INSTITUTE FOR HOMELAND SECURITY

Adversarial Actors and Artificial Intelligence:

Propaganda Processes and Threats to Critical Infrastructure

Institute for Homeland Security

Sam Houston State University

Andrew P. Davis, Ph.D

Abstract:

The extension of artificial intelligence (AI) into the propaganda efforts of adversarial actors, across a range of organizations be they state-sponsored cyber attackers, or homegrown extremist organizations, or foreign terrorist organizations, represents a significant evolution in propaganda strategies. AI capabilities enhance the reach of propaganda through targeted messaging and convincing deepfake technology, posing new challenges for security agencies and policymakers. Examination of reported cases of terrorist AI use, including ISIS' use of deepfakes, highlight the urgency of addressing these threats. Adversarial actors may also leverage AI technology to attack critical infrastructure systems. As AI technologies advance, counter-strategies must evolve to mitigate emerging threats effectively. Moreover, the ability of AI to automate propaganda dissemination, conduct sophisticated surveillance, and manipulate public perception through realistic fake content makes it a formidable tool for terrorists and challenge for counter terrorism operators. Addressing these challenges requires the development of robust AI detection tools, enhanced public awareness, and international collaboration to ensure a proactive and resilient approach to policing propaganda in the age of AI.

Keywords: Adversarial Actors, Propaganda, Deepfake Technology, Critical Infrastructure Security, Artificial Intelligence, Policing Strategies

Introduction and Overview

Artificial Intelligence (AI) has rapidly developed in recent years in ways that may dramatically affect our understandings of the propaganda processes for adversarial actors of many types (for ex. state-sponsored cyber attackers, homegrown extremist organizations/hate groups, or foreign terrorist organizations). Traditionally, such organizations have relied on relatively rudimentary print and digital tools to disseminate their ideologies, communicate organizational identity, and recruit followers. Such efforts have been the subject of a great deal of academic research (Kinney, Davis, and Zhang 2018). However, the advent of AI has transformed these efforts, introducing sophisticated techniques that significantly enhance the reach and impact of their messages. Academic research indicates that AI's capabilities in data analysis, natural language processing, and deepfake technology may markedly improve the efficacy of terrorist propaganda (Bazarkina 2023; Lakomy 2023). As there is evidence that adversarial actors are using AI technologies (Lakomy 2023), the implications of this shift are

profound, warranting an in-depth exploration of AI's role in the context of propaganda. The current research brief discusses the role of AI in propaganda processes, focusing on real-world examples of the practice as well as risks for critical infrastructure. This report will vacillate between identifying potential threats and discussing existing problems. Finally, I will discuss how AI is used in online policing operations and how AI detection practices may intervene in adversarial actors' propaganda efforts.

AI and Propaganda, General Risks

The main objectives of propaganda efforts are to radicalize, recruit, and incite individuals to commit acts that radically disrupt social order, make threats of violence, or otherwise destabilize social systems. Traditionally, these efforts have involved the dissemination of extremist narratives through various outlets including social media and other online forums (Kinney et al. 2018). These traditional methods often lacked the ability to precisely target and engage specific individuals based on their unique characteristics. The integration of AI has transformed these methods, enabling more targeted and personalized propaganda campaigns. AI algorithms can process vast amounts of data to identify vulnerable individuals, tailoring messages to resonate with their specific psychological and social profiles, a worry of previous research on this topic (Berger 2018). This enhanced targeting capability allows adversarial actors to craft messages that address the specific grievances, fears, and aspirations of individuals, making the extremist narratives appear more relevant and convincing. This precision not only increases the likelihood of radicalization but also enhances the efficiency of recruitment processes.

