Combating Deep Fakes

December 1, 2020

With social media hosting political discourse, scrutiny over social media companies' responsibility to maintain fact-based, honest means of dialogue has increased. Recently, and in light of the 2020 Presidential election, social media companies and their users have seen an increase in the use of deepfakes. Deepfakes are manipulated media sources meant for massive consumption deployed through machine learning (ML) and artificial intelligence (AI). An actor with a robust infrastructure and sophisticated operational capability may deploy this as a means of deception rising to military deception. As cyberwarfare becomes a traditional war fighting capability, concerted disinformation campaigns are no longer bound to the

physical realm but creeping into the digital space.

To stay one step ahead of deep fake operators, regardless of sophistication, tech giants are deploying ML and AI to increase awareness and detection by determining the veracity of the posted sources. This integrity check follows the security principle of zero trust. Zero trust security is an IT principle that holds all users attempting to access a network resource to the strictest identity verification standards. This principle can and should be extrapolated from the network to social media and other platforms for deepfakes.

Research aimed at combating deepfakes focuses on the automated nature of today's deception operations. Although inconspicuous to the naked eye, users should be aware of the following common tactics on the altered media:

- Distortion to voice and facial features
- Unnatural patterns to blinking and speech

Although difficult for a human to detect, tools and principles exist to aid in the fight against deepfakes. Traditional cybersecurity principles like zero trust can be used. Practically, zero trust can be developed as such:

- Match each device to a validated user and validate by deploying authentication codes and verifying their use among other account logins.
- Ensure the user is who they say they are and validate through enterprise authentication.
- With a zero-trust model, ensure each user is correctly given the right permissions and roles within the platform.

Whether it's a newsworthy event, like an election, or everyday use of social media technology, administrators' platform credibility is at stake, and they have an ethical imperative to ensure the veracity of information consumed by the public and their users. Failing to live up to that trust and allowing willful disinformation through a lack of due diligence is irresponsible.