CYBER CAPTAINS CAREER COURSE AND AUGUSTA UNIVERSITY'S MASTER OF ARTS IN INTELLIGENCE AND SECURITY STUDIES

INFORMATION WARFARE CONFERENCE 2024

This conference is sponsored by research supported by the U.S. Office of Naval Research under award number N000142212549, as well as from the Master of Arts in Intelligence and Security Studies, and the Department of Social Sciences.

Thursday, August 22nd

0930-1530

Augusta University JSAC Coffeehouse

Routledge
Taylor & Francis Group

Defense & Security Analysis

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cdan20

Artificial intelligence and information warfare in major power states: how the US, China, and Russia are using artificial intelligence in their information warfare and influence operations

Lance Y. Hunter, Craig D. Albert, Josh Rutland, Kristen Topping & Christopher Hennigan

To cite this article: Lance Y. Hunter, Craig D. Albert, Josh Rutland, Kristen Topping & Christopher Hennigan (2024) Artificial intelligence and information warfare in major power states: how the US, China, and Russia are using artificial intelligence in their information warfare and influence operations, Defense & Security Analysis, 40:2, 235-269, DOI: 10.1080/14751798.2024.2321736

To link to this article: https://doi.org/10.1080/14751798.2024.2321736

Published online: 05 Mar 2024.

DEFENSE & SECURITY ANALYSIS
2024, VOL. 40, NO. 2, 235–269
https://doi.org/10.1080/14751798.2024.2321736

Artificial intelligence and information warfare in major power states: how the US, China, and Russia are using artificial intelligence in their information warfare and influence operations

Lance Y. Hunter[a], Craig D. Albert[a], Josh Rutland[b], Kristen Topping[a] and Christopher Hennigan[c]

[a]Department of Social Sciences, Master of Arts in Intelligence and Security Studies (MAISS) Program, Augusta University, Augusta, GA, USA; [b]U.S. Army Cyber, Fort Eisenhower, GA, USA; [c]Department of Government and Public Service, Deloitte, Washington, DC, USA

ABSTRACT

Previous research in security studies contends that information warfare (IW) is becoming a critical element in states' overall security strategies. Additionally, many researchers posit that artificial intelligence (AI) is quickly emerging as an important component of digital communications and states' military applications worldwide. However, less is known regarding how states are incorporating AI in their information warfare and influence operations (IWIO). Thus, given the growing importance of AI and IW in global security, this paper examines how the United States, China, and Russia are incorporating AI in their IWIO strategies and tactics. We find that the US, China, and Russia are utilizing AI in their IWIO approaches in significant ways depending on each state's overall IW strategy, with important implications for international security.

KEYWORDS

Artificial intelligence (AI); information warfare (IW); information warfare and influence operations (IWIO); major power states; United States; China; Russia

Introduction

Many government officials, military leaders, and researchers acknowledge the growing importance of information warfare and influence operations (IWIO) in the realm of international security. As Khan states: "While information warfare is as old as military history, the revolution in communication sciences has changed its nature. It has become a double-edged sword equally important for the powerful states as well as technically poor states, non-state actors, and individual experts in software."[1] IWIO is critical to international security because it can shape global and domestic narratives that affect stability within states, international alliances, and the survivability of governments and leaders.[2] As military officials and researchers remark: "In current and future warfare, information superiority could be the single most decisive factor."[3] IWIO is evolving at a rapid pace due to the technological changes that have occurred in recent years that influence IWIO capabilities. One evolving technology of particular importance to IWIO is artificial intelligence (AI). AI

CONTACT Lance Y. Hunter lahunter@augusta.edu

developments have a significant impact on IWIO because they enhance the speed and effectiveness of IWIO operations, as well as shape the specific IWIO tactics that can be employed.[4] In addressing the role of AI in defence competition, Hurley remarks: "The onset of what is perceived to be the next global 'arms' race will position 'the winner' as the top superpower that could define and dictate future directions and priorities across the globe."[5]

Previous valuable IWIO scholarship has focused on the strategies and tactics behind IWIO,[6] and researchers have documented the effect AI has on military applications in major-power states.[7] However, minimal research has considered how AI is affecting the IWIO strategies of major-power states. This topic is important due to the expanding role of AI and IWIO in defence and security, and the rapidly growing influence AI is having on IWIO tactics. Perez and Nair note that "AI and its subcomponents … are serving as powerful tools for generating and amplifying disinformation about the Russia-Ukraine war, particularly on social media channels."[8] Based on the above statements, it is clear AI is already shaping the digital battlespace, and in the case of the war in Ukraine, it is also likely affecting the kinetic realm. Thus, this paper examines how the US, China, and Russia are incorporating AI in to their IWIO strategies and tactics and considers the implications for international security. In our analysis, we find that the US, China, and Russia are utilising AI in their IWIO in significant ways depending on each state's overall IWIO strategy. Additionally, we argue that the manner in which AI is being used in IWIO by the US, China, and Russia has significant effects on the stability and security of states that could be targeted in IWIO. Overall, we contend that considering the effect AI has on IWIO strategies and tactics in the major power states is a vital component of the modern security landscape (Table 1).

We focus on the US, China, and Russia due to the large amount of influence each state has in global politics and international security.[9] Even though other states could be considered major power states, we contend that the IWIO strategies employed by the US, China, and Russia significantly affect other states in their regions, as well as globally. More specifically, we argue that while other states' IWIO strategies are also important to consider, the US, China, and Russia play an outsized role in affecting international security based on their IWIO capabilities and operations.

The layout of the paper is as follows. First, we discuss our definition of AI and how it relates to IWIO. We then discuss our conceptualisation of IWIO and why it is important to global security. In this section we also highlight the distinct IWIO doctrines of the US, China, and Russia and how they motivate their IWIO strategies and tactics. Next, we analyse how surveillance capitalism, data collection processes, regime type, and grey zone activities affect the use of AI in IWIO for the US, China, and Russia. We then examine how the US, China, and Russia are incorporating AI in their specific IWIO strategies and tactics, discussing similarities and differences amongst the three states. Lastly, we discuss the ways in which the US, China, and Russia's application of AI in their IWIO strategies could affect international security.

Artificial intelligence

There are many different definitions of AI and there is much debate regarding how to define it. Based on prior scholarship, we advance a common definition of AI that has

Table 1. Artificial Intelligence (A.I.) Strategy, Development, and Integration in Major Power States.

State	Strategies and Tactics	Developers and Companies	Integration and Applications
US	Focus on information operations and the means through which the state pursues information warfare[a]	Booz Allen Hamilton[f]	Involve psychological operations, operations security, military deception, and electronic warfare[k]
	Train via complex wargame simulations generated and adapted by machine learning programmes and operational concept improvement bolstered by machine learning[b]	US Intelligence Community[g]	Monitor large streams of data to detect information patterns which can be identified as hostile information[l]
	Third Offset Strategy (TOS)[c]	National Artificial Intelligence and Initiative Office[h]	Incorporate learning machines, human-machine collaboration, assisted human operations, human-machine combat training, and network enabled autonomous weapons[m]
	Artificial Intelligence Use Case Inventory[d]	DARPA[i]	GEAR program[n]
	Pillar AI Strategy: Deliver AI-enabled capabilities that address key missions, scale AI's impact across DoD through a common foundation that enables decentralised development and experimentation, cultivate a leading AI workforce, engage with commercial, academic, and international allies and partners[e]	Joint Artificial Intelligence Center[j]	Use unmanned aerial systema (UAS) and artificial intelligence (AI)-enabled autonomy capability[o] Deepfake Detector[p] Multidimensional Anomaly Detection fusing HPC, Analytics, and Tensors (MADHAT)[q] MediFor[r]
China	Use "Asymmetric warfare" capable of offsetting technological inferiorities that might otherwise impact a state's ability to challenge geopolitical adversaries[s]	State Key Laboratory for Communication Content Cognition[y]	Invest heavily into information warfare capabilities[ab]
	Overcome superior forces by "robbing an army of its spirit" and a commander of his courage[t]	Comprehensive National Science Center[z]	Enhance population control, as well as to profile and control its ethnic minorities[ac]
	Utilise information technology in a wide variety of sectors and regions, including "disruption through trade wars, information manipulation in cyberspace and military integration of advanced technologies[u]	Academy of Military Medical Sciences (AMMS)[aa]	Strengthen specific advantages in social control and information management[ad]
	Focus on 'informatisation warfare,' or 'xinxihua,' the application of information technology to all aspects of military operations[v]		Exploit contradictions in interests and perceptions between groups and create division[ae]
	New Generation Artificial Intelligence Development Plan[w]		Use big data and artificial intelligence" to strengthen China's leadership and better understand the citizens[af]
	4 Key Sectors: Increased information-processing capabilities, rapid decision-making, the use of swarms, and cognitive warfare[x]		Co-ordinate campaigns of inauthentic posts to create the illusion of widespread grassroots support for a policy, individual, or viewpoint, when no such widespread support exists[ag]
Russia	Cover a wide swath of technology, where "jamming electronic communication and disrupting access to the electromagnetic spectrum, Cyber espionage, and distributed denial of services (DDoS) attacks are no different from (and work in tandem with) using trolls and bots to spread dis/misinformation, establishing pro-Russian media outlets, or supporting local sympathisers to propagate favourable messages[ah]	"Internet troll factory"[am]	Sow distrust in U.S. elections[ap]
	Revitalise "traditional values at the individual level and a focus on returning the glory of the Soviet Union on the national level[ai]	Russian Internet Research Agency [an]	Deepen pre-existing socio-political fault lines in Western societies[aq]
	Divide and polarise society, tear it into small pieces and fragments, and make these fragments sincerely hate each other in order to have them collide with each other thereby initiating a fight for destruction or combine their aggression into a uniform stream and direct it against the ruling government[aj]	FSB[ao]	Equip state backed propaganda facilities with AI powered Deepfake technology that can create more realistic false narratives by constructing fake images and even video clips involving key figures that support whatever narrative the "troll" is attempting to push[ar]

(*Continued*)

Table 1. Continued.

State	Strategies and Tactics	Developers and Companies	Integration and Applications
	Use internet connected technology to “undermine, manipulate, and mislead the information people consume as it believes this can advance its political and military objectives[ak] Do not differentiate between technologies and information, and instead of calling the digital-only system Cyberspace, refer to “as the ‘information space,’ which includes both computer and human information processing[al]		Inject dis/misinformation (partially attained through cyberattacks), and fake news stories that a majority of those exposed to believed true at the time[as] Gray Zone[at] Use AI instruments, including electronic warfare, influence operations, propaganda campaigns, and disinformation[au] Generative Adversarial Networks (GANs)[av]

[a] Army Techniques Publication 3-13.1, 2018, 1-1.
[b] McGrath, ‘Twenty-first century’; Pomerleau, ‘Pentagon AI Team Sets Sights on IWIO’.
[c] Work, ‘The Third U.S. Offset Strategy’.
[d] US Department of State, ‘AI Inventory’.
[e] CDAO, ‘About the JAIC’.
[f] Booze Allen Hamilton 2022.[14]
[g] National Security Commission on Artificial Intelligence 2021, 110.
[h] National Artificial Intelligence and Initiative Office 2022; Subcommittee on Networking and Information Technology Research and Development Committee on Science and Technology Enterprise 2021, 18.
[i] NSTC 2019.
[j] CDAO, ‘About the JAIC’.
[k] Theohary, ‘IWIO: Issues for congress’.
[l] McGrath, ‘Twenty-first century’; Pomerleau, ‘Pentagon AI Team Sets Sights on IWIO’.
[m] Hilner, ‘The Third Offset Strategy and the Army Modernization Priorities’.
[n] Pimentel 2022.[15]
[o] CDAO Public Affairs 2022.
[p] US Department of State, ‘AI Inventory’.
[q] C4ISRNET, ‘Pentagon AI team sets sights on IWIO’.
[r] Sybert, ‘DARPA Launches New Programs to Detect Falsified Media’.
[s] Wang, ‘Asymmetric war?’.
[t] Tzu, ‘The Art of War’, 108.
[u] Saalman, ‘China and its hybrid warfare spectrum’, 95.
[v] Buck, ‘China in the Asia-Pacific Cyber Domain’, 1.
[w] Kania, ‘The Strategic Support Force and the Future of Chinese Information Operations’, 3.
[x] Takagi, ‘The Future of China’s Cognitive Warfare’.
[y] Pollpeter and Kerrigan, ‘The China AI and Autonomy Report’, 5.
[z] Towey, ‘Researchers in China claim they have developed ‘mind-reading’ AI’.
[aa] Bureau of Industry and Security, Commerce, ‘Addition of Certain Entities’.
[ab] Cheng, *Cyber Dragon*.
[ac] Daniels and Chang, ‘National Power After AI’, 12.
[ad] Buck, ‘China in the Asia-Pacific Cyber Domain’, 4.
[ae] Pollpeter and Kerrigan, ‘The China AI and Autonomy Report’, 3.
[af] Ibid., 5.
[ag] US Department of State 2022.[16]
[ah] Bolton, ‘Targeting Ontological Security’, 130.
[ai] Ajir and Vailliant, ‘Russian IWIO’, 70.
[aj] Manoilo, ‘Modern-Day IWIO and Hybrid War Operations’, 3.
[ak] Topor and Tabachnik, ‘Russian Cyber IWIO’, 115.
[al] Ajir and Vailliant, ‘Russian IWIO’, 74.
[am] Wilde and Sherman, 2023, 34.[17]
[an] O’Donnell, ‘Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos’, 710.
[ao] Polyakova, ‘Weapons of the weak: Russia and AI-driven asymmetric warfare’.
[ap] Wojnowski, *Russian Interference in the U.S. Presidential Elections in 2016 and 2020 as an Attempt to Implement a Revolution-like IWIO Scheme. Part II*.
[aq] Cunningham, ‘A Russian Federation IWIO Primer’.
[ar] Marks and Bendett, ‘Russia Is Systematically Copying U.S. Military AI Robotics’.
[as] Bolton, ‘Targeting Ontological Security: IWIO in the Modern Age’, 136.
[at] Morgan et al., *Military applications of AI: ethical concerns in an uncertain world*, 83.
[au] Ibid.
[av] Ibrahim, ‘“We are not prepared”: Russia uses AI, deep fakes in propaganda warfare.’

been used by many previous researchers. We define AI as the ability of machines or computer programmes to execute tasks in a similar manner as humans in areas such as visual and spatial perception, audio, text, language, and speech recognition, decision-making, data collection and data analysis, and learning.[10] Within this definition, it is important to acknowledge that many of the functions we discuss are forms of machine-learning, which is a subset of AI.[11] As Kumar et al. remark, "Machine learning is a branch of artificial intelligence that aims at enabling machines to perform their jobs skilfully by using intelligent software."[12] These machine learning functions include, but are not limited to: algorithmic content moderation, algorithmic classification, machine-learning enabled content generation (including image, text, and algorithmic feeds), regression analysis, and clustering.[13] Thus, when using the term AI, we are referring to many types of machine-learning functions. We use the phrase AI because it encompasses these functions as well as additional activities that fall outside traditional machine-learning functions, or that are extensions of machine learning functions including but not limited to: symbolic AI, natural language processing, and expert systems.[18]

Information warfare and influence operations (IWIO)

This paper focuses on how states apply AI in their IWIO strategies and tactics. A large portion of the paper examines how states use AI in what would be considered information warfare (IW) operations in the digital space, specifically through social media sites, such as attempting to influence specific populations through purposeful narratives digitally. However, within this context, we also examine Cyber-operations (to a lesser degree) as they can be linked with broader IW campaigns. Typically, this is referred to as Cyber-enabled influence operations (CEIO) and is an example of IW in Cyberspace. We understand CEIO as "information operations that leverage means and dynamics unique to Cyberspace – with a particular focus on operations targeting social media."[19] CEIO is the cognitive hacking that occurs through digital media and is generally combined, or can work in unison with, physical Cyber-attacks, thus fitting within the larger framework of information warfare and the more specific terminology of IWIO. For example, a Cyber-attack that targets an adversary state to acquire data to use in a digital IW operation is also considered in the study since the two separate actions (i.e. the Cyber-attack and subsequent use of the data from the Cyber-attack to form propaganda narratives online) are part of a larger IW strategy. Thus, we use the term Information Warfare and Influence Operations (IWIO), which includes both Cyber-enabled influence operations (e.g. spreading propaganda narratives on social media), and Cyber activities, to account for the Cyber-component of the study. We elaborate on our conceptualisation of IW and IWIO below.

IW is a complex phenomenon that academics and military professionals have struggled to define. Researchers have conceptualised IW as "the deliberate manipulation or use of information by one party on an adversary to influence the choices and decisions the adversary makes in order for military or strategic gain."[20] Whilst broad, this definition highlights the fundamental elements of IW, namely the targeted and intentional desire to influence an adversary's decision-making through information.[21] US military doctrine tends to focus on information operations, the means through which the state pursues IW,[22] and these operations can fall within the realms of psychological operations, operations security, military deception, and electronic warfare.[23] By IWIO, we rely on Lin's definition, which is the "deliberate use of information (whether true of false) by one party on an adversary to confuse, mislead, and ultimately to influence the choices and decisions that the adversary makes."[24] Additionally, IWIO describes the integration of multiple aspects of IW (electronic warfare, Cyber-warfare, and psychological warfare) to achieve strategic objectives.[25] As mentioned, the focus of the paper is on how AI is being used in IW strategies and tactics as it pertains primarily to CEIO. However, we also discuss Cyber-attacks within the analysis. Thus, as previously mentioned, we use the term IWIO that incorporates the more broadly considered, full range of activities from CEIO, Cyber-attacks, and IW.

IWIO is considered significant in the context of modern great power competition for its ability to act as both a force multiplier and an alternative to traditional kinetic means of persuasion. China describes its value as a means of "asymmetric warfare" capable of offsetting technological inferiorities that might otherwise impact a state's ability to challenge geopolitical adversaries.[26] This is embodied in Sun Tzu's teachings on how to overcome superior forces by "robbing an army of its spirit" and a commander of his courage.[27] Russia has recognised the potential power IWIO offers as well, notably in its own efforts to sow distrust in U.S. elections via meddling[28] and attempting to deepen pre-existing socio-political fault lines in Western societies, notably within the US.[29] In essence, the capability of IWIO lies in its ability to redistribute power and negate the need for rapid advancements in military capabilities.[30]

It should be noted that as of this writing, each of these three states (US, China, and Russia) advance very different approaches to IWIO. The acronym DIME (Diplomatic, Informational, Military, and Economic) often refers to instruments of national power,[31] and unlike China and Russia, the US does not have an entity responsible for the informational component of DIME.[32] In fact, it can be argued that the three states view the IWIO domains through fundamentally distinct lenses. The US views IWIO as largely taking place in the Cyberspace domain. Russia views it through the information domain, and China uses a mixed domain preference.[33] Furthermore, the US generally separates peacetime from wartime activities in the information domain. In other words, the US limits its IWIO capabilities when it is not engaged in conflict with another state.[34] This places the US at a disadvantage when compared with Russia and China, who constantly engage in offensive IWIO activities as a matter of strategy and policy in the information domain. The US, however, considers IW activities as force multipliers within an already defined strategic conflict; in other words, the US only engages in IW activities during conflict, whereas China and Russia see IW as perpetual and persistent activities. Additionally, the US conducts IWIO through the military and DoD broadly while China and Russia employ IWIO through a whole of society approach. Russia's official stance is there is no distinction in their IWIO strategies and tactics in peacetime and wartime.[35] In other words, Russia is permanently in a state of conflict within the information environment, specifically against the United States, but also as a general strategy in the international arena. China also pursues IWIO differently than the US, but more akin to the Russian approach. As with Russia, China's Cyber-enabled influence operations revolve around the operational imperative of "peacetime wartime integration."[36] These differences are important to acknowledge when considering how each state applies AI to their IWIO strategies. We expand on each state's IWIO strategies and tactics in the later sections of the paper.