While computationally demanding and likely requiring some technical/computational expertise in the matter – as well as access to digital trace data obtained via a variety of methods,

AI has the potential to enhance recruitment targeting capabilities for adversarial actors. Here, organizations can potentially leverage data analytics and machine learning to analyze vast amounts of social media activity, online searches, and communication patterns of individuals. This allows organizations to identify individuals exhibiting signs of vulnerability or discontent, such as frequent searches for specific ideological content or participation in certain online forums. AI can segment these individuals based on psychological profiles, interests, and social networks, tailoring propaganda to address their specific grievances, fears, and aspirations. For example, an individual frustrated with economic hardships might receive messages emphasizing economic justice, while someone feeling isolated might be targeted with content offering a sense of community. AI optimizes the delivery of these messages through targeted ads, direct messages, and strategically placed posts, ensuring they reach the individual at the most impactful times. This precision not only increases the likelihood of radicalization but also enhances the efficiency of recruitment processes by continuously refining strategies based on real-time responses.

"Deepfakes" or hyper-realistic videos and audio recordings capable of manipulating public perception and deceiving audiences can be more easily developed with the use of generative AI technologies. Organizations may exploit this technology to fabricate speeches, political endorsements, and violent events, crafting false narratives that can incite further violence or sow discord. Previous work on this topic focusing specifically on terrorist organizations noted that they have not had access to the capabilities to produce convincing deepfakes (Citron & Chesney 2019), but access to generative AI makes this possible at relatively low cost for terrorist organizations. The potential for deepfakes to erode trust in media and institutions poses a significant threat to societal stability and security. Additionally, the

proliferation of deepfakes can complicate the verification processes for authentic information, further enabling malicious actors to manipulate public opinion and exacerbate social divisions (Vaccari & Chadwick, 2020). This challenge is compounded by the increasing sophistication of AI tools, which make deepfakes more convincing and harder to detect with traditional verification methods.

While AI technology has clear implications for the targeting of individuals at the micro-psychological level, there are more meso and macro-level implications as well. AI's capacity to quickly analyze large amounts of data allows organizations to identify patterns across diverse demographic groups. This ability enables them to better tailor messaging - creating highly persuasive propaganda geared toward individuals in specific cultural, religious, or social contexts. Use of this type of technology can identify and exacerbate existing social tensions - developing nuanced propaganda campaigns that can exploit local social tensions. By identifying and exploiting these macro-level patterns, adversarial actors may strategically incite discord and manipulate public sentiment on a larger scale to achieve their specific goals.

Risk to Critical Infrastructure

The integration of AI technologies into the efforts of adversarial actors more broadly writ poses risks to critical infrastructure, particularly in countries (such as the United States) where physical assets are most frequently tied to digital systems. Critical infrastructure typically refers to essential systems such as transportation networks, communication capacities, and energy grids that are crucial physical assets that help society function smoothly. AI can be exploited to identify vulnerabilities in these systems through advanced data analysis and pattern recognition in the digital systems upon which physical infrastructure depends. For instance, AI algorithms can analyze vast amounts of data from power grids, water supply networks, and communication

systems to detect weaknesses and potential targets for attacks. AI tools might be used to perform real-time monitoring of critical infrastructure, identifying abnormal patterns that could indicate a vulnerability or an ongoing attack. This could include unusual data flows in network traffic, unexpected shutdowns, or irregular usage patterns that deviate from the norm. By identifying weak points, actors can develop targeted strategies to launch cyber-attacks that disrupt essential services, causing significant damage and chaos. A successful attack on any of these infrastructures could lead to widespread chaos and panic, significantly impacting public safety and national security. The use of AI in this manner represents a new frontier in the threat landscape, necessitating robust cybersecurity measures to mitigate these risks (Malatji 2024).

While the threats of cyber-attacks on critical infrastructure are well known, less analysis is devoted to how propaganda processes may pose a risk to critical infrastructure. This risk is largely borne out through the erosion of public trust in critical infrastructure and related systems. By disseminating disinformation and deepfake content via AI, adversarial actors can create a sense of chaos and uncertainty. One can imagine many different scenarios. As an example, consider how a widely distributed deepfake video depicting a fabricated terrorist attack on a major transportation hub may lead to widespread panic among a population, causing disruptions in transportation and related economic activities. Alternatively, false information about the quality of a community's water supply may result in panic buying and resource hoarding that may strain supply chains. Such incidents could have cascading effects, undermining public confidence in the safety and reliability of critical infrastructure (Brundage et al., 2018).

