Clint Watts

- **Distinguished Research Fellow, Foreign Policy Research Institute**
- **Non-Resident Fellow, Alliance For Securing Democracy, German Marshall Fund of the United States**
- **Author, *Messing With The Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians and Fake News*[1]**

Statement Prepared for the U.S. House of Representatives – Permanent Select Committee on Intelligence

The National Security Challenges of Artificial Intelligence, Manipulated Media, and "Deepfakes" – 13 June 2019

All advanced nations recognize the power of artificial intelligence to revolutionize economies and empower militaries. But those countries with the most advanced artificial intelligence (AI) capabilities and unlimited access to large data troves will gain enormous advantages in information warfare. AI provides purveyors of disinformation the ability to rapidly recon American social media audiences to identify psychological vulnerabilities. AI powered systems can quickly generate modified content and digital forgeries advancing false narratives against Americans and American interests.

'Deepfakes', false audio and video content, grow in sophistication each day and their dissemination via social media platforms is far and wide. Historically, each advancement in media, from text to speech to video to virtual reality, more deeply engages information consumers enriching the context of experiences and shaping user reality. The falsification of audio and video allows manipulators to dupe audience members in highly convincing ways provoking emotional responses that can lead to widespread mistrust and, at times, physical mobilizations. False video and audio, once consumed and believed, can be extremely difficult to refute and counter.

Before the Kremlin's Internet Research Agency pushed bogus social media advertisements and manipulated content heading into the Presidential election of 2016[2], the Soviet Union authored and placed forged documents seeding conspiracies abroad. The most notable and possibly prolific claimed the U.S. created and proliferated the AIDS virus[3]. Last decade, manipulated video was disseminated to mainstream media outlets in an attempt to disparage an American diplomat serving in Russia.[4]

Moving forward, I'd estimate Russia, as an enduring purveyor of disinformation, is and will continue to pursue the acquisition of synthetic media capabilities and employ the outputs against its adversaries around the world. I suspect they'll be joined and outpaced potentially by China. China's artificial intelligence capabilities rival the U.S., are powered by enormous data troves to include vast amounts of information stolen from the U.S., and the country has already shown a propensity to employ synthetic media in television broadcast journalism.[5] These two countries along with other authoritarian adversaries and their proxies will likely use 'Deepfakes'

as part of disinformation campaigns seeking to 1) discredit domestic dissidents and foreign detractors, 2) incite fear and promote conflict inside Western-style democracies and 3) distort the reality of American audiences and the audiences of America's allies.

'Deepfake' proliferation presents two clear dangers. Over the long term, deliberate development of false synthetic media will target U.S. officials, institutions and democratic processes with an enduring goal of subverting democracy and demoralizing the American constituency. In the near and short term, circulation of 'Deepfakes' may incite physical mobilizations under false pretenses, initiate public safety crises and spark the outbreak of violence. The recent spate of false conspiracies proliferating via WhatsApp in India offer a relevant example of how bogus messages and media can fuel violence. The spread of 'Deepfake' capabilities will likely only increase the frequency and intensity of such violent outbreaks.

U.S. diplomats and military personnel deployed overseas will be prime targets for 'Deep Fake' disinformation conspiracies planted by adversaries. U.S. interests in the developing world, where information consumption has jumped from analog in-person conversations to social media sharing lacking any form of verification filter, will likely be threatened by bogus synthetic media campaigns.

Recent public discussions of 'Deepfake' employment heavily focus on foreign adversaries, but the greatest threat of inauthentic content proliferation may come not from abroad, but from home, and not from nation-states but from the private sector. Thus far, I've focused on authoritarian nation states, but a range of Advanced Persistent Manipulators[6] (APMs) will use their vast resources to develop or acquire 'Deepfakes' as needed in pursuit of their goals. Recent examples of disinformation and misinformation suggest it could be oligarchs, multi-national corporations, political action groups, public relations firms and activists with significant financial support that will seek out synthetic media capabilities and amplify 'Deepfakes' available in the wild. Regardless of whether the purveyor of 'Deepfakes' is international or domestic, the net effect will be the same: degradation of democratic institutions and elected officials, lowered faith in electoral processes, weakened trust in social media platforms, and potentially sporadic violence by individuals and groups mobilized under false pretenses.