Throughout the paper, we refer to states' IWIO strategies and tactics. Strategies refer to broad, long-term goals or plans that are advanced to achieve the desired political and military objectives of a state. Tactics are more detailed, immediate actions that are guided by strategies and are designed to accomplish shorter-term goals within the context of a given strategy and can vary based on circumstance.[37] Thus, IWIO tactics are the specific actions that are carried out based on the IWIO strategies of each state.

Governance structures, AI, and data collection

Before examining how the US, China, and Russia are using AI in their IWIO, it is important to consider how the type of government institution of each state affects their data collection processes, IWIO strategies, and how they apply AI within these strategies. First, while the US obtains a large amount of data on potential security threats domestically, China and Russia can collect data on their domestic populations to a greater extent through more aggressive methods due to the non-democratic nature of their governments. In other words, China and Russia can surveil their populations more freely without privacy, or civil liberty concerns, and use the data to train AI algorithms that can be used in IWIO.[38]

The authoritarian and centralised nature of the Chinese and Russian governments allow both states greater ability to collect data domestically and internationally compared with democracies such as the US. These data can be deployed to train AI algorithms to use in IWIO and grey zone activities.[39] In contrast, the US faces more difficulty in collecting and utilising large-scale data in the same manner due to its democratic principles that include some protection of privacy rights, as well as political and bureaucratic oversight regarding the decision-making processes that govern the collection and use of domestic surveillance data. This is also a component of the US' strategy regarding not employing IW more broadly outside of wartime, including against its strategic adversaries. This is not to argue that western democracies do not collect security data domestically, or that US multinational corporations (MNCs) do not collect large amounts of domestic data through surveillance capitalism frameworks. Rather, it is to argue that democracies such as the US cannot collect data and wield it in IWIO to the same degree as authoritarian states such as China and Russia based on the democratic and decentralised nature of US governance. Relatedly, an important point to highlight within the context of regime type, AI, and IWIO is how surveillance capitalism affects data collection, AI algorithms, and IWIO.

Surveillance capitalism and data collection

Surveillance capitalism refers to the practice of technology companies collecting and selling data and personal information while employing specifically tailored algorithms to predict and affect individual behaviour. Surveillance capitalism is based on the idea that consumer data is a driving force within the digital economy. Consumer data is often used by corporations in conjunction with algorithmic programmes to target individuals to affect their buying habits, as well as by political actors to influence individuals' political viewpoints. Some of the potentially negative effects of surveillance capitalism are that it can harm privacy, erode confidence in governmental systems, increase polarisation within societies, and exacerbate various forms of discrimination.[40] Surveillance capitalism is important to discuss regarding AI and IWIO because states such as China and Russia can more easily deploy AI algorithms to sow division and divide populations based on the data gathered through companies operating within the surveillance capitalist framework compared with western democracies.[41] The reason is that the data collected can allow the AI algorithms to identify and target individuals for IWIO based on their shopping and social habits and political viewpoints. As Dawson states: "governments must recognise microtargeting – data informed individualised targeted advertising – and the current advertising economy as enabling and profiting from foreign and domestic information warfare being waged on its citizens."[42] China and Russia can use data collected through the surveillance capitalism framework to a greater extent than democracies such as the US because their state identities and authoritarian governance structures allow for more aggressive data collection programmes, which drive AI algorithms that are useful for IWIO.

China is employing surveillance capitalism methods to collect large amounts of data to power AI algorithms that could be used for numerous purposes including, but not limited to: increasing domestic surveillance and population control of Chinese citizens, targeting Uighur minorities in Xinjiang for re-education purposes, and conducting IWIO to divide and polarise societies within democracies such as the US.[43] Through Chinese security forces, Chinese companies, and the Belt and Road Initiative (BRI), China is collecting massive amounts of data that can be used to tailor AI algorithms that can help China achieve its larger IW goals of controlling domestic populations, spreading the PRC's political narratives internationally, and increasing division within western democracies to undermine confidence in their governments.[44]

One example of how China exerts control over its domestic population through data collection and AI is through the application WeChat. WeChat is a Chinese social media application that can be used for a range of functions including instant messaging and mobile phone payment and fund transfers. WeChat was developed by the Chinese company Tencent. It is estimated that over 60% of transactions in China are conducted through WeChat.[45] The data from these transactions are used by the PRC to monitor and control Chinese citizens. China has identified 75 behavioural characteristics to identify if someone is considered susceptible to radicalisation,[46] and WeChat is the ideal platform to monitor and track individuals' behaviour to measure the extent they conform with PRC standards. Additionally, the issue is not limited to the domestic use of AI to monitor and control individuals. Researchers and policy-makers have raised concerns that companies such as Tencent, Byte Dance, and Zoom collect large amounts of data on citizens around the globe through online gaming, social media, and video conferencing platforms that could be potentially used for AI algorithms for targeted IWIO.[47] As Dawson states, "While Chinese data collection is perceived as a national security threat, domestic data collection is viewed as a digital privacy issue – these are not separate issues. Domestic digital privacy is fundamentally linked to national security."[48]

Russian influence operations have also used the surveillance capitalism framework to deploy specifically tailored algorithms to increase political polarisation within the

US. One example is Russia's use of data collection processes and AI algorithms to target military veterans and individuals more prone to support the military for propaganda campaigns.[49] It is also estimated that leading up to the 2016 election Russian IWIO efforts were likely designed to motivate some voters to turn out to the polls, whilst discouraging others.[50] The bipartisan US Senate investigation discovered that during the 2016 US Presidential elections, Russia conducted IWIO targeting US infrastructure using Facebook-targeted advertising and used social media to intensify social divisions in the US.[51] Additionally, Russian AI algorithms have been employed to encourage some individuals to attend protests, whilst encouraging others to attend counter-protests, thereby amplifying polarisation.[52] An important feature of the PRC and Russian IWIO efforts described above is the data collection efforts conducted within the surveillance capitalism framework significantly empowered the AI algorithms that were used to target particular individuals and groups for IWIO. Within this context, it is also important to examine each state's identity in considering grey zone activities and IWIO, and why democracies such as the US are likely to distinguish between wartime and peace time regarding IWIO while authoritarian states such as China and Russia are more prone to view IWIO within the framework of continuous conflict.[53]

State identity and grey zone activities

Grey-zone actions are those that are below the threshold of armed kinetic conflict and are designed to achieve specific goals. They often include, but are not limited, to Cyber-attacks, information warfare, economic coercion, and the use of proxy forces.[54] As Tiwari remarks: "The grey zone has been defined as the space between peace and war, characterised by the ambiguity of objectives, the participants involved, and the role of military force in response that remains below the level of war."[55] Whilst the US has focused much of its efforts on protecting vital infrastructure, China and Russia have viewed the US (and much of the West) as a threat to their security and geopolitical ambitions, thus prompting both states to take more aggressive actions in the grey zone compared with the US. In this way, China and Russia view IWIO within the grey zone in the context of a broader, ongoing conflict with the US and West where there is no distinction made between wartime and peacetime activities.[56]

Examples of Chinese grey-zone activities include using IWIO to disseminate propaganda regarding territorial disputes in the South China Sea and Taiwan reunification, incorporating psychological warfare into military operations, and controlling digital information and spreading online disinformation to decrease morale and increase polarisation in Western democracies.[57] Russia has employed similar IWIO strategies in the grey zone, as evidenced through its 2014 doctrine which prioritises the use of Cyber and IWIO, to assist its military as well as Russia's use of the Internet Research Agency (IRA) to employ wide-spread disinformation campaigns through social media. Specific examples include Russian IWIO in the 2016 US Presidential elections and attempted IWIO in the 2018 US mid-term elections along with numerous IWIO in Ukraine leading up to the invasion and during the conflict.[58]

Overall, the nature of China and Russia's political regimes allow for more aggressive data collection domestically and internationally compared with the US. The data can be used in AI algorithms for IWIO that spreads disinformation, propagates PRC and Russian narratives, and seeks to undermine the confidence in democratic governments. These activities can be incorporated in grey-zone operations across numerous fronts. Thus, China and Russia's state identities consider IWIO within the context of continual conflict with the west, and the authoritarian nature of the Chinese and Russian regimes leads to more aggressive use of AI in data collection efforts, IWIO, and grey-zone activities compared with the US. In contrast, the US' identity and democratic institutions place more restrictions on its ability to gather data to use for AI algorithms for IWIO. Thus, the US places more emphasis on developing AI programmes to detect and counter IWIO and adversarial grey-zone actions. Having examined how regime type and state identify affect data collection, AI algorithms, and grey-zone activities, along with the role of surveillance capitalism in IWIO, we now turn to analysing how each state is applying AI within their IWIO strategies and tactics.

US

IW background information

The origins of US IW can be dated to World War II. During World War II, President Franklin Roosevelt established the Office of War Information (OWI) to organise US propaganda. In addition, the Office of Strategic Services (under the Joint Chiefs of Staff) employed psychological warfare techniques in concert with overseas military operations. The Supreme Headquarters Allied Expeditionary Forces in Europe was also active in the IW space as was evinced by the successful subterfuge involved in concealing the actual location of the D-Day invasion.[59] In 1942, the US launched the Voice of America (VOA) to disseminate news to states in German occupied territory. The VOA was also used to spread American values and attempt to counter communist propaganda during the cold war.[60] During this era, the US also employed IW in the form of psychological and disinformation tactics in attempting to obtain a narrative advantage over the Soviet Union.[61] In later years, US Air Force Colonel John Boyd helped develop and solidify the notion of "information warfare" and argued that IW was not simply a way to spread disinformation or propaganda, but could also be used to a greater extent to assist the US in military and political activities due to the inherent value that emerged from utilising information in a particular manner.[62] In the 1960s and 1970s, the US employed information warfare tactics (psychological warfare specifically) in the Vietnam conflict.[63] In the 1980s, the US military, intelligence community, and US State Department began using computers and satellites as part of the US' IW efforts. In the 1990s, the US employed IW tactics against Saddam Hussein's regime in the Gulf War.[64]

The US and AI military application: information warfare

Current experts contend that the US does not have a clearly defined strategy regarding IWIO.[65] Part of this is related to the lack of a clear definition of IW by the US government, or single agency responsible for conducting IW for strategic advantage.[66] The US military is responsible for the US' IWIO. IWIO is gaining increasing recognition in the

US as an important and inter-related aspect of war. The US Department of Defense (DoD) published the Joint Publication (JP) 3-13, Information Operations in 2012 (i.e. Joint Doctrine), and updated it again in 2014. Though lacking a unified definition and doctrine regarding IWIO across the government, military, and civilian populations, the acknowledgment that "operations in and across land, sea, air, space, and the electro-magnetic spectrum … depend on … [and] create information" is becoming more widely recognised.[67] Regarding US military strategy and IWIO, there is one operational environment and three dimensions within it: the physical, informational, and human.[68] The US Army describes the physical as "connective infrastructure that supports the transmission, reception, and storage of information," and the cognitive as "the minds of those who are affected by and act upon information."[69] Taken as a whole, the US Army views IWIO as using information collected from the physical environment to influence an adversary's decisions. Similarly, the US Marine Corps manual describes it as "leveraging the power of information to influence the behaviour of others."[70] It should be noted that the US Army is currently out of line with the joint doctrine, because joint doctrine recognises the distinction of the information environment where the US Army does not.

The most recent publicly available doctrine guiding US IWIO is the US Army's ATP 3-13.1 Conduct of Information Operations document, published in 2018.[71] Regarding the US' IWIO strategy, or lack thereof, though there is doctrinal recognition that IWIO can be used offensively to influence others, the US appears more hesitant than states such as Russia or China to use such tactics against individuals. Thus, the US seems to approach IWIO from a more defensive posture seeking to protect itself from IWIO and respond to adversarial IWIO when targeted.[72] The US has traditionally viewed IWIO in military terms while attempting to differentiate between acceptable and unacceptable activities.[73]

Recently, the US Army has shifted from operating within the standard academic definitions of IW and has moved into what is being called Information Advantage and Decision Dominance (IA & DD).[74] Within this realm information advantage activities (IAA) are conceptualised as the condition of holding the information advantage over a relevant actor's behaviour, situational understanding, and decision-making by using all military capabilities.[75]

Within the US' IWIO strategy, one issue that places it at a disadvantage compared with states such as China is the lack of a centralised approach to data collection.[76] Outdated data collection directives, different agencies having disparate approaches to data collection processes, and the lack of inter-agency communication often produces redundant data collection efforts and unnecessary resource expenditures. In simpler terms, due to the lack of a centralised data collection process, US agencies often expend valuable resources collecting identical data, and unco-ordinated data collection processes can lead to difficulty analysing the data and producing actionable intelligence for IWIO.[77]

The US is applying AI in several areas related to IWIO. China and Russia's recent investments in IWIO operations have driven the US to prioritise AI research to bolster its own defensive and offensive IWIO capabilities. The focus appears to be on using AI for weapons systems, training purposes, and protecting networks and digital information from other states' IWIO attacks.[78] It should be noted that China has invested heavily in IWIO in lieu of AI technology to offset some of its military technological disadvantages compared to the US.[79]

Currently, AI is being applied by the US in a defensive standpoint to monitor large streams of data to detect information patterns which can be identified as hostile information campaigns and potentially countered.[80] McGrath and others believe the US can improve its own information operations through training via complex wargame simulations generated and adapted by machine learning programmes, and operational concept improvement bolstered by machine learning.[81] As McGrath[82] argues, this might help the US realise the goals of its Third Offset Strategy (TOS), which was adopted in 2015 with the aim of shifting the US military's mentality towards innovation and direct competition with other great power states in hopes of overcoming adversarial technologies in Russia and China. The TOS was announced in 2015 by Robert Work, Deputy Secretary of Defense.[83] The top technological priorities listed in TOS focused on: learning machines, human-machine collaboration, assisted human operations, human-machine combat training, and network-enabled autonomous weapons.[84]

The importance of expanding and utilising AI to protect against IWIO attacks is exemplified by the 2023 National Defense Authorization Act (NDAA). The NDAA "outlines the Pentagon's spending priorities" with "a $20 billion increase from the 2022 NDAA" going toward research and development for AI.[85] According to the 2023 NDAA, there is a five-year plan to apply AI to "warfighting cyber missions within DoD."[86] Department of the Navy CISO Tony Plater describes how "AI will impact vulnerability management, threat hunting, and boost network security … so [it] is highly sought after to help … secure … cloud services."[87] In this way, the DoD seeks to bolster its network defences against outside informational threats by increasing the areas in which AI is employed, as well as update how it is utilised to protect the information sphere. By expanding the use of AI in the IWIO space, the US seeks to protect itself from manipulation, privatise and compartmentalise its information, and secure its intellectual property from theft.

The US and AI diplomatic application against information warfare

The US Department of State (DoS) utilises AI in a multitude of ways both to inform diplomacy and protect the American public from IWIO tactics employed by other states. Recognising the issues posed within the information sphere, the DoS acknowledges that "competing strategically on a global stage demand[s] that data not only be produced, used, or stored, but leveraged as a strategic asset."[88] In a departmental first, an AI Use Case Inventory has been released. In it, the DoS reveals the multitude of ways AI contributes to national security regarding IWIO from "accessing and analysing large amounts of text data from Department reporting" to "countering disinformation."[89] The AI Inventory reveals that the Global Engagement Centre (GEC) is at the heart of many AI uses for operating against IWIO. For Disinformation Topic Modelling, the GEC uses "text clustering and topic modelling of documents and social media to determine possible disinformation subjects and topics" whilst image clustering is used to "identify similar images in order to predict likely disinformation."[90] Another way the DoS is using AI to combat disinformation is via a Deepfake Detector. This tool examines an image of a face "and classifies the image as either being real … or fake (synthetically generated face …) to predict disinformation activities."[91] Such programmes and tools could potentially help the American government recognise attempts to sow disinformation within the public more expediently, and adversarial IWIO tactics may be revealed in a timelier manner, thus increasing the likelihood of countering offensive information campaigns.

The pentagon and US cyber command's use of artificial intelligence for information warfare

In 2018, the Joint AI Centre (JAIC) was established by the Department of Defense to make use of AI and its potential as a valuable tool in the sphere of IWIO.[92] The DoD has five pillars of AI strategy: to "deliver AI-enabled capabilities that address key missions, scale AI's impact across DoD through a common foundation that enables decentralised development and experimentation, cultivate a leading AI workforce, engage with commercial, academic, and international allies and partners" whilst maintaining ethics and safety precautions.[93] One mission of the JAIC was to use AI to enhance joint warfighting efforts. In 2020, JAIC placed greater emphasis on ways to incorporate AI in the IWIO space.

By incorporating AI, the JAIC aimed to give the Department of Defense "an information advantage" by first refining its ability to combine commercial AI capabilities with government AI and then "improving the standardization of foundational DoD data needed to field high-performing AI-enabled capabilities to support operations in the information environment."[94] One programme the JAIC was using is the Multidimensional Anomaly Detection fusing HPC, Analytics, and Tensors (MADHAT). MADHAT "allows for the exploration of network data as a way of enabling more effective detection of nuanced adversarial threats."[95] By combining MADHAT's capabilities with established AI technology such as NLP and speech-to-text functions, the DoD aims to reduce the signal-to-noise ratio. When successful, using AI in this way allows analysts to devote their limited human resources to issues which require more nuanced interpretation rather than sifting through immeasurable data. JAIC was merged into the Chief Digital and Artificial Intelligence Office (CDAO) in 2022. Two of the primary goals of the newly formed CDAO are to: "1-Review and more tightly integrate the Department's policy, strategy, and governance of data, analytics, and AI, to include an integrated Data, Analytics and AI Strategy. 2-Provide the enterprise-level infrastructure and services that enable efforts to advance adoption of data, analytics, and AI, to include an expanded and more accessible enterprise data repository and data catalogue with designated authoritative data sources, common data models for enterprise and joint use cases, as well associated coding and algorithms to serve as a 'public good' as Department stakeholders put data on the offensive."[96]

In 2023, the CDAO reinstated experiments known as Global Information Dominance Experiments (GIDE) in collaboration with the Joint Chiefs of Staff (JCS). Members of the US military from all branches and civilian personnel made up the teams. The large-scale integration was made possible, in part, due to data and analytics connected to CDAO AI programmes. The most recent version of the GIDE (fifth iteration) included participation from combatant commands, the Pentagon, and international duty stations. The purpose of the GIDE was to provide information regarding Joint All-Domain Command and Control (JADC2) solutions pertaining to Joint data integration and AI and machine learning technology.[97] As Chief Digital and AI Officer Dr. Craig Martell stated, "We want to rapidly improve access to data across the Joint force – from the strategic level to our tactical warfighters. The intended outcome of these experiments is two-fold. First, we want to identify where we may have barriers in policy, security, connectivity, user-interface, or other areas that prohibit data sharing across the Joint force. Second, we want to show how data, analytics, and AI can improve Joint workflows in a variety of missions from global integrated deterrence through targeting and fires."[98]

Defense advanced research projects agency, artificial intelligence, and information warfare

The Defense Advanced Research Projects Agency (DARPA) aims "to a singular and enduring mission: to make pivotal investments in breakthrough technologies for national security."[99] DARPA is a collaborative effort between government employees and civilians with a storied connection to advancements across arenas from stealth technology to the Internet.[100] In April 2022, DARPA's Director Dr. Stefanie Tompkins stated the Department "is pursuing more than 39 programs that are exploring ways to advance the state-of-the-art in AI, pushing towards third wave contextual reasoning capabilities" while over "60 active programmes are applying AI in some capacity."[101] DARPA focuses on identifying and countering malicious deepfake technology, which uses AI to substitute one person's likeness for another in media such as photographs or videos. MediFor, DARPA's Media Forensics programme, "builds algorithms to detect manipulated images or videos, then produces a quantitative measure of integrity, which enables filtering and prioritization of media at scale."[102] The programme "uses detection algorithms, which analyse media content to determine if manipulation has occurred," as well as "fusion algorithms, which combine information across multiple detectors."[103] These algorithms contribute to an integrity score for each piece of data the programme analyses. A low score means the media was likely manipulated and is thus flagged for review by analysts – resulting in large volumes of media being analysed by AI, allowing analysts to concentrate their efforts when and where they are most needed. Though the MediFor programme is in its final stages, DARPA has a new programme called semantic forensics (SemaFor). Unlike MediFor, which focused on detecting discrepancies and anomalies in media, SemaFor aims to attribute and characterise these deepfakes.[104] SemaFor's semantic technologies "automatically analyse modal media assets to defend against large-scale, automated disinformation attacks" while its "attribution algorithms will infer if digital media originates from a particular organization or individual" and its "characterization algorithms determine whether media was generated or manipulated for malicious purposes."[105] These models may help bolster their deepfake defensive models which preserve individuals' facial expressions and how they move their head.[106] The defensive model would illuminate whether a video of a President, or dictator were legitimate while SemaFor could indicate who may be responsible for the particular deepfake episode. In this scenario, a deepfake (for example, a video of a world leader ordering the release of a nuclear, or biological weapon) could have serious ramifications for national security and international relations around the globe. Thus, the deepfake identification technology may play an important role in US AI-enabled IWIO defence.

The US, artificial intelligence, and information warfare overview

The US is employing AI in its overall IWIO strategy in numerous ways. The US is primarily focused on applying AI defensively rather than through offensive IWIO operations. This mirrors the US' overall IWIO strategy that is defensive in nature. Through collaborations with US technology companies and numerous government and military sectors, the US is using AI to identify, categorise, and counter a wide array of potential international IWIO threats. Examples include utilising AI technology designed to sift through large amounts of data to identify misinformation, propaganda, and intentionally divisive content that is intended to sow discord within the US domestically through social media and online content. Additional emphasis is being placed on using AI technology to counter AI-driven deepfake technology that could be used by adversaries for IWIO operations directed at the US. Furthermore, AI is being used to protect critical infrastructure from Cyber-attacks. This is being accomplished by employing machine learning programmes to sift through large amounts of data for indicators of possible attacks and generating AI programmes to defend against Cyber-attacks.

China

IW background information

Influenced by Sun Tzu and Mao Zedong, psychology is a central component of Chinese IW and is often employed as a key weapon rather than simply a support instrument.[107] Chinese IW is often conceptualised as consisting of "three warfares" that entail legal, psychological, and media operations. The aim of the warfares is to manipulate international legal regimes, affect public opinion, and undercut the morale of potential enemies. Within this framework, China employs IW operations pre-emptively. China often combines its IWIO tactics to include electronic warfare, precision-strikes, and Cyber-warfare with the goal being to injure the information capacity of its opponents.[108]

In engaging in IW, China incorporates Mao's notion of the "People's War" which consists of employing large amounts of Cyber-attacks combined with online disinformation. IWIO is a central component of China's military strategy given that China concedes it cannot match US military spending. China has placed significant emphasis on IW beginning as early as the 1950s, which has evolved into the current Strategic Support Force (SSF) and is a main component of China's IW capacity.[109] Numerous academies have been designed by China to expand China's IW capabilities, which include the Academy of Military Sciences Military Strategy Research Centre, the PLA Academy of Electronic Technologies, and the Xian Politics Academy. The Xian Politics Academy places a unique emphasis on psychological warfare training.[110] Researchers contend that China has employed IW simulation training for over a decade and IW units specialising in psychological warfare are embedded within the army.[111] Additionally, it is important to note that an important component of China's IW strategy includes operations in Cyber-space.[112] An example is the interconnected network of Chinese online influencers who reinforce Chinese narratives in countries that are targeted in Chinese IWIO.[113]

China actively employs its IWIO on social media. China utilises IWIO in its operations to attempt to weaken the perception of an enemy's leaders and its citizenry.[114] In a similar strategy as Russia, China employs psychological warfare to divide populations. This occurs through social media and by PRC agents purposely placed on social media platforms to propagate PRC narratives. Many of the programmes used by

PRC agents are AI-assisted. One example is China's use of the United Front, which is a sophisticated network of operators that carry out co-ordinated IWIO against specific individuals and institutions.[115] These actions allow the PRC to manipulate public narratives that are favourable to the party, domestically and internationally. China also controls online and social media content domestically to shape narratives and to ensure that it does not become the target of the type of influence campaigns it directs at adversary states.[116]

China and artificial intelligence military application: information warfare

China seeks to utilise information technology in a wide variety of sectors and regions, including "disruption through trade wars, information manipulation in cyberspace, and military integration of advanced technologies."[117] China created the Strategic Support Force (SSF) in 2015 with the aim of generating strategic advantages in the areas of space, Cyber-space, and the electromagnetic spectrum.[118] As Kania and Constello remark, "the SSF has integrated the PLA's capabilities for cyber, electronic, and psychological warfare into a single force within its Network Systems Department, which could enable it to take advantage of key synergies among operations in these domains."[119] China aims to implement an IWIO strategy that "focuses on 'informatization warfare,' or 'xinxihua,' the application of information technology to all aspects of military operations."[120] Daniels and Chang state that the government of China is actively "using AI technologies to enhance population control, as well as to profile and control its ethnic minorities."[121] They continue, stating that "China will likely export versions of these capabilities to authoritarian governments globally in the 2020s and 2030s, as it has already begun to do."[122] If social influencing can be altered and if "mass opinion can be decisively influenced by the clash between AI influence systems, for example, China may determine its best option for reabsorbing Taiwan is heavy investment in AI-empowered propaganda."[123] The integration of AI with nearly every facet of China's technology allows for specific advantages in social control and information management and is "enhanced with its 2017 Cyber-law that delivers unlimited avenues to virtually every network and piece of hardware operating in the Asia-Pacific."[124]

China has been accused of engaging in "cognitive warfare" against Taiwanese citizens by a Taipei think-tank and other observers in Taiwan.[125] As Taiwanese citizens, particularly the younger generation, have increasingly shifted away from China amid arguments that they have no connections to the mainland, China has engaged in "tactics ranging from military intimidation and propaganda to misinformation spread by its army of online trolls in a bid to manipulate public opinion."[126] Ultimately, this tactic is aimed at trying to coerce a reunification of Taiwan with mainland China without risking armed conflict.[127] This type of cognitive warfare falls within the realm of IWIO, particularly as the efforts seek to manipulate Taiwan's decision-making capacity.[128,129]

It has also been alleged that China has adopted the Kremlin's IWIO tactics "to highlight America's faults and weaponize the culture wars and identity politics currently buffeting the West,"[130] which some have alleged is a move "to distract away from Beijing's own rights abuses, including the internment of more than a million ethnic Muslim Uyghurs."[131] This marks a notable shift from China's previous methods of defending itself from accusations of human rights abuses, most of which involved pressuring foreign states to refrain from involvement in China's "internal affairs."[132] These new offensive tactics bare the hallmarks of an IWIO campaign, and the increasing incorporation of AI technology could intensity China's IWIO operations.

China recognises the potential AI holds as a facilitator for growth, disruption, and control in the information space. In 2017, China's "New Generation AI Development Plan elevated AI as a core priority, catalysing what has become a whole-of-nation strategic initiative."[128] AI falls under China's military strategy of "'intelligentised' warfare," which "is characterised by four key features: increased information-processing capabilities, rapid decision-making, the use of swarms, and cognitive warfare."[133] According to Chinese strategists, human cognition is the main battlefield in intelligentised warfare. A former deputy chief of staff of the PLA, Qi Jianguo, "stated that those who gain the upper hand in developing new-generation AI technologies will be able to control the lifeline of national security."[134]

The official paper of the PLA, the *PLA Daily*, published an article discussing how cognitive warfare could be employed to influence the PLA's opponents. First and foremost, "cognitive warfare is directed at human emotion" and "should focus on the use of … AI … to strike at 'cognitive gaps' between social groups, especially the alliance system of the 'strong power' (a euphemism for the US), to exploit contradictions in interests and perceptions between groups, and create division."[135] The US saw many cultural conflicts intensify in recent years between protests, heated election cycles, and dichotomous stances on Covid vaccines. The PLA is utilising AI to identify and target these fractures that could have significant ramifications for such intelligentised warfare. However, China does not singularly focus on using AI in its IWIO tactics directed against the US. China also plans to employ AI to monitor and control the information space as it pertains to Chinese citizens.

The 2022 China Internet Civilisation Conference was meant to bolster and encourage the People's Republic of China's ability to implement and increase internet authority and control within its borders. The Party secretary, Ye Zhenzhen, shared "that the State Key Laboratory for Communication Content Cognition … is working to develop cognitive computing applications to guide political direction, public opinion guidance, and values orientation into a 'national weapon in the digital era.'"[136] Zhenzhen implicates "the use of big data and AI" as a means to strengthen China's leadership and better understand the citizens.[137] Though the report and accompanying video were quickly deleted following massive public backlash and condemnation, China's Comprehensive National Science Centre in Hefei's researchers "claimed to have developed 'mind-reading' AI capable of measuring citizens' loyalty to the Chinese Communist Party."[138] According to the researchers, AI analysed facial expressions and brain waves, thus measuring viewers' reactions – both positive and negative – to political information.[139] Though this specific publication was deleted, the US Department of Commerce did add the Academy of Military Medical Sciences (AMMS) in China, along with nearly a dozen of its research institutes, to its Entity List "based on the body of information that AMMS and its eleven research institutes use biotechnology processes to support Chinese military end uses and end users, to include purported brain-control weaponry."[140] The Entity List, though originally focusing on items relating to WMDs, now also serves to notify the public of "activities contrary to U.S. national security and/or foreign policy interests." Thus, the AMMS's inclusion on a list for potential cognitive-monitoring and control tactics, and the potential to achieve its stated cognitive warfare goals, is concerning to many observers.[141]

China: artificial intelligence, information warfare, and the Uyghurs

Analysing China's manipulation of the information circulating regarding the treatment of Uyghurs in the Xinjiang region reveals a myriad of ways the state is utilising AI for IWIO purposes. It is alleged that China has detained "more than one million Uyghurs against their will" whilst many others in this majority Muslim community have been imprisoned.[142] In August 2022, the US Department of State released a report stating that "the People's Republic of China (PRC) actively attempts to manipulate and dominate global discourse on Xinjiang" in multiple ways via the internet and social media.[143] Of particular interest is their means of downplaying negative reports on the treatment of the Uyghur population while magnifying more positive, fabricated stories.

The Department of State report explains that "the PRC floods conversations to drown out messages it perceives as unfavourable to its interests on search engines and social media feeds."[144] Researchers analysed how often Chinese state media appeared in search results for key terms relating to Xinjiang and Covid over a four-month period for Google Search, Google News, YouTube, Bing Search, and Bing News.[145] They found that over the course of one hundred and twenty days, "Chinese state media featured prominently" in search engine results with "21.5% of the top results on Google News and Bing News" and a quarter of YouTube's results featuring state-backed media and accounts.[146] Simply searching a neutral term such as "Xinjiang … returned Chinese-state media in top results in 88% of News searches and 98% of YouTube searches."[147] By matching text and headlines word-for-word, nearly three dozen additional sources regurgitated Chinese state media reports – their inclusion in the report would have increased Chinese state influence by almost ten percent while YouTube videos posted by confirmed Beijing-supported users would add an additional twenty-seven percent of search results.[148] With "AI power[ing] almost every part of a search engine" and every single search result produced being "a direct result of decisions made by AI," some researchers contend that China's use of AI to manipulate information in the international arena must be examined more closely.[149]

Additionally, supporters of the PRC's IWIO mission also engage in astroturfing to promote more positive stories of what is happening to the Uyghurs in Xinjiang.[150] The term astroturfing describes "coordinated campaigns of inauthentic posts to create the illusion of widespread grassroots support for a policy, individual, or viewpoint, when no such widespread support exists."[151] The PRC accomplished this by using bots to spread quickly videos of content, such as the portrayal of happy Uyghur citizens on social media. When the *New York Times* and *ProPublica* analysed thousands of videos in 2021, they discovered numerous signs of astroturfing. Though "most of the clips carry no logos or other signs that they are official propaganda," analysis of over three thousand videos "found evidence of an influence campaign orchestrated by the Chinese government."[152] Most of the videos were shared on Chinese apps, but then began appearing other apps such as Twitter and YouTube – with English subtitles. All the videos possess similar or identical messaging, words, and phrases claiming that the Uyghur citizens being filmed were happy, prosperous, and free. In over one thousand of the videos, the people "say they have recently come across [Former Secretary of State Mike] Pompeo's remarks" regarding their treatment and that his declarations they are oppressed, and genocide is taking place, are "complete nonsense."[153] Numerous aspects of the videos indicate their scope and reach was propelled by Chinese AI technology. *ProPublica* and *The Times* discovered "the clips were shared by more than 300 accounts whose posts strongly suggested they were no ordinary users" due to the identical messaging save "for a random string of characters at the end with no obvious meaning."[154] The random characters being generated were meant to circumvent anti-spam filters employed by Twitter to identify such bots. The random characters were found in seventy-five percent of the tweets. Additionally, every account had been recently created, did not follow other accounts, had few – if any – followers, and most of the tweeting occurred during the daytime in Beijing.[155] Of particular importance is the fact that "the text of several of the accounts' tweets contained traces of computer code, indicating that they had been posted, sloppily, by software."[156]

The CCP has also used popular, female, minority social media influencers to spread CCP propaganda in Xinjiang, Tibet, and Inner Mongolia. When examining 1,741 videos published on 18 popular YouTube Accounts researchers found that the influencers propagated the CCP narrative that political, economic, and social conditions were ideal in these regions and rejected or ignored any human rights concerns.[157] Researchers contend that the influencers were likely manipulated by 'professional user-generated content,' or content that's produced with the help of special influencer-management agencies known as 'multi-channel networks (MCNs).'[158] These MCNs are directly controlled and funded by the CCP and are designed to propagate the CCP's narrative. These videos are often prioritised on search engines because the users generate a large amount of reoccurring posts and AI search-engine algorithms prioritise users that post frequently. Thus, posts from non-CCP affiliated users in these regions, which often raise genuine human rights concerns, are given lower priority by AI-search engine algorithms and the posts are viewed less often because the users are not able to post with the same volume and frequency as the MCN-assisted creators. Additionally, since YouTube is blocked in China, non-CCP affiliated social media creators cannot monetise social media content on platforms like YouTube where the MCNs can, due to their special agreements with China, thus providing the CCP with greater means to disseminate its propaganda.[159]

China: artificial intelligence, information warfare, and Hong Kong

Following a wave of protests and demonstrations in Hong Kong opposing China's new extradition law, the PRC media began to spread false narratives to attempt to delegitimise the Hong Kong protestors and portray them as participating in an independence, or separatist movement. It was discovered that numerous fake accounts were generated by the PRC to amplify the PRC's narrative that the demonstrators were violent separatists. The fake accounts produced large amounts of misleading information across numerous social media platforms including Twitter, Facebook, and YouTube. In August 2019, Facebook suspended 7 pages (with approximately 15,500 account followers) and 3 groups (with approximately 2,200 account followers). Additionally, Twitter suspended 200,000 accounts and You Tube suspended 210 channels related to

PRC misinformation efforts regarding the Hong Kong demonstrators. Furthermore, in 2020, Twitter suspended 23,750 main accounts and discovered that approximately 150,000 social media accounts were created to amplify the misleading content of the main accounts.[160] The suspicious accounts were identified due to the accounts reinforcing pro-PRC narratives and the activity of the accounts surged at the same time as the PRC began its propaganda campaign against the Hong Kong demonstrators. Additionally, many of the accounts did not have any followers and many account users claimed to be located in Hong Kong, but the account locations were set in other countries. After removing the suspicious accounts, Twitter announced that the suspended accounts were attempting to "sow political discord in Hong Kong" by "undermining the legitimacy and political positions of the protest movement on the ground."[161] As with the IWIO tactics employed by China regarding Taiwan and the Xinjiang region, AI algorithms were likely involved in the bot activity pertaining to the Hong Kong protests in respect to the content shared, the frequency of postings, and attempts to evade spam detection protocols.

China, artificial intelligence, and information warfare overview

China is using AI in the area of IWIO through multiple channels. China has incorporated AI into its offensive IWIO strategy by attempting to increase social and political tensions and divisions in the US through social media. China has also used AI to attempt to manipulate public sentiment in Taiwan, international opinion regarding the Hong Kong demonstrations, and the international community's perception of the treatment of the Uyghurs in Xinjiang. To accomplish these objectives, China has appeared to use AI to eliminate negative press while manipulating information by filming propaganda videos, disseminating the videos globally, and employing AI to circumvent spam detectors while flooding social media platforms with misinformation. These tactics indicate that China is willing to take aggressive actions to control political narratives and utilise AI to achieve its IWIO goals. China has also increasingly used AI to surveil its domestic population and spread political propaganda that is favourable to the PRC within its borders.

Russia

IW background information

Russia, since the fall of the Berlin Wall, has sought to revitalise "traditional values at the individual level and a focus on returning the glory of the Soviet Union on the national level."[162] To do so, Russia has used information and technology as part of its IW approach, where the purpose of such warfare when directed at adversary states is "to divide and polarise society, tear it into small pieces and fragments, and make these fragments sincerely hate each other in order to have them collide with each other thereby initiating a fight for destruction or combine their aggression into a uniform stream and direct it against the ruling government."[163] As Wilde and Sherman remark, "its core tenet might well be that regime security has historically been indivisible from information warfare in Russian strategic thought. Rather than an aggressive, or expansionist, expression of Moscow's foreign policy, the Kremlin's so-called information war should primarily be viewed through a domestic and regime security prism – it's as much a counterinsurgency as an expeditionary strategy, less an escalation than a projection."[164] Some of the common IW techniques employed by Russia include disinformation, propaganda, and psychological operations. Among the more well-known Russia IW operations, Russia has allegedly employed IW to influence elections in the US, France, and Germany, and has been accused of deploying IW to aid the Russian military in Syria and Ukraine.[165]

Topor and Tabachnik explain that the focus of Russia's IWIO strategy is to use internet connected technology to "undermine, manipulate, and mislead the information people consume as it believes this can advance its political and military objectives."[166] The key to this style of warfare is the creation of unsecured or permissive information spaces, "wherein discourse or debate lines favourable to Moscow permeate a targeted society."[167] Ajit and Vailliant state that the use of IWIO is nothing new to Russia, where the "first known use of the words 'active measures' was in a Bolshevik document in 1919."[168] The use of manipulating, influencing, and controlling information has been a constant tool used by all versions of Russia throughout the most recent century.

From a constructivist viewpoint, Russia perceives itself as a disrupter. Since the 1970s, Russia strategists have been considering how the Digital Age would affect warfare and society. Russia has long considered the digital information age as a new type of battlefield where information can be wielded as a weapon. However, Russia has realised it cannot compete commercially in the digital space with other western states such as the US. Thus, it has employed a strategy of disruption, denial, and delay regarding IWIO. This strategy has included Cyber-warfare and influence operations, especially disinformation. The ultimate aim of the strategy is to undermine public confidence in the US and western political systems through the surveillance capitalism model.[169]

An important element to consider in relation to Russia, AI, and IWIO, is how technologies such as AI, as well as globalisation and changing economic landscapes, affect cultural backlash in western democracies, and how the configuration of these factors impacts the types of IWIO Russia employs as well as the ultimate success of Russia's IWIO strategies. One aspect of cultural backlash theory is based on how some individuals in western states may become disenchanted with the erosion of traditional ideals and beliefs and the emergence of more progressive and secular trends, thereby increasing their political grievances and support for populism.[170] A second aspect is centred on possible grievances that emerge in western states due to rising economic inequality tied to changes that transform economic patterns and labour markets, which could also increase grievance formation and support for populism.[171] In considering these potential economic changes, researchers have noted that technologies such as AI can affect labour markets[172] leading to possible increases in social, economic, and political divisions and potentially greater instability and support for populist movements.[173] A potential effect of cultural backlash, whether driven by social, economic, or technological factors, is that states such as Russia can more easily deploy AI driven IWIO campaigns to target individuals and groups that are discontent, leading to greater societal divisions, polarisation, and support for populist movements. Having examined how historical factors and Russia's state identity affect its IWIO, we now turn to examining how Russia is applying AI in its IWIO strategies and tactics.

Russia and artificial intelligence military application: information warfare

Despite Russia's AI developments, Russia is currently lagging behind the United States and China in terms of incorporating AI technology into its military overall.[174] However, Russia has demonstrated an intense focus on further developing its already advanced IWIO tactics with the assistance of AI technologies. This is likely due to Russia's strategic focus on IWIO as a primary security strategy. Russia's internet sponsored propaganda manufacturing facilities, or "troll farms," are now equipped with AI powered Deepfake technology that can create more realistic false narratives by constructing fake images and even video clips involving key figures that support whatever narrative the "troll" is attempting to push.[175] This software has made it possible to create ultrarealistic depictions of events that never happened. Experts have stressed the massive risk posed if Russia begins doctoring images and videos for political gain.[176]

Russia's view of IWIO covers a wide swath of technology, where "jamming electronic communication and disrupting access to the electromagnetic spectrum, Cyber-espionage, and distributed denial of services (DDoS) attacks are no different from (and work in tandem with) using trolls and bots to spread dis/misinformation, establishing pro-Russian media outlets, or supporting local sympathisers to propagate favourable messages."[177] Russia exploits information ecosystems by "interjecting dis/misinformation (partially attained through Cyber-attacks), and fake news stories that a majority of those exposed to believed true at the time."[178] Faking and altering digital materials can be used in many scenarios, including political ones, as shown when "during the 2017 French election, Russia stole documents from the Macron campaign and edited them to include fake, damaging information."[179] Botnets, trolls, and deepfakes are tools often utilised in the information space with decent success rates, so much so that "Russian authorities have set up the so-called 'Internet troll factory' in St. Petersburg – young people who pretend to be real members of the Internet, widely concentrating and disseminating provocative and outrageous information."[180] O'Donnell provides another example of Russian disinformation, stating that the "Russian Internet Research Agency has launched sophisticated campaigns to create the appearance of a chemical disaster in Louisiana and an Ebola outbreak in Atlanta."[181] The nature of AI-assisted deep fake technology has "accentuated perceived differences between the realities of partisan groups and accelerated the prevalence of, and discussion on, 'fake news.'"[182] The use of AI-assisted deep fake technology can even create physical, real-world events, as seen when Russia "successfully organised a fake protest prior to the 2016 election that was attended by thousands of people in New York and another in Florida."[183]

In present times, unlike competitors, Russia does not differentiate between technologies and information, and instead of calling the digital-only system Cyber-space, refers to it "as the 'information space,' which includes both computer and human information processing."[184] This integrated viewpoint allows Russia to command a hybrid information and digital technology suite with real-world applications such as "the recycling and spreading of a YouTube video of Russian soldiers with the title 'Punitive Ukrainian National Guard Mission' throwing dead bodies near Kramatorsk (Donetsk region) on 3 May 2014."[185] Not just regulated to using combat footage to influence viewers on the World Wide Web, Russia has used social influencing and communications to sway public opinion when Russian agents tweeted "pundits call on @Theresa_May to disrupt possible Russia-US thaw. No trust in Britain's best friend and ally?"[186] Further IWIO tactics involve stories shared on social media, where prior to the Netherland's 2016 trade deal referendum with Ukraine, Russia subjected Dutch citizens to online articles consisting of "Ukrainian soldiers crucifying a child and reports from individuals, purporting to be experts, portraying Ukraine as a 'bloodthirsty kleptocracy, unworthy of Dutch support.'"[187] By using various AI-assisted methods and tools in the information space, Russia can influence consumers of web-connected systems, public figures, and private citizens alike.

Russia has been flexible with its AI applications, often deploying the technology "in situations that may not constitute either war or peace," commonly referred to as the "grey zone."[188] Cyber-warfare, electronic warfare, influence operations, propaganda campaigns, and disinformation are prime examples of instruments that fit the Russian models of AI.[189] Russia has increased the use of AI in the digital world and "cyberwarfare, electronic warfare (EW), influence operations, propaganda campaigns, and disinformation are prime examples of instruments that fit the Russian modus operandi and are ripe to be integrated with AI."[190]

Though there are reports that Russia is lagging behind the US in military-AI integration, it is important to note that "unconventional tools – Cyber-attacks, disinformation campaigns, political influence, and illicit finance – have become a central tenet of Russia's strategy toward the West and one with which Russia has been able to project power and influence beyond its immediate neighbourhood."[191] By using AI in the information sphere, Russia can significantly improve the scope of their IWIO campaigns. Polyakova's assertion that "unlike in the conventional military space, the United States and Europe are ill-equipped to respond to AI-driven asymmetric warfare (ADAW) in the information space" requires serious consideration by policymakers (Polyakova 2018). With Russia trailing the US in integrating AI into the military, it is understandable that Russia would focus on ADAW as asymmetric warfare involves "conflicts between nations or groups that have disparate military capabilities and strategies."[192] Additionally, the Russian state utilises its AI IWIO capabilities at home as much as they do abroad.

In 2016, the Yarovaya amendments were instituted. These Russian laws "required telecom providers, social media platforms, and messaging services to store user data for three years and allow the FSB access to users' metadata and encrypted communications."[193] Although there is no consensus or knowledge about what Russia wants with such data, "their very collection suggests that the Kremlin is experimenting with AI-driven analysis to identify potential political dissenters."[194] Additionally, in Moscow, officials are using AI facial recognition systems called Sfera to target and surveil journalists.[195,196] By utilising such surveillance and biometric data, "the system has seen the preventative detention of dozens that the regime suspects to be potential instigators of public unrest." If IWIO is evaluated according to an entity using capabilities "to influence, disrupt, corrupt, or usurp the decision making of [target audiences]," then Russia's detention and intimidation of journalists from independent outlets can influence and disrupt the information Russian citizens obtain from independent journalists.[197]

Russia, artificial intelligence, information warfare, and Ukraine

Russia uses AI to conduct influence campaigns against citizens to sow civil discord or garner support for their own actions (such as the war in Ukraine). On February 15th and 16th, nine days prior to the Ukrainian invasion, it is alleged that Russia carried out Distributed Denial-of-Service (DDoS) attacks targeting Ukrainian banks, government websites, and the Ukrainian Ministry of Defence and Foreign ministry. In conjunction with the DDOS attacks, Ukrainians began receiving SMS spam messages containing disinformation that indicated Ukrainians could not withdraw funds from ATMs due to technical issues.[198] In addition, numerous Russian disinformation campaigns have been identified and reported by social media companies since the Ukrainian invasion began. The Russian social media disinformation campaigns have directed inauthentic behaviour on social media platforms, temporarily seized control of social media channels, and sought to compromise the integrity of social media accounts.[199] A specific example was in September 2022 in which Meta removed a large network of fake accounts impersonating major news outlets publishing pro-Kremlin articles. These articles "[accused] the Ukrainian government and military of corruption and warning of dire consequences from European sanctions on Russia."[200] The report indicates "many of the fake accounts used profile pictures generated by AI."[201] Twenty-three hundred accounts were removed. These accounts, their pictures, and the websites created for the fake news stories, may have been the result of generative adversarial networks (GANs). This "branch of AI can be trained to produce realistic-looking data … [and] can disseminate that disinformation like rapid fire, while at the same time tracking its performance online by counting clicks and engagement."[202] GAN is the same software that can produce deepfakes and is a concerning area for those seeking to combat AI-enabled IWIO tactics employed by states such as Russia.

Russia uses GANs in its IWIO tactic in many areas as the Kremlin's Internet Research Agency (IRA) "is becoming increasingly decentralised and is gaining 'incredible traction' on TikTok with misinformation aimed at sowing doubt over events in Ukraine."[203] The IRA has a history of using trolls to post online and/or create bots that can spam social media sites with repetitive messaging. In May 2022, officials with the United Kingdom's government revealed that recent "research suggested Moscow's operation was 'designed to manipulate international public opinion' in favour of its military campaign in Ukraine."[204] Social media sites such as TikTok, Facebook, and Instagram attempt to remove accounts posting inconsistently with legitimate, non-bot users, but it can be difficult to keep up.[205] Though Twitter reports removing 100,000 accounts "for violations of its platform manipulation and spam policy" between February to May 2022, these types of information operations are becoming more common, gaining momentum, and appear more authentic than ever before.[206] The Ukrainian Secret Service confirmed in March 2022 that it had neutralised five different bot farms that were spreading disinformation on more than 100,00 active social media accounts.[207]

Russia, artificial intelligence, and information warfare overview

Russia, similar to China, utilises AI in its IWIO tactics to increase domestic tensions in the US and other Western states, as well as to divide and confuse antagonistic agencies and organisations. The overarching view of IWIO by Russia is "that information is the most important object of operations, independent of the channel through which it is transmitted."[208] As analysis and co-ordination of information and data from disparate channels requires exponential effort from IWIO analysts, Russia is researching the application of AI to IWIO data streams. Bendett explains that the focus of Russia is on merging separate sectors into a unified front, stating that "many public efforts originate from the Russian Ministry of Defence (MOD), which is dedicating financial, human, and material resources toward AI development across its vast technical, academic, and industrial infrastructure."[209] Additionally, many technology events have been hosted by Russia over the past few years as they seek to merge AI research and application with existing IWIO implementations. These IWIO and AI events include "the 2018 Intellectual Systems in Information Warfare symposium" as well as workshops held on a regular basis by the "Russian AI Association."[210]

In summary, Russia has pursued aggressive IWIO operations targeting the US and other democracies ranging from seeking to manipulate elections, to spreading politically motivated deepfake videos, to attempting to increase political and societal polarisation within states. Russia has devoted more of its energy and resources to utilising AI in its overall IWIO strategy compared with the US and China. This may be an extension of Russia's overall military doctrine that places greater emphasis on offensive IWIO tactics compared with the US.

Discussion

AI can play a significant role in affecting IWIO strategies and tactics. AI can significantly enhance capabilities for automating IWIO operations, especially in reaching mass audiences and influencing public perceptions. AI can affect IWIO by increasing the speed of IWIO operations and it can be applied to a wide range of IWIO applications. As AI technology continues to evolve, its influence on IW tactics and techniques will undoubtedly grow and affect the landscape of modern security competition. This study has found that the US, China, and Russia are applying AI in their IWIO strategies and tactics in unique and impactful ways.

The United States is applying AI in its overall defensive IWIO strategy through numerous techniques. The US is utilising AI in many governmental and military areas the better to identify and counter IWIO threats pertaining to disinformation spread over social media and through other online channels. Specific AI applications seek to identify particular texts, themes, images, and videos that are part of foreign governments' IWIO operations. The aim is to reduce threats that seek to spread misinformation, propaganda, intensify polarisation and division within the population, and increase discontent with the government. While the US is incorporating AI in many different sectors within its defensive IWIO framework, frequent discussions within the government are centring around whether the US should continue to approach IWIO from a defensive posture, or advance a more offensive approach, as illustrated in the new doctrine of IA & DD.[211] If the US decides to adopt a more offensive-minded IWIO strategy, it possesses the technical sophistication to incorporate AI in its operations in several potentially effective ways based on pre-existing technology that can be adapted for offensive tactics. The decision by government leaders to pursue a more aggressive IWIO strategy may ultimately be determined by whether, and to what degree, Russia and China continue to target the US and its allies in future IWIO operations.

China is incorporating AI in its offensive IWIO operations through multiple avenues. China is using AI algorithms to spread information on social media to highlight divisions and political tensions in democracies as part of its overall "divide and conquer" strategy.

In addition, China is using AI in its cognitive warfare tactics to attempt to manipulate public opinion in Taiwan regarding reunification. This is being done in part through AI-powered programmes and bots that target Taiwanese citizens through the spread of misinformation and propaganda on social media. China is employing similar strategies internationally in its efforts to manipulate global opinion regarding the Uyghur situation in the Xinjiang region. China has mounted significant international IWIO operations ranging from AI bots generating misleading content on social media, to altered videos depicting the treatment of Uyghurs in China, to the manipulation of propaganda information posted on social media to evade AI misinformation and spam detectors on social media platforms. China is also advancing initiatives to further incorporate AI into the monitoring and control of its domestic population through expanding surveillance techniques and propaganda messaging. Concerns exist that China will export these AI technologies, designed to control and manipulate domestic populations, to other authoritarian regimes in the coming years.

Russia, employing a similar strategy as China, has used AI in its IWIO in attempting to sow political discord in several democratic states. Russia has relied on many different AI technologies to spread disinformation against its perceived adversaries in hopes of internally weakening those states. AI algorithms, bots, and deepfake technology have been employed to undermine the functioning of targeted governments. AI has also assisted Russian IWIO in identifying targets (e.g. specific citizens and groups) within democracies for precisely tailored propaganda messaging. Russia is also actively incorporating AI in its IWIO operations regarding Ukraine by attempting to manipulate international public opinion on the Ukrainian invasion using misinformation, bots, and altered videos that are AI driven. Lastly, Russia is likely to be employing AI technology to monitor journalists and potential opposition groups in Russia in a larger effort to minimise public dissent regarding the Ukrainian invasion and Putin administration.

In summary, AI is playing a pivotal role in affecting the IWIO tactics employed by the US, China, and Russia. Each state's overarching IW strategy is guiding the types of IWIO deployed. AI is providing states with a greater spectrum of possible tactics ranging from more complex IWIO detection software, to more powerful misinformation techniques (e.g. social media propaganda, deep fake videos, and bot proliferation), and increased capacity to conduct domestic surveillance and manipulate public opinion, both domestically and internationally. AI is significantly expanding the types of IWIO states can employ and altering the existing IWIO landscape. Given the growing sophistication of AI-supported IWIO tools, states will have to decide what types of IWIO strategies and tactics most appropriately match their values and maximise their security.

Notes

1. K. Khan, 'Understanding IWIO and Its Relevance to Pakistan', *Strategic Studies* 32 (2012): 138–159, p. 138.
2. W. R. Gery, S. Lee, and J. Ninas. 'Information Warfare in an Information Age', *Joint Force Quarterly* 85, no. 2 (2017): 22–9, p. 24.
3. Ibid., 24.
4. G. Yan, 'The Impact of AI on Hybrid Warfare', *Small Wars & Insurgencies* 31, no 4 (2020): 898–917.

5. J. S. Hurley. ‘Enabling Successful AI Implementation in the Department of Defense’, *Journal of IWIO* 17, no. 2 (2018): 65–82, p. 65.
6. M. Bishop and E. Goldman, ‘The Strategy and Tactics of IWIO’, *Contemporary Security Policy* 24, no. 1 (2003): 113–39.
7. L. Y. Hunter, C. D. Albert, C. Hennigan, and J. Rutland, ‘The Military Application of AI Technology in the United States, China, and Russia and the Implications for Global Security’, *Defense and Security Analysis* 39, no. 2 (2023): 207-232.
8. C. Perez and A. Nair, ‘Information Warfare in Russia’s War in Ukraine: The Role of Social Media and Artificial Intelligence in Shaping Global Narratives’, *Foreign Policy* (August 2022). https://foreignpolicy.com/2022/08/22/information-warfare-in-russias-war-in-ukraine/
9. Congressional Research Service (CRS) Renewed Great Power Competition: Implications for Defense – Issues for Congress (2022); A. Radin et al., ‘China-Russia Cooperation. Determining Factors, Future Trajectories, Implications for the United States’, *RAND Corporation* (2021).
10. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (Cambridge: Massachusetts, MIT Press, 2016); I. A. Joiner, *Emerging Library Technologies: It’s Not Just for Geeks* (Chandos Publishing, 2018).
11. Y. Kumar, K. Kaur, and G. Singh, ‘Machine Learning Aspects and Its Applications Towards Different Research Areas’, (2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM), 2020), 150–6.
12. Ibid., 150.
13. A. Alimadadi and others, ‘Artificial Intelligence and Machine Learning to Fight COVID-19’, *Physiological Genomics* 52, no. 4 (2020): 200–2; Murphy, K. P. Murphy, *Probabilistic Machine Learning: An Introduction* (Cambridge: MA, MIT Press, 2022).
14. Booz Allen Hamilton, ’AI for National Security’. (2022). https://www.boozallen.com/markets/intelligence/ai-for-national-security.html (accessed September 27, 2022).
15. Pimentel, Elias, ‘2nd MAW Marines Train Using Video Games’. US Government: Marines. (2022). https://www.2ndmaw.marines.mil/News/Article-View/Article/3125552/2nd-maw-marines-train-using-video-games/
16. US Department of State, ’AI Inventory’ (2022), https://www.state.gov/data-strategy/ai_inventory/ (accessed September 15, 2022).
17. G. Wilde and J. Sherman, ’No Water’s Edge: Russia’s Information War and Regime Security’. Carnegie Endowment for International Peace. (2023). https://carnegieendowment.org/2023/01/04/no-water-s-edge-russia-s-information-war-and-regime-security-pub-88644
18. M. Harré, T. Bossomaier, and A. Snyder, ‘The Development of Human Expertise in a Complex Environment’, *Minds and Machines*, 21 (2011): 449–64; H. Honda and M. Hagara, ‘Question Answering Systems with Deep Learning-Based Symbolic Processing’, *IEEE Access* 7 (2019): 152368–78.
19. B. Nakayama, ‘Democracies and the Future of Offensive (Cyber-Enabled) Information Operations’, *Cyber Defense Review* 7, no. 3 (2022): 49–65, p. 50.
20. C. A. Whyte, T. Thrall, and B. M. Mazanec, *IWIO in the Age of Cyber Conflict* (Oxfordshire: UK, Routledge, 2020, 344).
21. Ibid.
22. US Army, U.S. Army Techniques Publication No. 3-13.1: *The Conduct of Information Operations*. Headquarters, Department of the Army (2018), p. 1–1.
23. C. A. Theohary, ‘IWIO: Issues for Congress’, *Congressional Research Service* (2018): 7–5700.
24. H. Lin, ‘Russian Cyber Operations in the Invasion of Ukraine’, *The Cyber Defense Review* 7, no. 4 (2022): 31-46, p. 166.
25. CRS, Renewed Great Power Competition: Implications for Defense—Issues for Congress (2022)
26. V. W. Wang, ‘Asymmetric War? Implications for China’s IWIO Strategies’, *American Asian Review* 20 (2002): 167–207.

27. S. Tzu, *The Art of War*, trans. Sammeul Griffith (London: Duncan Baird Publishers, 2005), 108. (Original work published 5th century BC).
28. M. Wojnowski, *Russian Interference in the U.S. Presidential Elections in 2016 and 2020 as an Attempt to Implement a Revolution-like IWIO Scheme. Part II* (Warsaw Institute, 2021).
29. C. Cunningham, 'A Russian Federation IWIO Primer', The Henry M. Jackson School of International Studies, University of Washington (2020). https://jsis.washington.edu/news/a-russian-federation-information-warfare-primer/ (accessed March 5, 2022).
30. Cunningham, 'A Russian Federation IWIO Primer'; Wang, 'Asymmetric War? Implications for China's IWIO Strategies'.
31. US Army, U.S. Army Techniques Publication No. 3-13.1
32. Cybersecurity and Infrastructure Security Agency (CISA), 'Chinese Ministry of State Security-Affiliated Cyber Threat Actor Activity', (2020), https://www.cisa.gov/news-events/cybersecurity-advisories/aa20-258a; G. Wilde and J. Sherman, 'No Water's Edge: Russia's Information War and Regime Security'. Carnegie Endowment for International Peace (2023), https://carnegieendowment.org/2023/01/04/no-water-s-edge-russia-s-information-war-and-regime-security-pub-88644
33. B. Nakayama, 'Democracies and the Future of Offensive (Cyber-Enabled) Information Operations'.
34. C. D. Albert and others, 'Weaponizing Words: Using Technology to Proliferate Information Warfare', *Cyber Defense Review* (2023); D. Morabito, 'National Security and the Third-Road Threat: Toward a Comprehensive Theory of Information Warfare', *Air & Space Power Journal* 35, no. 3 (2021)*:* 19–39.
35. M. Ajir and B. Vailliant, 'Russian IWIO: Implications for Deterrence Theory', *Strategic Studies Quarterly* 12, no. 3 (2018): 70–89; C. Francois and H. Lin. 'The Strategic Surprise of Russian Information Operations on Social Media in 2016 in the United States: Mapping a Blind Spot', *Journal of Cyber Policy* 6, no. 1 (2021): 9–30.
36. E. B. Kania and J. K. Costello, 'The Strategic Support Force and the Future of Chinese Information Operations', *The Cyber Defense Review* 3 no. 1 (2018): 105–21; R. Diresta and others, *Telling 'China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives* (Stanford International Observatory. Cyber Policy Center: Hoover Institution, 2020).
37. C. Botan, 'Grand Strategy, Strategy, and Tactics in Public Relations', *Public relations Theory II* (Oxfordshire, Routledge, 2006).
38. J. Dawson, 'Who Controls the Code, Controls the System: Algorithmically Amplified Bullshit, Social Inequality, and the Ubiquitous Surveillance of Everyday Life', *Sociological Forum* (2023); J. Dawson and T. Wheeler, *How to Tackle the Data Collection Behind China's AI Ambitions* (Brookings, 2022). https://www.brookings.edu/techstream/how-to-tackle-the-data-collection-behind-chinas-ai-ambitions/
39. See Note 34; S. Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics', *The Defense Horizon Journal* (2022).
40. S. Zuboff, *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (New York: NY, Public Affairs, 2020).
41. J. Dawson, 'Microtargeting as Information Warfare', *Cyber Defense Review* (2021), https://cyberdefensereview.army.mil/CDR-Content/Articles/Article-View/Article/2537110/microtargeting-as-information-warfare/; Dawson, 'Who Controls the Code'.
42. Dawson, 'Microtargeting as Information Warfare', 63.
43. See Note 34.
44. Ibid.
45. K. Strittmatter, *We have been Harmonized: Life in China's surveillance state* (New York: NY, HarperCollins, 2020).
46. See Note 37; Strittmatter, *We have been Harmonized: Life in China's Surveillance State.*
47. Dawson, 'Microtargeting as Information Warfare'; J. Leonard, S. Mohsin, and D. McLaughlin, 'Tencent's Gaming Stakes Draw U.S. National Security Scrutiny', (Bloomberg, 2020). https://www.bloomberg.com/news/articles/2020-09-17/tencent-s-game-investments-draw-u-s-national-security-scrutiny#xj4y7vzkg

48. Dawson, 'Microtargeting as Information Warfare', 69.
49. Dawson, 'Who Controls the Code'; K. Goldsmith, 'An Investigation into Foreign Entities Who Are Targeting Service members and Veterans Online', *Vietnam Veterans of America* (2019).
50. See Note 34.
51. 116th Congress, Report of the Select Committee on Intelligence United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election. Senate Report. 116–XX (Washington, DC: United States Senate Intelligence Committee, 2017); Dawson, 'Microtargeting as Information Warfare'.
52. 116th Congress, Report of the Select Committee on Intelligence United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election; Dawson, 'Who Controls the Code'.
53. See Note 30.
54. Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'.
55. Ibid.
56. See Note 30; Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'.
57. J. Costello and J. McReynolds, 'China's Strategic Support Force: A Force for a New Era', *China Strategic Perspectives* 13 (2018); Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'.
58. M. Kaminska, 'Restraint under Conditions of Uncertainty: Why the United States Tolerates Cyberattacks', *Journal of Cybersecurity* 7, no. 1 (2021). https://doi.org/10.1093/cybsec/tyab008; Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'.
59. Thomas Cripps and David H. Culbert. *Information Control and Propaganda: Records of the Office of War Information* (University Publications of America, 1986).
60. Voice of America, 'History of VAO', (2023), https://www.insidevoa.com/p/5829.html
61. See Note 55.
62. B.R. Price, 'Colonel John Boyd's Thoughts on Disruption', Marine Corps University Press. *MCU Journal, JAMS* 14 no. 1 (2023). https://www.usmcu.edu/Outreach/Marine-Corps-University-Press/MCU-Journal/JAMS-vol-14-no-1/Colonel-John-Boyds-Thoughts-on-Disruption/
63. J. Whitaker, 'Psychological Warfare in Vietnam', *Political Psychology* 18, no. 1 (1997): 165–79; Whyte Jeffrey, 2018. 'Psychological War in Vietnam: Governmentality at The United States Information Agency', *Geopolitics* 23, no. 3 (2018): 661–89.
64. R. Bouchard, 'Information Operations in Iraq', *Defense Technical Information Center. Access Number*: ADA363160 (1999).
65. See Note 30; Tiwari, 'The Reality of Cyber Operations In The Grey Zone – The Emerging Geopolitics'.
66. H. Lin, 'Russian Cyber Operations in the Invasion of Ukraine'.
67. M. D. Phillips and T. A. Drohan, 'Informatizing Operations: The Other Half of All-Domain Warfare', *Small Wars Journal* March 03 (2020). https://smallwarsjournal.com/index.php/jrnl/art/informatizing-operations-other-half-all-domain-warfare
68. R. Stenson, US Army Training and Doctrine Command updates Army capstone doctrine, codifying shift to multidomain operations. U.S. Army (2022). https://www.army.mil/article/260943/us_army_training_and_doctrine_command_updates_army_capstone_doctrine_codifying_shift_to_multidomain_operations
69. Department of the Army, *FM 3-13: Information Operations*, 1–2, 2016.
70. USMC. *Information Doctrine*, 4–2, 2022.
71. US Army, U.S. Army Techniques Publication No. 3-13.1: The Conduct of Information Operations (Headquarters, Department of the Army, 2018).
72. See Note 30; Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'.

73. A. Wanless and J. Pamment, 'How Do You Define a Problem Like Influence?', *Journal of Information Warfare* 18, no. 3 (2019): 7.
74. R. J. Ross, 'Information Advantage Activities: A Concept for the Application of Capabilities and Operational Art during Multi-Domain Operations', *The Cyber Defense Review* 6, no. 4 (2021): 63–74.
75. Ibid.
76. T. Blagovest, M. Purcell, and B. McLaughlin, 'Russia's Information Warfare: Exploring the Cognitive Dimension', *MCU Journal* 10, no. 2 (2019): 133.
77. See Note 30; S. Tiwari, 'The Reality of Cyber Operations in The Grey Zone – The Emerging Geopolitics'; T. Blagovest, M. Purcell, and B. McLaughlin, 'Russia's Information Warfare: Exploring the Cognitive Dimension'.
78. J. R. McGrath, 'Twenty-First Century IWIO and the Third Offset Strategy', *Joint Force Quarterly* 82, no. 3 (2016): 16–23.
79. D. Cheng, *Cyber Dragon: Inside China's IWIO and Cyber Operations: Inside China's IWIO and Cyber Operations* (ABC-CLIO, 2016).
80. McGrath, 'Twenty-First Century IWIO and the Third Offset Strategy'; M. Pomerleau, 'Pentagon AI Team Sets Sights on IWIO', *CYISRNET* (2020).
81. Ibid.
82. McGrath, 'Twenty-First Century'.
83. R. Work, 'The Third U.S. Offset Strategy and its Implications for Partners and Allies', As Delivered by Deputy Secretary of Defense Bob Work, January 28, 2015. https://dod.defense.gov/News/Speeches/Speech-View/Article/606641/the-third-us-offset-strategy-and-its-implications-for-partners-and-allies/
84. E. P. Hilner 'The Third Offset Strategy and the Army Modernization Priorities. Director's Action Group', *Center for Army Lessons Learned* (2019).
85. A. Obis and K. Macri, 'The 2023 NDAA Emphasizes AI Investment for Cybersecurity, JADC2', *Government CIO Media*. December 2022. https://governmentciomedia.com/2023-ndaa-emphasizes-ai-investment-cybersecurity-jadc2
86. Ibid.
87. Ibid.
88. US Department of State, 'Data Informed Diplomacy' (2022), https://www.state.gov/data-strategy/
89. US Department of State, 'AI Inventory' (2022), https://www.state.gov/data-strategy/ai_inventory/
90. Ibid.
91. Ibid.
92. CDAO, 'About the JAIC: The JAIC Story', (2022) https://www.ai.mil/about.html
93. Ibid.
94. C4ISRNET, 'Pentagon AI team sets sights on IWIO', July 22, 2020 https://www.c4isrnet.com/smr/information-warfare/2020/07/22/pentagon-ai-team-sets-sights-on-information-warfare/
95. Ibid.
96. CDAO, 'Chief Digital and Artificial Intelligence Office', (2023) https://www.ai.mil/
97. US Department of Defense, 'DoD Chief Digital and Artificial Intelligence Office Hosts Global Information Dominance Experiments' (2023). https://www.defense.gov/News/Releases/Release/Article/3282376/dod-chief-digital-and-artificial-intelligence-office-hosts-global-information-d/
98. Ibid.
99. DARPA, 'About DARPA', (2022) https://www.darpa.mil/about-us/about-darpa
100. Ibid.
101. S. Tompkins, Accelerating Innovation for the Warfighter: Statement Submitted to the U.S. Senate Armed Services Committee – Subcommittee on Emerging Threats and Capabilities, 117th Cong. (2022) (statement by Dr. Stefanie Tompkins, Director, Defense Advanced Research Projects Agency (DARPA)). https://www.armed-services.senate.gov/imo/media/

doc/PASSBACK%20DARPA_%20Tompkins%20SASC-ETC%20testimony%206%20Apr%202022_DARPA_FIINAL%200031.pdf

102. S. Sybert, 'DARPA Launches New Programs to Detect Falsified Media', *GovCIO Media & Research* September 16, 2021. https://governmentciomedia.com/darpa-launches-new-programs-detect-falsified-media
103. Ibid.
104. Ibid.
105. Ibid.
106. Ibid.
107. J. Yin and P.M. Taylor, 'Information Operations from an Asian Perspective: A Comparative Analysis', *Journal of Information Warfare* 7, no. 1 (2008): 3.
108. L. Wortzel, 'The Chinese People's Liberation Army and Information Warfare', *US Army War College, Monographs, Books, and Publications* 30 (2014): https://press.armywarcollege.edu/monographs/506
109. See Note 30.
110. Yin and Taylor, 'Information Operations from an Asian Perspective: A Comparative Analysis', 1–23.
111. T. Thomas, 'The Chinese Way of War: How Has it Changed?' *The MITRE Corporation*. US Army Futures and Concepts Center 47 (2020).
112. Kania and Costello, 'The Strategic Support Force and the Future of Chinese Information Operations'.
113. L. Turner and N. Hinkis, 'Chinese State Media's Global Influencer Operation', Miburo (2022) (currently part of the Digital Threat Analysis Center), https://miburo.substack.com/p/csm-influencer-ops-1
114. (Harold, Beauchamp-Mustafaga, and Hornung 2021) Harold, Scott W., Nathan Beauchamp-Mustafaga, and Jeffrey W. Hornung. *Chinese Disinformation Efforts* (RAND, 2021).
115. Diresta et al., *Telling 'China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives'*.
116. See Note 30.
117. L. Saalman, 'China and Its Hybrid Warfare Spectrum', in *Hybrid Warfare: Security and Asymmetric Conflict in International Relations*, eds. M. Weissmann, N. Nilsson, B. Palmertz, and P. Thunholm (London: I.B. Tauris, 2021), 95–112. Bloomsbury Collections. p. 95.
118. E. B. Kania and J. K. Costello. 'The Strategic Support Force and the Future of Chinese Information Operations', *The Cyber Defense Review* 3, no. 1 (2018): 105–21.
119. Ibid, p. 105.
120. D. Buck, 'China in the Asia-Pacific Cyber Domain', *Marine Corps Gazette* 105, no. 4 (2021): 1–5, p. 1.
121. M. Daniels and B. Chang, 'National Power After AI', *Center for Security and Emerging Technology* (2021): 1–30, p. 12
122. Ibid, p. 13.
123. Ibid, p. 14.
124. D. Buck, 'China in the Asia-Pacific Cyber Domain', *Marine Corps Gazette* 105, no. 4 (2021), p. 4.
125. J. Huang, 'China Using 'Cognitive Warfare' Against Taiwan, Observers Say', Voanews.com (2021), para. 1.
126. Ibid, para. 3.
127. Ibid.
128. C. A. Theohary, 'IWIO: Issues for Congress', *Congressional Research Service* (2018): 7–5700.
129. J. Dettmer, 'China Adopts Kremlin's 'IWIO' Tactics', Voanews.com (April 5, 2021), https://www.voanews.com/a/east-asia-pacific_china-adopts-kremlins-information-war-tactics/6204171.html, para. 1.
130. Ibid, para. 2.
131. Ibid, para. 11.

132. E.B. Kania, ‘Chinese Military Innovation in AI’, *Testimony before the U.S-China Economic and Security Review Commission Hearing on Trade, Technology, and Military-Civil Fusion* (2019), 3.
133. K. Takagi, ‘The Future of China’s Cognitive Warfare: Lessons from the War in Ukraine’, *Hudson Institute* (July 2022), https://www.hudson.org/research/17991-the-future-of-china-s-cognitive-warfare-lessons-from-the-war-in-ukraine
134. Ibid.
135. K. Pollpeter and A. Kerrigan, ‘The China AI and Autonomy Report’, *CNA* 22 (2022): 1–9, p. 3.
136. Ibid., 5.
137. Ibid., 5.
138. H. Towey, ‘Researchers in China Claim they have Developed ‘Mind-Reading’ AI that can Measure Loyalty to the Chinese Communist Party, Reports Say’, *Business Insider*, (July 10, 2022): https://www.businessinsider.com/china-says-mind-reading-ai-can-gauge-political-loyalty-reports-2022-7
139. Ibid.
140. Bureau of Industry and Security, Commerce, ‘Addition of Certain Entities to the Entity List and Revision of an Entry on the Entity List’, *Federal Register* (December 07, 2021), https://www.federalregister.gov/documents/2021/12/17/2021-27406/addition-of-certain-entities-to-the-entity-list-and-revision-of-an-entry-on-the-entity-list
141. Bureau of Industry and Security, ‘Entity List’ (2022), https://www.bis.doc.gov/index.php/policyguidance/lists-of-parties-of-concern/entity-list
142. BBC, ‘Who are the Uyghurs and why is China being accused of genocide?’ *BBC News* (May 24, 2022), https://www.bbc.com/news/world-asia-china-22278037
143. US Department of State, ‘PRC Efforts to Manipulate Global Public Opinion on Xinjiang’, August 24 (2022) https://www.state.gov/prc-efforts-to-manipulate-global-public-opinion-on-xinjiang/
144. Ibid.
145. J. Brandt and V. Wirtschafter, ‘How China Uses Search Engines to Spread Propaganda’, *Brookings* (July 6 2022): https://www.brookings.edu/techstream/how-china-uses-search-engines-to-spread-propaganda/
146. Ibid.
147. Ibid.
148. Ibid.
149. M. Kaput, ‘AI in Search Engines: Everything You Need to Know’, *Marketing AI Institute* (March 7, 2022): https://www.marketingaiinstitute.com/blog/how-search-engines-use-artificial-intelligence
150. US Department of State, ‘PRC Efforts to Manipulate Global Public Opinion on Xinjiang’.
151. J. Kao and others, ‘How China is Using Social Media Propaganda to Whitewash the Repression of the Uyghurs’, *Scroll.in* (June 25, 2021) https://scroll.in/article/998385/how-china-is-using-social-media-propaganda-to-whitewash-the-repression-of-the-uyghurs
152. Ibid.
153. Ibid.
154. Ibid.
155. Ibid.
156. F. Ryan, D. Impiombato, and H. Pai, ‘China is Using Ethnic-Minority Influencers to Spread Its Xinjiang Narrative on Social Media’, *Australian Strategic Policy* (2022).
157. Ibid.
158. Ibid.; P. Mozur and others, ‘How Beijing Influences the Influencers’, *The New York Times* (2021) https://www.nytimes.com/interactive/2021/12/13/technology/china-propaganda-youtube-influencers.html; Ryan et al., ‘China is Using Ethnic-Minority Influencers to Spread Its Xinjiang Narrative on Social Media’.
159. Mozur et al., ‘How Beijing Influences the Influencers’; Ryan et al., ‘China is Using Ethnic-Minority Influencers to Spread Its Xinjiang Narrative on Social Media’.

160. Diresta et al., *Telling 'China's Story: The Chinese Communist Party's Campaign to Shape Global Narratives'*.
161. Ibid., 20.
162. Ajir, Media, and Bethany Vailliant, 'Russian Information Warfare: Implications for Deterrence Theory', *Strategic Studies Quarterly* 12, no. 3 (2018): 70–89.
163. A. Manoilo, 'Modern-Day IWIO and Hybrid War Operations', *Analytical Dossier* 1 (2021): 1–69, p. 3.
164. G. Wilde and J. Sherman, 'No Water's Edge: Russia's Information War and Regime Security', Carnegie Endowment for International Peace (2023), https://carnegieendowment.org/2023/01/04/no-water-s-edge-russia-s-information-war-and-regime-security-pub-88644
165. C. Cunningham, 'A Russian Federation IWIO Primer', The Henry M. Jackson School of International Studies, University of Washington (2020), https://jsis.washington.edu/news/a-russian-federation-information-warfare-primer/; B. Lilly, 'Russian Information Warfare: Assault on Democracies in the Cyber Wild West', *Naval Institute Press* (2022).
166. Topor, Lev, and Alexander Tabachnik, 'Russian Cyber IWIO: International Distribution and Domestic Control', *Journal of Advanced Military Studies* 12, no. 1 (2021): 112–27, p. 115.
167. D. Bolton, 'Targeting Ontological Security: IWIO in the Modern Age', *Political Psychology* 42, no. 1 (2021): 127–42, p. 130.
168. See note 158, 72.
169. Z. Rogers, 'The Geopolitics of Surveillance Capitalism', Chesterfield Strategy (2019), https://chesterfieldstrategy.com/2019/09/16/the-geopolitics-of-surveillance-capitalism/
170. Norris, Pippa, and Ronald Inglehart, 'Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash', *Harvard JFK School of Government Faculty Working Papers Series* (2016): 1–52.
171. Ibid.
172. A. Agrawal, J. Gans, and A. Goldfarb, 'Economic Policy for Artificial Intelligence', *Innovation Policy and the Economy* 19, no. 1 (2016): 139–159.
173. See Note 7.
174. See Note 7; S. Petrella, C. Miller, and B. Cooper, *Russia's AI Strategy: The Role of State-Owned Firms* (Foreign Policy Research Institute, 2020).
175. R. J. Marks and S. Bendett, 'Russia is Systematically Copying U.S. Military AI Robotics', *Mind Matters* (2020), https://mindmatters.ai/2020/10/russia-is-systematically-copying-u-s-military-ai-robotics/.
176. F. E. Morgan and others, *Military Applications of AI: Ethical Concerns in An Uncertain world* (Santa Monica, CA: Rand Project Air Force, 2020), 89.
177. D. Bolton, 'Targeting Ontological Security', 130.
178. Ibid., 136.
179. Ibid., 136.
180. O. S. Sheremet, 'Political and Legal Aspects of the IWIO', *Revista Amazonia Investiga* 10, no. 45 (2021): 31–41, p. 34.
181. N. O'Donnell, 'Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos', *University of Illinois Law Review* 3 (2021): 701–40, p. 710.
182. Bolton, 'Targeting Ontological Security', 137.
183. O'Donnell, 'Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos', 710.
184. See note 158, 74.
185. Ibid., 75.
186. Ibid., 76.
187. Bolton, 'Targeting Ontological Security', 135.
188. Morgan et al., *Military Applications of AI: Ethical Concerns in an Uncertain World*, 83.
189. Ibid., 83.
190. Ibid., 88.

191. A. Polyakova, ‘Weapons of the Weak: Russia and AI-Driven Asymmetric Warfare’, *Brookings Institute*, (November 15, 2018): https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/. https://www.rand.org/topics/asymmetric-warfare.html
192. The RAND Corporation, ‘Asymmetric Warfare’.
193. Polyakova, ‘Weapons of the Weak: Russia and AI-Driven Asymmetric Warfare’.
194. Ibid.
195. D. Salaru, ‘Russia: Facial Recognition Software used to Target Journalists’, *International Press Institute* (June 23, 2022): https://ipi.media/russia-facial-recognition-software-used-to-target-journalists/
196. Ibid.
197. Joint Staff, *Joint Publication 3-13: Information Operations,* (2012), II-1
198. James Andrew Lewis, ‘Cyber War and Ukraine’, CSIS (2022); NOTE: The 2/15-2/16 DDOS attacks were attributed to Russia by Australian, US, and UK. The 2/15 disinformation attack remains unattributed.
199. H. Lin, ‘Russian Cyber Operations in the Invasion of Ukraine’, *The Cyber Defense Review* 7, no. 4 (2022): 31-46.
200. S. Bond, ‘Facebook takes down Russian network impersonating European news outlets’, *NPR* (September 27, 2022): https://www.npr.org/2022/09/27/1125217316/facebook-takes-down-russian-network-impersonating-european-news-outlets
201. Ibid.
202. N. Ibrahim, ‘"We Are Not Prepared": Russia Uses AI, Deep Fakes in Propaganda Warfare’, *Global News* (March 30, 2022), https://globalnews.ca/news/8716443/russia-artificial-intelligence-deep-fakes-propaganda-war/
203. R. Booth, ‘Russia’s Trolling on Ukraine Gets ‘Incredible Traction’ on TikTok’, *The Guardian* (May 01, 2022), https://www.theguardian.com/world/2022/may/01/russia-trolling-ukraine-traction-tiktok
204. Ibid.
205. Ibid.
206. Ibid.
207. Lin, ‘Russian Cyber Operations in the Invasion of Ukraine’.
208. V. Akimenko and K. Giles, ‘Russia’s Cyber and Information Warfare’, *Asia Policy* 15, no. 2 (2020): 67–75
209. S. Bendett, ‘The Development of Artificial Intelligence in Russia’, *Air University Press* 1, (2019): 168–77, 168.
210. Ibid.
211. Ross, Robert J., ‘Information Advantage Activities: A Concept for the Application of Capabilities and Operational Art during Multi-Domain Operations’, *The Cyber Defense Review* 6, no. 4 (2021): 63–74.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Lance Y. Hunter, PhD, is a Professor of International Relations in the Department of Social Sciences and Master of Arts in Intelligence and Security Studies programs at Augusta University located in Augusta, GA, USA. His expertise is in security studies and democratization. His research focuses on the causes and effects of terrorism and conflict, democratization, and the relationship between evolving technology and conflict.

Dr. Craig D. Albert, PhD, is Professor and Director of the Master of Arts in Intelligence and Security Studies at Augusta University. He received his PhD from the University of Connecticut in 2009. His areas of concentration include international security studies, ethnic conflict, cyberterrorism, and cyberwar.

Josh Rutland is a graduate of Augusta University's Master of Arts in Intelligence and Security Studies program. He is an Information Technology Specialist at the U.S. Army Cyber Command. His research focuses on information warfare, cybersecurity, terrorism, and biosecurity.

Kristen Topping is a recent graduate of Augusta University's Master of Arts in Intelligence and Security Studies program. Her research focuses on information warfare, social influence, and how in-depth cross-cultural understanding and language acquisition affect intelligence analysis.

Christopher Hennigan is an AI and Data Operations consultant in Deloitte's Government & Public Services group. He is a graduate of Augusta University's Master in Intelligence and Security Studies program. His expertise is in Six Sigma, business risks and mitigation, data and process analysis, and ML process automation. His research focuses on machine learning, cybersecurity, terrorism, and AI game modeling theory.

Routledge
Taylor & Francis Group

Defense & Security Analysis

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cdan20

The military application of artificial intelligence technology in the United States, China, and Russia and the implications for global security

Lance Y. Hunter, Craig D. Albert, Christopher Henningan & Josh Rutland

To cite this article: Lance Y. Hunter, Craig D. Albert, Christopher Henningan & Josh Rutland (2023): The military application of artificial intelligence technology in the United States, China, and Russia and the implications for global security, Defense & Security Analysis, DOI: 10.1080/14751798.2023.2210367

To link to this article: https://doi.org/10.1080/14751798.2023.2210367

Published online: 11 May 2023.

DEFENSE & SECURITY ANALYSIS
https://doi.org/10.1080/14751798.2023.2210367

Routledge
Taylor & Francis Group

The military application of artificial intelligence technology in the United States, China, and Russia and the implications for global security

Lance Y. Hunter[a], Craig D. Albert [a], Christopher Henningan[b] and Josh Rutland[a]

[a]Department of Social Sciences, Master of Arts in Intelligence and Security Studies (MAISS) Program, Augusta University, Augusta, GA, USA; [b]Department of Government and Policy, Deloitte, Washington, DC, USA

ABSTRACT

A number of studies have considered the theoretical role that Artificial Intelligence (AI) may play in shaping the global balance of power in the future. While these studies are informative, we currently lack an understanding regarding the precise manner AI technologies are being applied and incorporated in militaries in major power states. Thus, in this study, we examine how AI technology is being applied in the militaries in the US, China, and Russia and analyse the implications for the future of AI, global military competition, and international security. We examine current research on the military application of AI technology in the US, China, and Russia and conduct expert interviews with leading AI experts in academia, think tanks, multinational technology companies, and the military to better understand how AI technology is being applied in the three major powers states and the implications for global security.

KEYWORDS

Artificial intelligence; military applications; US; China; Russia; global security

Introduction

In a now infamous quote, in speaking to the importance of Artificial Intelligence (AI) and how it will affect the international balance of power over time, alleged war criminal and Russia Premier Vladimir Putin stated in 2017 that “the one who becomes the leader in this sphere will be the ruler of the world.”[1] While this quote is often referenced to signify the importance of AI development regarding the international balance of power and global security, the gravity of the quote must be taken seriously. Many additional government leaders and researchers also acknowledge the important role AI technology will play in shaping global security in the coming years. Jayshree Pandya, founder and CEO of the security and technology company Risk Group states:

> Technological development has become a rat race. In the competition to lead the emerging technology race and the futuristic warfare battleground, artificial intelligence (AI) is rapidly becoming the center of global power play. As seen across many nations, the development in autonomous weapons systems (AWS) is progressing rapidly, and this increase in the weaponisation of artificial intelligence seems to have become a highly destabilising development.

CONTACT Lance Y. Hunter ✉ lahunter@augusta.edu

> It brings complex security challenges for not only each nation's decision-makers, but also for the future of the humanity.[2]

In a similar sentiment, Physicist and MIT Professor Max Eric Tegmark states:

> The more automated society gets and the more powerful the attacking AI becomes, the more devastating cyber-warfare can be. If you can hack and crash your enemy's self-driving cars, auto-piloted planes, nuclear reactors, industrial robots, communication systems, financial systems, and power grids, then you can effectively crash his economy and cripple his defenses.[3]

These statements highlight the pivotal role AI may have in shaping the international balance of power and security in the coming years.

Many previous studies have examined the theoretical role AI will have in shaping global military competition. While these studies are important and informative, less research has considered the precise manner in which the three major power states are developing and applying AI technology in their militaries. Thus, we seek to fill a gap in this research by examining how each major power state is applying AI in their militaries and consider the implications for the international balance of power and global security. To accomplish this task, we examine how the US, China, and Russia are applying AI within their militaries, the programmes currently under development regarding the military application of AI, and future programmes that are being considered pertaining to the military application of AI within each state. To conduct this analysis, we interviewed 10 international AI experts from academia, multinational technology companies, think tanks, and the military to assess how AI is being applied in the three major power states and the implications for international security.

We examine the three major power states (the US, China, and Russia) due to the role each state plays in affecting global security and the great power competition that currently exists amongst the three states.[4] As the Congressional Research Service (CRS) states: "The emergence of great power competition with China and Russia has profoundly changed the conversation about U.S. defense issues from what it was during the post-Cold War era."[5] While non-major power developed states and developing states are also important to consider as their levels of AI development can significantly impact international competition and security as well,[6] many researchers contend that the US, China, and Russia are the current major powers and have significant influence over international relations,[7] and analysing the precise manner they are applying AI technology within their militaries can increase our understanding regarding how global AI competition can potentially affect international security. However, we do not argue that AI development in non-major power states and developing states is unimportant for global security. Examining the military application of AI technology in non-major power developed and developing states is important to international security, but is outside the scope of the current study due to the time and space needed to examine these additional states in an accurate and comprehensive manner.

The layout of the paper is as follows. First, we explain our process for selecting and interviewing the 10 AI experts included in the study. Second, we discuss the role AI technology plays in affecting global security by examining current research at the intersection of AI and defence. Third, we analyse the development and application of AI technology within militaries in the US, China, and Russia and discuss the results from our AI expert interviews regarding the application of AI technology in the three major power states to assess the effects for the international balance of power and international security. Lastly, we consider the implications of our research findings and discuss how the development and application of AI technology in the major power states affects global security.

AI expert interviews

To assess the importance of AI technology to the balance of power and develop a deeper understanding of the military application of AI technology in the US, China, and Russia we conducted expert interviews with 10 AI experts in multiple countries. We selected international experts from academia, think tanks, multinational technology corporations, and the military to obtain a broad range of expertise and perspectives regarding the application of AI technology in major power states and the implications for global security. The breakdown of AI experts is as follows: three experts were from academic institutions, three experts were from think tanks, two experts were from the military, and two experts were from leading multinational technology companies. The identities of the experts are kept anonymous in this article to protect the individuals participating in the study and to ensure they could provide answers to our questions free from concerns that their personal or professional reputations would be compromised. We refer to the experts as Expert A, B, C, etc. for purposes of anonymity. We asked our AI experts a wide range of questions regarding the potential effects of AI regarding the international balance of power, global security, and great power competition regarding the US, China, and Russia. However, due to space limitations and the focus of this study, we discuss the questions that are pertinent to the study and not every question that was included in our interviews. Our research project was approved by the Institutional Review Board (IRB), and we obtained written and verbal consent from each expert prior to each interview.

AI, global security, and the international balance of power

The International Balance of Power (BOP) refers to how the distribution of economic and military power between states affects international relations.[8] The study of the BOP considers how changes in relative power between states affects state behaviour. Many researchers contend that as states increase the amount of power they have relative to other states, they obtain greater influence (political, economic, military) in the international system, as well as in respect to specific regional spheres of influence (e.g. South China Sea, Eastern Europe, Latin America).[9] Thus, variations in levels of relative power amongst the US, China, and Russia affect how much influence each state has internationally and in respect to their regional spheres of influence. The development and military application of AI is important to BOP discussions because many researchers contend that AI is an important factor that affects military capabilities and the ability of states to project power regionally and globally.[10] Therefore, AI is now arguably central to the analysis of relative power and the BOP.

Scholars contend that AI will drastically affect military strategy, as well as impact international power balances[11] and could intensify strategic rivalries amongst great power states and possibly increase instability.[12] The rational for these positions is based on the effects AI technology has on military capabilities and the distribution of military power between states.[13] As Ayoub and Payne remark: "Our argument is not that AI will prove strategically infallible, or that it will somehow replace humans, but rather that it will address such unbounded and complex problems in ways that confer distinct advantages to those who can employ them."[14]

More specifically, militaries that can develop and apply superior AI enabled programmes with greater optimisation on the battlefield will have significant advantages over their opponents.[15] In addressing the importance of AI regarding military capabilities, Davis remarks:

> the nation which can more quickly and effectively harness AI across the spectrum of war will have an advantage by being able to marshal more resources, more efficiently posture forces, and implement their actions in ways which are subtly more effective.[16]

Additional researchers from government institutions and think tanks echo the idea that AI technology will play a vital role in affecting state security, international economies, and military conflict:

> Artificial intelligence (AI) likely is one of the most dramatic technological game changers of our time with the potential to transform human life from daily social interactions to how we conduct warfare. Specifically, AI will play a critical role in driving change in military, information, economic superiority, and the nature of security risks … "[17]

Scholars also contend that the proliferation of AI technology will have profound consequences for existing defence strategies including nuclear deterrence. Based on this scholarship, a larger AI presence in decision-making functions could complicate traditional deterrence strategies and increase the likelihood of uncertainty in the area of nuclear deterrence due to the introduction of a greater number of unknown factors regarding AI decision-making processes.[18] In more specific terms, AI technology can potentially affect the surveillance of nuclear threats and the decision-making processes that control the use of offensive nuclear weaponry. Thus, scholars contend that AI will significantly affect the nature of warfare,[19] how decision-makers approach potential international crises, nuclear deterrence, and strategic stability.[20]

Every expert we interviewed considered AI to be important to the distribution of relative power amongst states. When asked the question: *Regarding the international balance of power, how important is the development of AI technology on a 1–10 scale (1 = Not Important at all; 10 = Extremely Important)*, the average response was 9.40. Six experts answered 10, two experts answered 9, and two experts answered 8. Thus, every expert answered 8 or higher. The experts stated that AI has the ability to affect production and efficiency within states in significant ways in numerous areas including general economic purposes, education, medicine, and defence. The experts noted the ability of AI to be used for multiple purposes including, but not limited to, administering information flows, decision-making, surveillance, intelligence gathering, communications, data management and analysis, and decreasing production costs, and increasing efficiency across countless sectors.

Given the importance many scholars, researchers, and government leaders place on the ability of AI to impact the distribution of relative military power amongst states and international security from a theoretical perspective, we contend that it is important to analyse the precise manner the major power states are developing and applying AI technology within their militaries. Doing so allows us to move beyond a purely theoretical discussion of the implications of the AI arms race and identify the specific manner the major power states are developing and applying AI.

Table 1 displays information regarding AI spending levels for the US and China. Data are unavailable for Russia.[21] Table 2 displays information regarding the military application of AI within each major power state and the strengths and weaknesses of each state's AI strategies. We now turn to examining the military application of AI within each major power state by examining how the US, China, and Russia are applying AI within their respective militaries and discuss how each state's application of AI technology compares to one another and the implications for international security.

The US and AI military application: land, air, and sea forces

Land forces

The US's 2018 *National Defense Strategy* [22]places significant emphasis on AI development, and the 2020 National Defense Authorization Act (NDAA) references AI eleven times.[23] In 2015 Robert Work, the then-US deputy Secretary of Defense, emphasised the importance of "human-machine collaboration combat teaming" and argued that its "early adoption will be a key competitive advantage" for those that seize it and a major blow in competitive viability to those states who let the opportunity pass them by. This set the stage for the United States to increase its focus on the development of AI technology as reported by the Congressional Research Service, which detailed that "all U.S. military services are working to incorporate AI into semiautonomous and autonomous vehicles, including fighter aircraft, drones, ground vehicles, and naval vessels."[24] The US's establishment of the Joint Artificial Intelligence Center (JAIC) in 2018 to "accelerate the delivery of AI-enabled capabilities, scale the Department-wide impact of AI, and Synchronize Department of Defense (DoD) AI activities to expand Joint Force advantages" further represents Washington's commitment to advancing the AI-powered tools at it's military's disposal.[25] The JAIC has thus far taken the lead in developing and implementing AI into all branches of the US military. Heller details that according to the DoD budget report for 2019, the Army reserved $6.5 million for training purposes relative to AI, to include simulations and virtual reality and the Navy set aside $6.5 million for similar training purposes to that of the army involving

Table 1. AI spending across sectors for the US and China[a] (2015–2021).

Country	US	US	US	China	China	China
Year	Civilian	Military	Total	Civilian	Military	Total
2015	$3,294,000,000	N/A	$3,294,000,000	$1,600,000,000	N/A	$1,600,000,000
2016	4,093,000,000	N/A	4,093,000,000	2,100,000,000	N/A	2,100,000,000
2017	5,425,000,000	N/A	5,425,000,000	3,400,000,000	N/A	3,400,000,000
2018	9,334,000,000	N/A	9,334,000,000	6,200,000,000	N/A	6,200,000,000
2019	16,500,000,000	973,000,000	17,473,000,000	10,200,000,000	N/A	10,200,000,000
2020	N/A	1,300,000,000	1,300,000,000	14,300,000,000	N/A	14,300,000,000
2021	N/A	6,000,000,000	6,000,000,000	N/A	N/A	N/A

Sources: Doubleday (2020); Harper (2022); Statista (2022).

[a]Data not available for Russia, US Dollars (USD).

Table 2. Artificial intelligence (AI) military developments and applications in major power states.

State	Land	Air	Sea	Command and Control	Information Warfare	Strengths	Weaknesses
US	High-speed Anti Radiation Missile[a]	X47-B autonomous flying systems – canceled due to funding[b]	Tomahawk Anti-Ship Missile (TASM)[c]	Information management	Use of AI and machine learning to wargame information warfare campaigns and achieve Third Offset Strategy (TOS)	Large economy and defense budget	Cultural differences between Silicon Valley leaders and the military regarding the military use of AI
	Robotics development aimed at sensing, navigation, decision-making	AI piloting programmes capable of assuming operational command and autonomously operating planes with little to no human supervision[d]	LRASM[e]	Intelligence processing	Information Operation (IO) training and operational concept improvement using machine learning Pattern detection	Joint AI development centre is a focal point of DoD AI strategy and is targeting development around “key missions” to maximise usefulness	So far, funding for AI research has been lower than expected
	“Self-healing” systems that can repair themselves autonomously when damaged/attacked[f]		Data analysis, strategic insight and pattern identification to overcome the fog of war	Financial planning, departmental coordination, orders development conducted by AI		Access to Silicon Valley Developments and DARPA Assistance	AI projects may lack Innovation sufficient to ensure long term advantage over adversaries
			Asset management and monitoring, prescriptive maintenance to allocate funding and resources responsibly	Transportation logistics and supply route planning			
			CLAWS[g]	VR Training Simulations			
				Disaster response (1CONCERN)[h]			

China	Sharp Claw I/II: Small unmanned tank-like drone. Receives input from a human operator when necessary	UAVs	Jinghai USV: Unmanned patrol boat for harbour and fleet defense	Use AI to help commanders make quicker decision	Automated monitoring systems designed to observe civilians/persons of interest (could be used to control information distribution)	China is considered to be close or equal to the US regarding AI development	Global strategy somewhat isolates China, decreasing chance of cooperation with other states regarding development
	CH-901: Man-portable, remote controlled, tube launched drone capable of carrying an ISR pod or warhead	Cruise Missile (automated launch and targeting)		Automatic target identification, automatic defense systems, AI driven reconnaissance systems		Doctrine prioritising security and military modernisation has accelerated development	"Domestic AI talent shortage, relaxed data privacy policies, brutal internal market and foreign competition, and irregular distribution of funds across important sectors," which may have stunted China's AI developmental growth in some ways
		DF-ZF craft: world's first hypersonic glide vehicle/missile system (Designed to Break through US defenses) Stealth Attack Drones: ID targets, aim and fire without human intervention Sky Eye				Secretive posture may help protect technological developments from being discovered or imitated by rivals	
Russia	Nerekhta unmanned ground vehicle	Drone swarming technology in development	AI piloted sea-mines with auto-targeting capabilities	NTechLab has developed facial recognition software for surveillance of civilians and military use	Russia's strategic focus is on information warfare	Focus is more on information warfare rather than diverse AI programmes. This doctrinal decision could place Russia on uneven footing with other major power states if AI becomes a central factor in global security competition	Russian strategy is to copy US AI developments. This may allow Russia to develop AI capabilities without incurring as many developmental expenses

(*Continued*)

Table 2. Continued.

State	Land	Air	Sea	Command and Control	Information Warfare	Strengths	Weaknesses
	Roboticisation/ automation of weaponry	UAV technology is still a focal point of development	"Galtel" unmanned underwater vehicle (hunt for mines, unexploded ordinance, enemy vessels)	FindFace app by NTechLab is being tested for use in "aggression detection"	Troll farms incorporate Deepfake technology into propaganda campaigns, and AI that can generate propaganda with human direction can significantly boost propaganda output		However, this places Russia at a developmental disadvantage in the long term. Russia's focus on AI as it pertains to information warfare may harm the diversification of AI development across sectors
				Satellite imagery and radar monitoring, processing, analysis			

[a]This missile variety is capable of targeting radiation signals such as those emanated by air defence systems. It is capable of operating autonomously and selecting targets on its own, but is designed so that a human operator must approve any strikes conducted by the missile.

[b]The X47-B autonomous piloting system is capable of conducting autonomous landing procedures and ariel refueling. However, it suffered from a lack of stealth and has since been cancelled due to funding issues (Azeem, 'Autonomous Unmanned Aerial Vehicles').

[c]The TASM is capable of loitering mid-air to search for ships to engage but has been known to cause accidents due to targeting issues (unintended targets were struck as a result).

[d]The U.S. air-force has test-flighted AI copilots capable of piloting U-2 spy planes on their own absent of human supervision. μZero, an AI programme designed to play board games such as chess, was trained to pilot the U-2 spy plane through over a million virtual training runs, assigning it the callsign ARTUμ. The AI was given "operational command" and piloted the craft successfully (Roper, 'Exclusive: AI Just Controlled a Military Plane').

[e]LRASM is a stealth based anti-ship cruise missile. The weapon is in development as a cooperative project between the U.S. Navy and Air-Force. The missile is capable of autonomous targeting absent of any human guidance post-launch via built-in, AI driven targeting systems. Weapon is vulnerable to human errors that can occur when uploading the software. The missile instructions have to be specific (Lockheed Martin, 'Long Range Anti-Ship Missile').

[f]Lee et al., 'Industrial Artificial Intelligence'.

[g]CLAWS is a prototype autonomous submarine in development by the Office of Naval Research. The subs are expected to come armed with 12 torpedo tubes that can be autonomously activated by the CLAWS AI absent of any human input. The sub will have its own authority to initiate or respond to attacks. The described goal is to "clandestinely extend the reach of large UUVs [unmanned underwater vehicles] and increase the mission areas into kinetic effects" (Macaulay, 'The US Navy is Developing AI-Powered Submarines', para. 6).

[h]A mapping tool designed to address civilian crises and natural disasters. Goal is to develop a "common and comprehensive picture during emergence operations" (Heller, 'Near-Term Applications of Artificial Intelligence').

AI, with the addition of experimentation with AI systems for combat purposes.[26] In addition, the DOD 2021 budget continues to focus on Advanced Capability Enables (ACEs) that enable high-end operational functions. The DOD allocated $1.5 billion in funds for Microelectronics/5G technology that provide critical infrastructure for AI pacing technology. The budget also provides $1.7 billion for autonomous system that generate enhanced speed and maneuverability for autonomous and semi-autonomous vehicles and greater collaboration between human and machine programmes. The budget also provides 0.8$ billion for AI pathfinders, JAIC, and Project Maven.[27] According to Heller, the Army also planned to begin field testing new unmanned combat systems referred to as Next Generation Combat Vehicle (NGCV) by late 2019, with new systems expected to be assigned to units with the intent to "replace both the M1 Abrams tank and the M2 Bradley infantry fighting vehicle … eventually."[28] In, 2018, the NGCV was reclassified as the Optionally Manned Fighting Vehicle (OMFV). The OMFV vehicles have not been deployed to date due to issues regarding the AI industry's ability to respond to the Army's timeline for acquisition and implementation.[29]

Air forces

The Air Force reserved $87 million to experiment with AI for wargames and field training in 2019.[30] The US Air Force has begun researching multiple autonomous piloting systems. These programmes are expected to be fully capable of assuming operational command and autonomously operating planes.[31] The system is designed with the AI programme serving as a pilot that functions as one part of a crew that is still primarily composed of humans. In addition, the programme includes settings that would specify risk acceptance levels regarding flight operations that the AI would follow when piloting its missions. One such programme, μZero, was designated with the callsign ARTUμ and successfully piloted a U-2 stealth craft in 2019.[32] Another similar programme, the X47-B autonomous piloting system, is capable of conducting autonomous landing procedures and aerial refueling.[33] The system performed poorly in stealth routines, however and has since been cancelled due to funding issues.[34] It should also be noted that the Air-Force and Navy have engaged in multiple cooperative AI development ventures as well, such as Long-Range Anti-Ship Missile (LRASM).[35]

Sea forces

In 2019 the Navy set aside $6.5 million for similar training purposes to that of the army involving AI (e.g. simulations and virtual reality), with the addition of experimentation with AI systems for combat purposes.[36] Heller also notes that the Navy, through its "rapid prototype development program," has dedicated $49 million to integrate AI with its combat systems "such as new submarine combat assets."[37] The US Navy has conducted extensive research into autonomously guided missile systems, resulting in Tomahawk Anti-Ship Missiles (TASM) (Scharre 2018) and LRASM stealth anti-ship cruise missiles (Lockheed Martin, 'Long Range Anti-Ship Missile'). The Marine Corps has also assigned $7.1 million for an "unmanned warning system to provide commanders with increased situational awareness."[38]

The Cognitive Lethal Autonomous Weapons Systems (CLAWS) includes a prototype autonomous submarine currently under development by the Office of Naval Research and is anticipated to have a significant impact on sea-based warfare.[39] It is expected to come armed with 12 torpedo tubes that can be autonomously activated by it's on-board AI, which will have authority to operate and manage threats autonomously.[40] Applications beyond combat are also being researched by the U.S. Navy, most notably asset management and monitoring. Specifically, AI will be used to monitor the real-time status and repair needs of Navy vessels and equipment, which will allow for more responsible allocation of funding and resources through prescriptive maintenance rather than preventative or reactive maintenance plans.[41] While the United States has accelerated its AI development in recent years, it has been significantly involved in the development and application of AI technology in the military prior to the JAIC's creation. As early as July 2016, the US Marine Corps tested "the modular advanced armed robot system, which uses sensors and cameras to control gun-toting robots based on AI."[42] Robotics development is also progressing towards systems capable of sensing, navigation, and decision-making (Feickert et al. 2018). It is also important to note that in 2015 the then Secretary of the Navy, Ray Mabus, expressed the Air Force's interest in progressing towards AI-assisted aircraft, noting that: "the F-35 should be, and almost certainly will be, the last manned strike fighter aircraft the Department of the Navy will ever buy or fly."[43]

The US and AI military application: decision-making capacity

Decision-making capacity regarding AI refers to how AI can assist, or control processes related to decision-making functions. Decision-making capacity entails the following components: communications, information flows, data analysis, co-ordination, and prioritisation. In basic terms, decision-making capacity refers to how AI affects military decision-making processes. Regarding US command and control developmental strategies, the US has been conducting AI research regarding "self-healing" AI systems that can repair themselves absent of human technicians when they are damaged and attacked in battle.[44] When realised, this development would provide greater safety for human operatives by removing the need for their presence on the battlefield, and it could also protect operational assets by ensuring their continued functionality and effectiveness despite losses.[45] This would allow military decision-makers greater protection against physical, battlefield threats. Alongside these developments, the United States is also implementing AI into military decision-making processes.[46] The development is geared towards information management, intelligence analysis and report generation, financial planning, departmental co-ordination, order generation and dissemination, transportation logistical planning and supply route planning, and VR training simulations that more accurately reflect the constantly evolving combat scenarios soldiers may encounter on the battlefield.[47] According to the Center for Security and Emerging Technology, "advances in big data analytics, speech recognition systems, natural language processing, neural networks, reinforcement learning, and other techniques will help commanders process and assess more options for action in complex situations."[48] Sisson also notes that the US military, and others globally, are likely to use AI to assist with decision-making, and that this may be through "providing information to humans as they make decisions, or even by taking over the entire execution of decision-making processes."[49] AI may also help new commanders and their units as they transition into active engagements previously held by other personnel by offering incoming units "an opportunity to have the system ingest all unstructured text generated by the departing unit during its deployment."[50] Therefore, new personnel could consult an experienced AI-enhanced knowledge base for insights about the current mission which would reduce orientation times. This could shorten the length of time required for the previous unit to remain in place and ensure the incoming unit could send troops to strategic positions faster and more seamlessly. The ability to brief and reassign troops more quickly and efficiently would provide militaries with advantages regarding speed and maneuverability compared with states that do not possess the same AI tools.

The US military application of AI: overview

The current US AI development plan has several strengths. The large economy of the US (20.9 trillion USD in 2020) combined with its extensive defence budget (778.2 billion in 2020)[51] provides opportunities for significant advances in the development and military application of AI technology.[52] However, despite the availability of resources, funding for AI development and application has been less than expected thus far, and the current level of innovation stemming from AI projects may not be sufficient to ensure long-term advantages over US adversaries.[53] In short, an increased prioritisation of AI could benefit US strategy to maximise its technological resources. The developmental guidance and resources stemming from the Defense Advanced Research Projects Agency (DARPA) and access to developments from the private sector in Silicon Valley means that cutting edge technology and intuitive development plans are accessible to the government in some instances. However, there are cultural differences between Silicon Valley leaders and the military regarding the military use of AI, which can limit collaboration amongst Silicon Valley companies and the military.[54] This is further complimented by the DoD's most recently published National Defense Strategy that was presented by then Secretary of State Jim Mattis which prioritises the maximum use and efficiency of AI, as well as joint development across branches of the military.[55] The National Defense Strategy includes a focus on autonomous weapons systems, as stated: "The Department will invest broadly in military application of autonomy, artificial intelligence, and machine learning, including rapid application of commercial breakthroughs, to gain competitive military advantages."[56] If significant military advances in autonomous AI developments are to be achieved, collaboration between Silicon Valley and the military is likely necessary.

Regarding AI development and the military application of AI in the US and our experts' analysis, when asked the question: *What country/state is currently leading in AI development?* – the experts were evenly divided regarding whether the US or China was leading in the development and application of AI. Five experts believed the US was ahead and five believed China was leading. Some of the experts stated that it was too early in the AI arms race to determine the leading state. In addition, whilst the experts varied regarding whether they believed the US or China was ahead in AI development, many of the experts expressed that much remains to be determined regarding the AI arms race and many factors could affect whether the US, China, or Russia, take a significant lead in the development and application of AI. In addition, many of the experts highlighted specific ways that each major power state was pursuing AI development.

Expert A stated that some of the top experts in AI and machine learning are in the US and there is significant funding available from federal agencies to facilitate their research and attract elite talent. Also, according to the expert, China produces top experts in data sciences as well.

The expert stated that many other elite AI researchers are dispersed around the globe, although the US and China have a large amount.

Expert B stated the US is leading in terms of AI innovation and research, but regarding real-world deployments and industrial integration, China is the leading state in the world. The expert stated that large sums of money are available for research and innovation in the US. US universities and large US companies, that are focusing on AI, are benefiting from the pool of money more than researchers in China or Russia, which allows the US to excel in innovation. The expert stated that US AI innovations are frequently open source. Thus, they are often published in venues that are fully accessible to anyone who is interested in using them, including researchers and developers from China and Russia. This means that the innovation costs are often shouldered by the US, but everyone in the world benefits from that investment, including adversaries of the US. In addition, according to the expert, China and Russia have a common advantage over the US, because they do not have to contend with privacy issues or data governance issues as much as the US, which means non-democratic states like China, Russia and others have greater data access and are better configured to deploy innovations that have been made with AI user interface (UI).

Expert C stated that authoritarian states, such as China, have some advantages regarding AI development due to their ability to control their societies. However, democracies, such as the US, also have certain advantages regarding AI development due to their cultures of innovation and competition. A separate expert (expert D) stated that Russia and China are superior to the US in the mathematical aspects of AI development and algorithms, and they have more momentum in those areas.[57] According to this expert, the US's primary strength is it is more capable regarding the application and implementation of AI technology in a more user-friendly manner and China and Russia's strengths are in development.