The erosion of public trust in critical infrastructure caused by AI-generated propaganda can have sweeping consequences. Beyond the immediate panic and disruptions caused by deepfake media about critical infrastructure, such disinformation campaigns can undermine the

trust that society places in essential services like power grids, water supplies, and communication networks. Disinformation campaigns of this ilk are likely to pressure existing social tensions, leveraging messaging campaigns that heighten hostilities based on social inequality. For example, AI generated propaganda may exaggerate resource disparities in critical resources (such as clean water) leading to reduced social trust and lower levels of social cohesiveness.

Real-world Examples

While some of the threats described above remain hypothetical, there are several instances of AI use among well-known terrorist organizations specifically that are noteworthy. While this section focuses on terrorist organizations specifically, similar capacities have been developed by a range of adversarial actors that should be taken seriously. The Washington Post reports that Islamic State of Iraq and Syria (ISIS) has taken its already sophisticated propaganda efforts to new heights with the help of AI-driven tools, particularly deepfake technology. Here, following the ISIS claimed attack in Moscow in March 2024 that killed over 130 concert-attendees (see Ables, Ebel, Ilyushina, and Dixon 2024), the organization used AI-generated news reports in the form of short (roughly 90 second) clips as part of a AI generated news program called “News Harvest” – made to mimic the style of a Al Jazeera news broadcast (Harwell 2024). Reporters note that the news-cast style of the AI generated messaging makes it difficult for tech companies to moderate, making it easier to disseminate on social media (Harwell 2024). These programs have become integrated into the propaganda wing of the organization and may be a crucial messaging tool for them going forward.

The use of generative AI for propaganda is likely to diffuse across a range of terrorist organizations – a process driven by isomorphic and competitive pressures (Kinney et al. 2018).

For example, The International Centre for Counter-Terrorism (ICCT) notes the increasing use of generative AI by terrorist groups to produce high-quality propaganda material (Nelu 2024). The report suggests that groups as wide ranging as Hezbollah, Hamas, ISIS, and Al-Qaeda are increasingly exploring ways to integrate AI tools into their propaganda processes (Nelu 2024). The report notes al-Qaeda has produces posters with images that are likely AI generated, while Hamas has reportedly produced or manipulated media depicting disturbing content related to the current war in Gaza (Nelu 2024). The use of AI-generated content allows these groups to maintain a more continuous flow of propaganda, effectively targeting different demographic segments with tailored messages to support a given goal.

The Cost of Action with AI

While this report has noted practical implications of adversarial actors' adoption of AI technologies, there are academic implications that may be discussed. Rational choice approaches to understanding social action (see Anderton and Carter 2005) note that actors engage in a meticulous cost-benefit analysis before undertaking actions, weighing the potential gains against the associated risks and expenditures. This strategic calculation involves assessing the financial, human, material, and reputational resources required for an operation, alongside the anticipated social impacts. Scholars in this vein consider factors such as the likelihood of success, likelihood of counter-measures, and the expected value of the act in terms of their public image as especially important considerations for various groups. By optimizing their resources to maximize impact while minimizing risk, these groups can execute more effective and sustainable operations. Thus, tools like AI may be attractive tools to enhance benefits and reduce costs. This analytical approach reflects a sophisticated understanding of organizational behavior, strategic impact in the operational planning of adversarial actors.

AI technology fundamentally alters the cost-benefit calculus traditionally associated with disruptive activities by automating processes, reducing human costs, and potentially scaling up operations. With these capabilities, organizations may conduct operations that would have previously required substantial human resources and financial investment. For example, automated propaganda dissemination through AI-driven bots on social media and various online forums significantly reduces the need for human actor to manage online efforts. AI can generate and distribute content at a scale and speed unattainable by other means, allowing organizations to reach wider audiences without proportionally increasing their costs. The cost-benefit approach to terrorism argues that organizations most-often seek to maximally achieve their goals while minimizing their expenditures and risks. By lowering the costs of conducting various disruptive activities, AI may enhance the attractiveness of disruption as a strategy. The reduced costs associated with such operations can lead to an increase in the frequency and sophistication of various attacks, as organizations can allocate their resources more efficiently. Given the increased attractiveness of disruption as a tactic, operators interested in policing efforts should become attentive to technological advancements in the realm of generative AI in terms of terrorist efforts, but also policing measures that leverage AI to fight adversarial actors.

Practical Applicability: Countering AI of Adversarial Actors

As previously discussed, AI use by adversarial actors presents significant risks to populations in terms of the ability to scale up operations. Yet, AI use also creates new challenges for practitioners that should be discussed. The task of detecting and mitigating the activities of hostile actors becomes more complicated - but also more urgent – as groups adopt AI technology to enhance their capacities. Traditional methods for identifying online materials produced by such groups (such as flagging content) may not be useful given the deceptiveness of AI-driven

content, including convincing deepfakes and social media bots. Understanding the problem of AI detection in the context of adversarial actors and exploring tools to address this challenge is crucial for developing effective policing strategies. Likewise, practitioners are increasingly adoption AI-driven systems in their policing operations. Thus, practitioners should familiarize themselves with how these systems might be useful in fighting against hostile actors.

For practitioners, the challenge of detecting deepfake content at scale lies in the powerful ability of AI systems to generate realistic content that can deceive humans and automated detection systems. Deepfake technology, for instance, uses advanced machine learning algorithms to create realistic videos, images, and audio recordings in an effort to achieve some desired goal. Deepfake detection tools have been developed that use deep learning methods to identify deepfakes across different media types (video, image, audio, etc.) (Heidari et al. 2023). Algorithms that detect irregularities in video compression, inconsistencies in lighting and shadows, or unnatural movements in facial expressions are crucial for identifying deepfakes (Verdoliva 2020). Such detection tools leverage machine-learning algorithms to identify inconsistencies in digital content, such as subtle artifacts in deepfake videos that can serve as clues to their inauthenticity. When fake content is indistinguishable from authentic content, it can be useful to refer to a tamper-proof ledger to store content in order verify the integrity of original digital files, making it harder for -malicious actors to alter content undetected. Meanwhile, collaborative platforms can facilitate information sharing and coordinate efforts to combat online extremism, enhancing the overall effectiveness of policing strategies (West 2017).

Another promising area of development for practitioners is the use of AI in policing measures themselves. By using AI and related technologies to analyze vast amounts of data - from social media and other online platforms – practitioners can detect patterns indicative of

possible disruptive activity. Machine learning algorithms can identify suspicious behaviors, flagging potential threats for further investigation. Additionally, natural language processing (NLP) techniques that are embedded in AI designs allow practitioners to analyze communication for signs of radicalization or plotting, enabling preemptive interventions (Benigni and Carley 2016; Chee et al. 2023). By integrating AI and related methods into policing operations, practitioners can enhance their ability to predict, detect, and prevent terrorist activities more effectively. The specificity in detecting AI-driven threats is advancing with the development of forensic tools that can analyze minute details in digital content with some degree of accuracy. Moreover, continuous advancements in AI-driven cybersecurity measures are essential for protecting critical infrastructure from AI-augmented attacks (Yaseen 2023). By utilizing these advanced detection tools and integrating AI into their policing strategies, practitioners can significantly enhance their ability to safeguard against and respond to the evolving landscape of AI-driven threats.

In terms of using AI and related techniques to defend critical infrastructure, practitioners can pursue several specific strategies. Here, I am referring specifically to digital systems upon which physical assets rely. Practitioners should leverage AI trained on historical data to detect anomalies in network traffic and behavior to detect potential attacks and coordinate responses (Sommer and Paxson 2010). This may involve using data to identify patterns and trends from prior attacks, a task possible before the AI revolution. Novel approaches may involve the processing of intelligence data from various sources to predict potential threats – here, data may be collated from multiple sources including public-facing social media or dark web forums to identify potential outside threats. Insider threats may also be detectable by monitoring unauthorized data access incidents or unusual file transfers. Finally, AI can automate the

response to detected threats, reducing the time between detection and response. Automated systems can also isolate compromised segments of a network, deploying security patches, or redirecting traffic to minimize damage. This is crucial, as rapid response capability is paramount in limiting the impact of attacks and restoring normal operations quickly (Sharbaf, 2011). By integrating these AI-driven strategies, practitioners can enhance their ability to protect critical infrastructure from sophisticated and evolving cyber threats.

The Importance of Successful AI Detection for Interventions

Successful detection of AI content may empower practitioners to intervene in propaganda efforts. In instances where deepfakes and other manipulated content engineered by AI have already been identified by practitioners, they can easily stop their spread and reduce the intended impact of propaganda efforts on targeted populations. This capacity is critical to maintaining public trust in information sources and preventing social conflict. One important tool for intervention is the for removing or flagging AI-generated propaganda content over social media platforms. Here, social media companies can assist in the identification and takedown of such content, thus reducing its reach and preventing it from influencing large audiences. For instance, companies like Facebook and Twitter have recently used AI-driven moderation tools that detect and remove hurtful content in real-time (Mozur 2018). These tools do so by reassessing text, images, and videos for manipulation or extremism telltale signs to ensure that such malicious content is detected before it spreads virally.

Successful AI detection can also drive forward proactive policing with actionable intelligence that can be acted upon regarding the methods and tactics of adversarial actors. Even the comments on AI-generated material might point toward the architecture of actors' networks and strategies, making much more effective interventions possible. For example, knowledge of

specific themes and narratives that the propaganda articulates can inform counter-messaging efforts aimed at directly addressing and undermining extremist ideologies (Kinney et al. 2018; Torok 2013). Intelligence-driven policing operations make the operations stronger overall and diminish the otherwise powerful appeal that propaganda may have.

Successful AI detection can also protect critical infrastructure from threats using AI. Deepfake videos and disinformation campaigns can easily raise panic, especially when sensitive locations or events become targets. Practitioners, through the detection and neutralization of these threats before they have the possibility of causing harm, can ensure public safety and the continued running of critical services (Akhtar 2023). For instance, AI detection tools can detect signals of coordinated disinformation campaigns that would be used to misguide elections or other major political events that would antagonize the masses to gain control and provide warnings prior to such situation to the concerned authorities in order to take precautionary measures using counter-information and fact-checking (Kertysova 2018; Santos 2023).

Effective AI detection can foster public trust in policing efforts and the depth of the institutions of society. When people understand that authorities have the proper tools and skills to deter and counter threats by malicious actors using AI technology, it may in turn increase trust in the government and related social institutions. This is especially important in an era of eroding institutional trust and an environment with rife with political misinformation (Allcott & Gentzkow 2017).

Extended Analysis and Future Implications

Given the rate of technological development, it is likely that the sophistication and effectiveness of propaganda will increase. As previously discussed, organizations are already using AI technology to produce propaganda-media in video, image, and audio form---and one

may infer that they are using this technology to revolutionize their recruitment efforts (at least indirectly through propaganda messaging, but perhaps in direct communications). Automating the recruitment and propaganda generation process could make organizations operate with more efficiency and scale to become a much greater challenge to practitioners**.** The rapid development of NLP technologies makes it possible to devise more credible and persuasive narratives in the context of propaganda. AI-powered chatbots and virtual influencers could engage individuals on social media, simulate human-like behavior, and over time change their beliefs or behavior. To be more believable to the recruits and more relatable to them in order to convince them of the cause, these AI-enabled agents could be designed to imitate the speaking styles of friends and acquaintances.

