AI and manipulated media pose significant challenge in the security sphere

[Content preview – Subscribe to **Jane's Intelligence Review** for full article]

The proliferation of accessibility to artificial intelligence (AI) capabilities by state and non-state actors is having an increasing impact on information dissemination. *Tate Nurkin* assesses the potential impact of AI-related media manipulation and deepfakes on the security sphere

Key Points

- Although artificial intelligence is becoming an increasingly prevalent tool in the military and security spheres, its use by malicious actors poses stark challenges for practitioners.
- Deepfakes have garnered significant public attention due to their largely misogynistic uses online, but they have also been shown to have potential for use in disinformation campaigns.
- Manipulated media content is not a new phenomenon, but the sophistication of AI-related manipulation means that military and security organisations will have to continually adapt to meet the challenge.

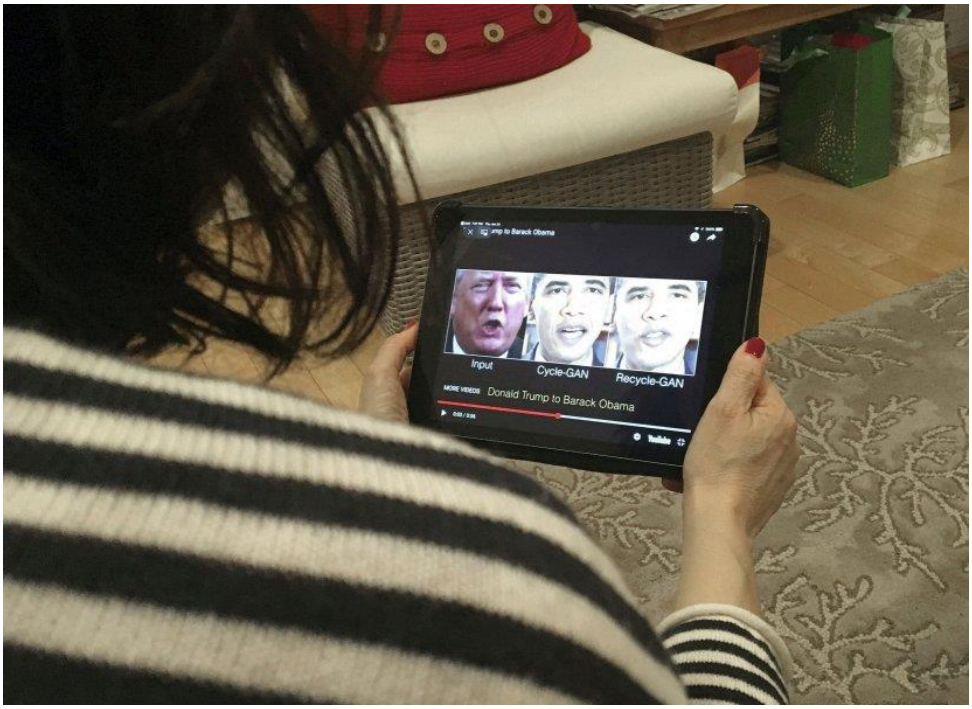
Artificial intelligence (AI) is at the centre of the advancement of military capabilities throughout the world: from dramatically speeding up intelligence processing and improving situational awareness and decision-making, to creating step-changes in simulation and training and human performance, to the enablement of autonomous platforms and systems, and beyond. New AI-enabled capabilities are already shaping accelerating competition for strategic, operational, and tactical advantage on the emerging cognitive battlefield.

However, the impact of AI-enabled capabilities for defence and security communities goes well beyond new or enhanced platforms and systems. Perhaps even more strategically affecting is the shrewd employment of AI-enabled tools in support of disinformation and deception efforts designed to disrupt the stability of polities and economies and divide societies, especially in open and democratic states. These efforts are already occurring, and the threat posed by AI-enabled disinformation tools is certain to intensify as their effectiveness is demonstrated, novel operational concepts for the use of these tools is adapted and refined, and the underlying AI technologies themselves develop further and diffuse.

The four fusions

In an Atlantic Council strategy paper entitled *A Candle in the Dark: US National Security Strategy for Artificial Intelligence* released in December 2019, Stephen Rodriguez and Tate Nurkin argued that the strategic context facing defence and security communities is characterised by the "fusion" of

four previously mostly separate concepts or conditions: peace and conflict, physical and digital worlds, reality and perception, and defence/security within commercial/consumer priorities. The intersection of these concepts has created a strategic and operational environment conspicuously vulnerable to exploitation by the employment of AI-driven disinformation, distortion, and disruption campaigns.



A woman views a manipulated video on 24 January 2019 that changes what is said by President Donald Trump and former president Barack Obama, illustrating how deepfake technology can deceive viewers. (Rob LeverAFP via Getty Images)

1761941

The fusion of the states of peace and conflict, for example, has created an unsettled world that verges on conflict, albeit frequently sub-threshold and non-kinetic conflict. Much of this conflict is taking place in the information domain, which offers state and non-state actors powerful asymmetric means of penetrating or degrading at-risk critical private sector and government network infrastructure and manipulating narratives, elections, and perceptions of their adversaries and competitors. General counsel of the US National Security Agency (NSA) Glenn Gerstell portrayed this in a September 2019 essay in *The New York Times* as one of "incessant, relentless, omni-present cyber conflict" in which technology-driven and targeted information operations feature prominently.

The fusion of the physical and digital realms is creating a hybrid environment in which distinguishing between interactions with real people and increasingly convincing digital entities and assets is already difficult. The US Intelligence Community and other expert assessments of Russian interference in the 2016 US presidential election prominently referenced the role of online bots in

driving social media activity in the months leading up to the election. In September 2019, researchers at the University of Southern California published analysis of approximately 31,000 Twitter accounts believed to be bots that showed that by 2018, bots were "better aligned with humans' activity trends, suggesting the hypothesis that some bots have grown more sophisticated".

The fusion of objective reality and misdirected perception is leading to a manipulated world that is vulnerable to the reinforcement of even easily discredited "alternative facts". As outlined in the Atlantic Council report "Influence operations will exploit the degradation of the truth to create and intensify divisive polarities and also offer sufficient justification for the instinct to retrench, to double down on interpretation and perspective in the face of established – but somehow still debated – facts".

The confluence of defence and security priorities and technologies of interest with those of the commercial and consumer sectors is a perennial concern as AI techniques such as machine learning, neural networks, generative adversarial networks, natural language processing, and other Fourth Industrial Revolution technologies and know-how are diffusing widely through licit and illicit channels. The resulting diffusion of the power to disrupt to a much broader set of state and non-state actors complicates the task for defence and national security communities throughout the world.

Individually and collectively, these fusions have created an environment in which state and non-state actors have more effective means of manipulating the political, societal, and cultural environments, fissures, and tensions of their adversaries and competitors. Indeed, so prominent and urgent is the challenge of advanced information and hybrid warfare that militaries and security communities throughout the world are now pivoting, at least at an organisational level, to better align to cope with intensifying information domain competition.

[Continued in full version…]

(720 of 3599 words)

For the full version and more content: