated using artificial intelligence/machine learning (AI/ML)—presenting plausible and realistic videos, pictures, audio or text of events which never happened.

[ii] How Deepfake Videos Are Used to Spread Disinformation - The New York Times (nytimes.com); Deepfake It Till You Make It (graphika.com)

[iii] Deepfake scams have arrived: Fake videos spread on Facebook, TikTok and Youtube (nbcnews.com)

[iv] Threat Actors are Interested in Generative AI, but Use Remains Limited | Mandiant

[v] How a fake network pushes pro-China propaganda (bbc.com); Facebook finds disinformation and hacking campaigns targeting Ukraine : NPR

[vi] UNC1151 Assessed with High Confidence to have Links to Belarus, Ghostwriter Campaign Aligned with Belarusian Government Interests (mandiant.com)

[vii] Unusual CEO Fraud via Deepfake Audio Steals US$243,000 From UK Company - Noticias de seguridad (trendmicro.com)

[viii] Scammers are now using AI to sound like family members. It's working. - The Washington Post

[ix] Threat Actors are Interested in Generative AI, but Use Remains Limited | Mandiant

[x] Cybercriminals train AI chatbots for phishing, malware attacks (bleepingcomputer.com)

[xi] Email spoofing is a technique used in spam and phishing attacks to trick users into thinking a message came from a person or entity they know or trust.