Furthermore, AI can make it easier to develop more sophisticated cyber-attack strategies. By realizing the patterns and vulnerabilities that characterize cybersecurity systems, AI may allow hostile actors to formulate better attacks against critical infrastructure. This ranges from disrupting communication networks to apocalyptic scenarios such as causing a shutdown in energy grids, thereby paralyzing activities. This exemplifies how AI can be used not only to enhance the strategic but also operational capacities of hostile actors; hence, the need to have well-rounded cybersecurity measures in place-themselves informed by AI and related technologies.

The international community has a role in responding to these emerging threats with coordinated efforts for the development and implementation of state-of-the-art strategies for detection and mitigation. This would include investments in research for AI tools to detect deepfakes, etc., and enhancing coordination between governments, technology companies, and academia for developing robust policing strategies that leverage the very latest in AI technology

advancements (Buchanan 2021). Additionally, there is a need for a regulatory environment that can aid in controlling malicious use of AI technologies. This regulatory environment may provide frameworks ensuring ethical applications of AI as well as other accountability and oversight mechanisms. Finally, it is crucial to educate the public in techniques for identifying deepfake content as well as the dangers of manipulation online via AI (Calo 2017). A multi-dimensional approach to the technological, regulatory, and educational aspects of AI can better equip citizens to live with new technologies.

The integration of AI into the propaganda efforts of hostile actors represents a significant evolution in the strategies employed by extremist organizations. The ability to leverage AI for data analysis, targeted messaging, and deepfake technology has enhanced the reach and impact of terrorist propaganda, posing new challenges for security agencies and policymakers. The examination of well-known cases, such as the activities of ISIS and the use of deepfakes, underscores the urgency of addressing these threats. Furthermore, the potential risks to critical infrastructure highlight the broader implications of AI-driven propaganda, necessitating comprehensive and coordinated efforts to safeguard societal stability and security. As AI technologies continue to advance, it is imperative that policing strategies evolve in parallel to effectively mitigate these emerging threats.

References

Ables, K, Ebel, F, Ilyushina, M., & Dixon, R. (2024). Death toll from Moscow concert attack rises to 133 as more bodies found. *The Washington Post*. Access: https://www.washingtonpost.com/world/2024/03/23/russia-concert-shooting-attack-isis/?itid=lk_inline_manual_1

Akhtar, P., Ghouri, A. M., Khan, H. U. R., Amin ul Haq, M., Awan, U., Zahoor, N., ... & Ashraf, A. (2023). Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions. *Annals of operations research*, *327*(2), 633-657. https://doi.org/10.1007/s10479-022-05015-5

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-236. DOI: 10.1257/jep.31.2.211

Anderton, C.H., and Carter J. (2005). On rational choice theory and the study of terrorism." *Defence and Peace Economics* 16, no. 4: 275-282. https://doi.org/10.1080/1024269052000344864

Bazarkina, D. (2023). Current and Future Threats of the Malicious Use of Artificial Intelligence by Terrorists: Psychological Aspects. In: Pashentsev, E. (eds) The Palgrave Handbook of Malicious Use of AI and Psychological Security. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-22552-9_10

Benigni, M., & Carley, K. M. (2016). From tweets to intelligence: Understanding the islamic jihad supporting community on twitter. In *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings 9* (pp. 346-355). Springer International Publishing.

Berger, J. M. (2018). Extremism. *MIT Press*.

Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint arXiv:1802.07228*.

Buchanan, B. (2021). The Hacker and the State: Cyber Attacks and the New Normal of Geopolitics. *Harvard University Press*.

Calo, R. (2017). Artificial Intelligence Policy: A Primer and Roadmap. *California Law Review*, 109(1), 59-113.

Chee, S. J., Khoo, B. L., Muthunatarajan, S., & Carley, K. M. (2023). Vulnerable, threat and influencer characterisation for radicalisation risk assessment. *Behavioral Sciences of Terrorism and Political Aggression*, 1-19. https://doi.org/10.1080/19434472.2023.2206455

Citron, D. & Chesney, R. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*, 98(1), 147-155.