The U.S. government should rapidly develop policies to promote appropriate use of artificial intelligence in media content creation and support technological development to verify the authenticity of video and audio content. First, Congress should implement legislation prohibiting U.S. officials, elected representatives and agencies from creating and distributing false and manipulated content. The U.S. government must always be the purveyor of facts and truth to its constituents assuring the effective administration of democracy via productive policy debate from a shared basis of reality.

Second, policymakers should work jointly with social media companies to develop standards for content accountability. Protecting account anonymity for those producing authentic content and exercising their free speech rights should be the goal for Western democratic societies. But

there is no public good in permitting the proliferation of inauthentic content from inauthentic accounts. For those producing and promoting inauthentic synthetic media from authentic accounts, they should be held responsible for their content and any violations of platform terms of service.

Third, the U.S. government should partner with the private sector to implement digital verification signatures designating the date, time and physical origination of content. Time stamping will help information consumers understand the authenticity of content and will help ensure a collective public reality.

Fourth, social media companies should enhance their labeling of synthetic content across platforms and work as an industry to codify how and when manipulated or faked content should be appropriately marked. Not all synthetic media is nefarious in nature. But, information consumers should be able to determine the source of information and whether it is an authentic depiction of people and events.

Fifth, the U.S. government, from a national security perspective, should maintain intelligence on adversaries capable of deploying 'Deepfake' content or the proxies they employ to conduct such disinformation. The Departments of Defense and State should develop immediate response plans for 'Deepfake' smear campaigns and 'Deepfake' inspired violent mobilizations overseas in an attempt to mitigate harm to U.S. personnel and interests.

Sixth, public awareness of 'Deepfakes' and its signatures will greatly assist in tamping down attempts to subvert U.S. democracy and incite violence. Public-private partnerships could develop educational materials regarding 'Deepfakes' which could then be delivered to Americans via the Internet and social media. Public awareness might likely be the best inoculation to the ill effects of fake audio and video content.

---

[1] Clint Watts, *Messing With The Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians and Fake News.* Harper, May 2018. Available at: https://www.amazon.com/Messing-Enemy-Surviving-Terrorists-Russians/dp/0062795996/ref=redir_mobile_desktop?_encoding=UTF8&qid=&ref_=tmm_pap_title_0&sr=

[2] See Volume I of *Report On The Investigation Into Russian Interference In The 2016 Presidential Election,* Special Counsel Robert S. Mueller, III, U.S. Department of Justice, March 2019. Available at: https://www.justice.gov/storage/report.pdf.

[3] For a breakdown of how Soviet and Russian disinformation works, see Adam B. Ellick and Adam Westbrook, "Operation Infektion - Russian Disinformation: From Cold War to Kanye." *The New York Times*, 12 November 2018. Available at: https://www.nytimes.com/2018/11/12/opinion/russia-meddling-disinformation-fake-news-elections.html

[4] See Jill Dougherty, "U.S. calls purported sex tape 'doctored' and 'smear campaign'," CNN, 24 September 2009. Available at: http://www.cnn.com/2009/US/09/24/russia.us.sextape/

[5] Lily Kuo, "World's first AI news anchor unveiled in China," *The Guardian*, 8 November 2018. Available at: https://www.theguardian.com/world/2018/nov/09/worlds-first-ai-news-anchor-unveiled-in-china

[6] Clint Watts, "Advanced Persistent Manipulators, Part One: The Threat to the Social Media Industry." Alliance for Securing Democracy, 12 February 2019. Available at: https://securingdemocracy.gmfus.org/advanced-persistent-manipulators-part-one-the-threat-to-the-social-media-industry/