Harwell, D. (2019). Faked Pelosi videos, slowed to make her appear drunk, spread across social media. *The Washington Post*. [online] Available at: https://www.washingtonpost.com/technology/2019/05/23/faked-pelosi-videos-slowed-make-her-appear-drunk-spread-across-social-media/.

Heidari, A., Navimipour, N.J., Dag, H., & Unal, M. (2023) Deepfake detection using deep learning methods: A systematic and comprehensive review. *WIREs Data Mining and Knowledge Discovery*, Volume 14, Issue 2 e1520. https://doi.org/10.1002/widm.1520

Kertysova, K. (2018). Artificial intelligence and disinformation: How AI changes the way disinformation is produced, disseminated, and can be countered. *Security and Human Rights*, *29*(1-4), 55-81. doi:10.1163/18750230-02901005

Kinney, A., Davis, A. P., Zhang, Y. (2018). Theming for Terror: Organizational Adornment in Terrorist Propaganda. *Poetics*, (69), 27-40. https://doi.org/10.1016/j.poetic.2018.05.001

Lakomy, M. (2023). Artificial Intelligence as a Terrorism Enabler? Understanding the Potential Impact of Chatbots and Image Generators on Online Terrorist Activities. *Studies in Conflict & Terrorism*, 1–21. https://doi.org/10.1080/1057610X.2023.2259195

Malatji, M. (2024). Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. *AI and Ethics.* *https://doi.org/10.1007/s43681-024-00427-4*

Mozur, P. (2018). A genocide incited on Facebook, with posts from Myanmar's military. *The New York Times*. Available at: https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

Nelu, C. (2024). Exploitation of Generative AI by Terrorist Groups. *International Centre for Counter-Terrorism*. Accesse at: https://www.icct.nl/publication/exploitation-generative-ai-terrorist-groups

Santos, F. C. C. (2023). Artificial intelligence in automated detection of disinformation: a thematic analysis. *Journalism and Media*, *4*(2), 679-687. https://doi.org/10.3390/journalmedia4020043

Torok, R. (2013). Developing an explanatory model for the process of online radicalisation and terrorism. *Security Informatics*, 2(1), 1-10. https://doi.org/10.1186/2190-8532-2-6

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1) https://doi.org/10.1177/2056305120903408

Verdoliva, L. (2020). Media forensics and deepfakes: an overview. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 910-932. Access: https://cybersecurity.uniroma1.it/sites/default/files/Media%20Forensics%20and%20Deepfakes%20-%20An%20overview.pdf

Ward, J. (2023). Generating Terror: The Risks of Generative AI Exploitation. *Combating Terrorism Center at West Point*. Access at: https://ctc.westpoint.edu/generating-terror-the-risks-of-generative-ai-exploitation/

West, D. M. (2017). How to combat fake news and disinformation. *Brookings Institution*. Available at: https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation/.

Yaseen, A. (2023). AI-driven threat detection and response: A paradigm shift in cybersecurity. *International Journal of Information and Cybersecurity*, 7(12), 25-43. https://publications.dlpress.org/index.php/ijic/article/view/73

INSTITUTE FOR HOMELAND SECURITY

Sam Houston State University

The Institute for Homeland Security at Sam Houston State University is focused on building strategic partnerships between public and private organizations through education and applied research ventures in the critical infrastructure sectors of Transportation, Energy, Chemical, Healthcare, and Public Health.

The Institute is a center for strategic thought with the goal of contributing to the security, resilience, and business continuity of these sectors from a Texas Homeland Security perspective. This is accomplished by facilitating collaboration activities, offering education programs, and conducting research to enhance the skills of practitioners specific to natural and human caused Homeland Security events.

Institute for Homeland Security
Sam Houston State University

Davis, A. P. (2024). Adversarial Actors and Artificial Intelligence: Propaganda Processes and Threats to Critical Infrastructure. (Report No. IHS/CR-2024-1040). The Sam Houston State University Institute for Homeland Security.

https://doi.org/10.17605/OSF.IO/XSBVU