China and AI military application: land, air, and sea forces

Land forces

Currently, some observers consider China to be the US's closest competitor regarding the development and application of AI technology in the military.[58] China's State Council has stated that it intends to lead the world in AI development by the year 2030 with a significant focus on incorporating AI into its military.[59] China's aim to develop AI superiority by 2030 and competition with the US is a major driving forces behind its overall AI strategy. This is supported by De Spiegleire, Maas, and Sweijs's assertion that "the Chinese military's initial approach to artificial intelligence is still strongly informed by its examination of US developments and initiatives."[60] Thus far, China has invested in and field-tested numerous unmanned vehicles for reconnaissance missions and even "developed AI in missile technology."[61] Alongside the goal of AI superiority by 2030, China also "aims to have an AI industry worth $150 billion USD," with a single technology centre in Beijing's suburbs receiving $2.1 billion of investment as of 2019.[62] The domestic AI market in China increased 24 percent to 160.69 billion yuan ($24.955 billion USD) in 2020 and is projected to increase 26 percent to over 203 billion yuan ($31 billion USD) by 2021.[63] The size of the domestic AI market is relevant for discussions on the military development of AI since AI-related products that are produced in the domestic arena can potentially be used for military purposes.

China is utilising AI across a number of different platforms with a diverse range of goals. These include "aggressively developing autonomous robots, both to increase the effectiveness of current weapons and tactics and to gain entirely new capabilities," but also the eager development of "autonomy at rest" which refers to "advanced data processing and decision support systems."[64] Similar to the US, China's military leaders seem to believe that these capabilities will enable them to "find hidden platforms, turn sensor data into a common operating picture, and speed up decision-making by serving as a digital staff officer for the commanders."[65] Thus far, publicly, China appears to have avoided the direct implementation of weapons systems into its military that can: "identify targets, take aim, and fire without human intervention," though, like the US, several of the weapons' systems it has developed could be easily repurposed or modified to operate autonomously with some relatively simple software adjustments.[66] According to Human Rights Watch,[67] China has officially supported banning the use of LAWS but has explicitly refrained from supporting a ban on their development and manufacturing, likely due to their own profitable development and sale of these technologies according to former US Defense Secretary Mark Esper's 2019 comments.[68] In addition, as previously, stated, China has developed and sold LAWS to other states, and LAWS could be incorporated by the Chinese military without public knowledge. China's apparent unwillingness to adopt LAWS weapons publicly thus far should not be interpreted as a signal that they will be unwilling to adopt these weapons in the future. Concerns over losing competitive advantages to adversaries could prompt states such as China to incorporate LAWS even though ethical issues remain surrounding their usage. Currently, the PLA has focused on swarms and on "utilizing AI to empower the flight guidance and target recognition systems in new generations of its own cruise missiles," which could enhance their operational versatility and allow military commanders to programme missiles to function within specific and rapidly altering battlefield conditions.[69] The Sharp Claw I/II are one of the few AI driven pieces of technology China has developed to bolster its ground forces. These small, tank-like drones receive limited input from a human operator only when necessary and can function in an armored support capacity for ground forces.[70]

Air forces

China has developed air, land, sea, and undersea autonomous vehicles.[71] China has also engaged in extensive AI development to bolster its air forces, including various missile and stealth craft technologies such as UAV's and cruise missiles capable of automated launching and targeting.[72] The PLA's development of stealth attack drones capable of target identification and neutralisation without human intervention could also represent a significant threat on the battlefield.[73] The Chinese air force has also publicly demonstrated its DF-ZF stealth craft, which is capable of hypersonic gliding and comes equipped with automated missile systems.[74] This craft, which is designed specifically to penetrate U.S. defensive systems, is indicative of China's willingness and desire to challenge the US from a military and technological standpoint.[75]

There is some evidence to suggest that China is developing, or has developed LAWS, most notably Defense Secretary Mark Esper's statement that "the Chinese government is already exporting some of its most advanced military aerial drones to the Middle East."[76] The drones Secretary Esper is referencing are those such as the Chinese company Ziyan's "Blowfish A3" which is marketed as capable of "autonomously performing more complex combat missions, including fixed-point timing detection, fixed-range reconnaissance, and targeted precision strikes."[77] China explicitly advertises the drones as capable of "full autonomy, including the ability to conduct lethal targeted strikes."[78] With this in mind, one could reasonably assume that China has "deployed" LAWS given that they have sold these technologies to their strategic allies, though their incorporation into the Chinese military is a point of debate. Additionally, in 2017, a Chinese university with ties to the military demonstrated an "AI-enabled swarm of 1,000 uninhibited aerial vehicles at an air show."[79] Swarming refers to the deployment of many drones simultaneously on a specific target-set; experts contend that swarms can provide militaries with significant advantages due to the volume, speed, coordination, and intelligence they offer.[80] Paul Scharre, a pioneer of the swarming technique and autonomous weapons expert at the Center for New American Security (CNA) states: "collectively, swarms of robotic systems have the potential for even more dramatic, disruptive change to military operations."[81]

The PLA has also devised a small, human carried portable drone that can be deployed through a launching tube and remote controlled once airborne.[82] The drone is capable of carrying and deploying an ISR pod, or small warhead, making it a credible threat to ground troops.[83] This device, known as the CH-901, is aimed at assisting terrestrial forces, but its avian nature is representative of China's developmental focus on "airborne robotic systems [rather than] terrestrial or maritime systems" and its concentration on "using autonomy to improve the effectiveness of existing platforms and tactics."[84]

Sea forces

China's military and navy are alleged to be working on several AI related projects based on public declarations from President Xi Jinping.[85] One such operation is being labelled "the first-ever AI-run colony on Earth" in the South China Sea's Manila Trench.[86] The Colony is alleged to be a research facility capable of ecological disaster detection and monitoring, but could also serve as a launch platform for military operations and a method of tracking foreign vessels.[87] The Chinese Academy of Sciences has also been working to create large, AI-controlled submarine vehicles capable of traversing thousands of nautical miles to engage in missions ranging from whale tracking to anti – carrier operations.[88] Each of these endeavours could be used to bolster China's "underwater great wall" initiative aimed at creating "a network of submarine detectors for national security in the South China Sea."[89] The Sharp Claw I/II drones and the

Jinghai USV are also AI tools developed by China primarily developed for reconnaissance and defence.[90] The Jinghai USV is an unmanned small naval craft that is capable of patrol functions as well as harbour and fleet defence.[91]

China and AI military application: decision-making capacity

China has tailored its AI command and control developmental strategy towards assisting commanders in providing rapid and accurate decision-making, as well as automatic target identification and reconnaissance.[92] China is also developing "more innovative systems in efforts to create novel capabilities, especially the ability to conduct penetrating strikes."[93] The PLA hopes that AI will help its platforms to quickly find and identify hidden targets whilst autonomously fusing multiple intelligence sources "including open-source intelligence and possibly human intelligence, into a single common operating picture."[94] With this combination of automatic systems, the PLA also hopes to use AI to help commanders "make better decisions faster."[95] These goals are similar to the United States' own desire to make the 'relief in place process' is as efficient as possible, with autonomy a key ingredient for this. Like in the US, debate exists as to the extent to which actual command decisions should be delegated to AI, but a significant number of PLA scholars have argued in favour of automating command processes on the grounds that further automation will be necessary if the PLA hopes to infiltrate adversaries' defences.[96] The rationale for this position is that automated decision-making can speed the human decision-making process by allowing leaders to make more objective and precise decisions due to the enhanced ability to analyse large data stores. This can allow for the penetration of adversaries' OODA loops.[97] In addition, automated decision-making programmes generate more sophisticated wargame simulations that depict adversaries' potential strategies (often termed "red team strategies") and produce more difficult and realistic war game scenarios that allow for the testing of AI functions that can be applied to command operations.[98]

China and the military application of AI: overview

China is considered to be equal to the US in terms of AI military development according to some researchers.[99] However, the ability of China to consistently apply AI programmes in effective and advantageous ways in real-world scenarios remains undetermined. Thus, the practical usability of China's AI applications remains a question for the Chinese military, as many researchers and our AI experts explain.

The PLA has significantly fewer military officers with combat experience compared to the United States.[100] Thus, China may be more inclined to rely on strategies calculated by AI programmes that have been trained through wargaming scenarios in potential conflicts. The PLA traditionally approaches innovation within its military with the mindset that technological capability should determine tactics. Meaning, China may be more willing to relinquish human control and rely on an AI systems for guidance than US operatives with combat experience.[101] The PLA's drive to "achieve ever-greater cognitive speed in battlefield decisions" may be a deciding factor in their strategic approach to utilising AI in the coming years.[102] In addition, China may be more willing to green light automatic weapon systems development and adoption in the near future as the global AI arms race intensifies to close the gap in military capabilities with the US.

The PLA's strategy prioritising security and military modernisation has proven favourable, as it has accelerated development beyond that of many of its regional and global rivals.[103] China's secretive posture has assisted the PLA in securing its technological developments from prying eyes and preventing imitation or early discovery by its adversaries. However, China's posture has left it somewhat isolated in recent years, robbing it of some potential benefits stemming from co-operation with other states regarding AI development.[104] It also suffers from "Domestic AI talent shortage, relaxed data privacy policies, brutal internal market and foreign competition, and irregular distribution of funds across important sectors," which may have stunted China's AI developmental growth in some ways.[105] Overall, the Chinese government's doctrine has succeeded in accelerating its technological and military developments in AI and preventing much of the rest of the world from realising China's exact capabilities. However, this posture has been interpreted as revisionist and has prevented China from obtaining many technological benefits the global community has to offer. Thus, China's ability to maintain its current rate of AI developmental is questionable, particularly if domestic issues regarding AI and human rights persist or worsen in the coming years.[106] More specifically, China appears to be developing and applying AI for domestic public surveillance systems capable of identifying and tracking persons of interest.[107] This raises concerns over human rights issues which could potentially affect Chinese AI development in the long-term by affecting China's ability to attract and retain top AI scientists.[108]

As previously stated, when asked the question: *What country/state is currently leading in AI development?* – five experts believed that China was leading and five believed the US was ahead. Expert E noted that China has been able to transition certain AI tools to everyday (domestic) use more so than the US in respect to functions like payment applications or delivery systems, and China is ahead of the US regarding these types of applications. The expert also highlighted the potential strengths and weaknesses of China's economic and political systems as it relates to AI development. Regarding the potential strengths, the expert stated that the Chinese government is better able to draw upon innovations from the private sector, in part because of the nature of its economic model in which the state and private sector are closely interlinked, and the large role state owned enterprises (SOE) play in the economy. Thus, there is a closer relationship between Chinese technology companies, such as Tencent, and the Chinese government than between US companies like Google and Facebook and the American Government. Also, the Chinese government is likely more willing to pressure technology companies in China to pursue the state's goals, which could give China an advantage regarding how it controls the direction of AI research. However, according to the expert, these relationships generally require more resources on behalf of the state, and the state-led semi-planned system is generally less dynamic than a capitalist system where there is more entrepreneurship and innovation.

Expert F noted that China is the most interesting major power state regarding AI development because it can gather and harness data together much easier than the US and many other states given that its citizens do not have the same privacy rights and ownership of data.

Expert G stated that the US and European states are still ahead in terms of the development and implementation of AI because they have environments where people believe they can harvest the benefits of the technology. Thus, commercially, people feel they can profit from AI on the back end, and in China there is less potential profit on the back end due to a weaker presence of private market forces. Thus, AI advancements in China do not appear to be as fast as those in the US and Europe, according to the expert.

Expert G stated that China is spending enormous sums on AI development and is educating AI experts to a greater degree than the US or Russia. However, the expert stated that some of the great AI minds may come from states outside the major power states. Thus, simply increasing spending on AI may not produce extraordinary results unless the major power states attract and retain the top minds working on AI projects, according to the expert.

Contrary to experts A and B discussed earlier in the paper, expert H stated that China has already prevailed regarding international AI competition because of its integration of AI into the global economic supply chain and the size of its data stores. Expert I stated that China is moving at a breakneck pace in AI development and is investing far more heavily in AI than any other state in the world and is significantly investing in the military application of AI. According to the expert, the balance of creativity remains on the side of the United States, but China seems to be substantially ahead of other states regarding the movement of AI technology into real-world applications. The expert also stated that China is very agile in determining its goals and investment priorities and then pursuing that direction because they have a heavily centralised government and can quickly and vigourously pursue their technological aspirations compared with the US that has a decentralised system of governance and 50 states with different priorities that makes pursuing a unified direction more difficult. Also, the expert stated that "blue sky" funding can significantly assist AI research, due to the spillover effects that often emerge from such projects. However, the expert stated that there are fewer such funding opportunities in the US in many instances compared to opportunities in Europe, except for DARPA. The expert stated that the lack of ambitious, long-term AI funding opportunities in the US is problematic because many breakthroughs emerge from "blue sky" type projects.

Russia and AI military application: land, air, and sea forces

Land forces

Whilst many scholars contend that Russia currently trails the US and China in AI development,[109] it is worth emphasising that President Vladimir Putin has repeatedly acknowledged the importance of AI, most notably with his 2018 claim that "whoever becomes the leader in this field will rule the world," as noted in the beginning of the article.[110] "In 2018, Russia released a 10-point AI agenda which outlines its goals for AI development including the establishment of an 'AI and Big Data consortium.'"[111] Former Deputy Defense Minister Yuri Borisov stated that military victories increasingly depend on advanced technologies such as AI.[112] Reflecting Putin's sentiments, Russia has heavily invested in AI development since this public declaration and has openly admitted to using AI military technology regarding its air and missile defense systems.[113] This information was likely made public for purposes of deterrence as Putin stated in respect to the capabilities of the semi-autonomous and autonomous missile systems that "they could reach anywhere in the world [and that Russia was not] bluffing."[114] In addition, many researchers contend that the development and application of AI technology is pivotal to the future viability of Russia's military.[115] One reason for the increased need for the incorporation of AI and automation in the Russia military is due to the declining Russian population and need for additional labour in the military.[116] "The Russian Military Industrial Committee has approved an aggressive plan that would have 30% of Russian combat power consist of entirely remote-controlled and autonomous robotic platforms by 2030."[117] However, despite this ambitious objective, according to a Congressional Research Service report in 2018, Russia lags the US and China regarding the military application of AI.[118]

Regarding specific AI military application, an AI powered Russian ground vehicle, nicknamed Nerekhta, "saw use by forces operating in Syria, amongst others for clearing booby traps and IEDs left behind by ISIS forces in the wake of the March 2016 regime offensive against the city."[119] The Russian armed forces have also begun "using AI for border protection, developing a system which will automatically interact with cameras, sensors, radars and drones to monitor [its] Eastern and Southern borders."[120] These developments are part of a larger trend towards robotisation and automation of weaponry in the Russian armed forces.[121] Further examples of this trend include the Kalashnikov defense manufacturing company's 2018 "prototype of an autonomous armored turret capable of independently acquiring, identifying and engaging targets" and expanded involvement with "the use of remotely controlled land robotic systems in actual combat" such as the Uran-9 vehicles, which are unmanned light tanks equipped with 30-millimetre guns and anti-tank guided missiles.[122] However, it should be noted that these vehicles have displayed some limitations when operating alongside ground troops, including "inadequate situational awareness of the robot's operator, which leads to difficulties when interacting with other units."[123]

Air forces

The Russian air force possesses several AI-powered weapons, many of which have been tested in Syria.[124] The KUB-BLA drone developed by the Kalashnikov Group is designed to destroy ground targets using co-ordinates that can be set manually or using images from the drone's built in guidance system.[125] Another drone model, the KYB – UAV AI enabled drone developed by ZALA Aero, self-destructs upon striking a target.[126] The Russian air force has also been working on AI-guided missiles "that could switch targets mid-flight since at least early 2017" in an effort to emulate U.S. missile technologies and close the technological gap between the two states' militaries.[127]

Sea forces

The Russian navy has prioritised the development of "unmanned underwater vehicle" (UUV) technology with a new generation of UUVs tasked with strategic missions. One example is the Poseidon; an unmanned, nuclear-powered platform with intercontinental range, which has been undergoing sea trials since 2018.[128] While its precise date of deployment is unknown, one of its planned deployments will be as a nuclear strategic delivery system, designed to "enhance Russia's second-strike capability."[129] The "Galtel" UUV, while not combat oriented, will allow Russia to protect better its naval assets by hunting for mines, unexploded ordinance, and enemy vessels within a given zone.[130] In a slight departure from more traditional naval craft, the Russian navy has also engaged in the development of AI-piloted sea-mines capable of automatically targeting and engaging enemy craft.[131] These mines could dramatically alter Russia's capacity to control aquatic territory, but may prove too expensive to manufacture on a large scale given the sacrifice of resources used to build the mines' piloting system.

Russia and AI military application: decision-making capacity

Russia's AI related command-and-control developments have been split between more traditional intelligence collection formats and innovative monitoring and various facial recognition programmes for both civilian and military applications.[132] The Russian tech firm NTechLab has been responsible for much of its more unique software developments. The FindFace APP represents one of the lab's more significant developments, as the AI powered facial recognition software is currently being tested for its ability to detect signals of "aggression" in humans.[133] This technology can be transitioned to the military to use for defence purposes to allow senior officers the ability to identify threats and make decisions regarding possible responses to perceived hostilities. More conventionally, Russia is also experimenting with automated satellite imagery and radar monitoring, as well as automated processing and analysis of intelligence collected from its traditional sensors and monitoring systems. This functions as an extension of Russia's recent experimentation with machine learning technologies that has resulted from global attention on the technology.[134] The end goal for these experiments is likely an automated data collection and processing system capable of threat monitoring, early warning, and more streamlined intelligence collection and processing.

Russia and AI military application: overview

Russia is challenged by its own financial limitations in contrast to its political rivals and industrial competitors who have access to greater resources and AI experts.[135] Russia's economy was 1.4 trillion USD in 2020, and the US economy was 20.9 trillion, and the Chinese economy was 14.7 trillion. In addition, in 2020 Russian military spending was 61.7 billion USD (4.41% of GDP) compared with 778.2 billion (3.7% of GDP) in military spending by the US and 252.3 billion (1.7% of GDP) by China.[136] As such, Russia's strategy thus far appears to attempt to mimic US AI development, a strategy that offers potential benefits as well as drawbacks.[137] Russia benefits by not having to incur the full range of developmental expenses associated with creating AI technologies since it frequently attempts to copy AI developments made in the US, but it suffers because it remains largely dependent on US advancement, forcing it to lag the US and robbing it of many opportunities to secure technological advantages prior the US.[138] In addition, Russia's decision to place greater focus on information warfare over other forms of AI technology may produce positive or negative results depending on the future direction of military technology. If AI becomes a dominant form of offensive military weaponry, Russia may struggle to close the gap due to financial and resource limitations. However, if information warfare becomes a more central element of future warfare, Russia may have an advantage in AI information warfare technology compared to its adversaries.

When we asked our experts the question: *What country/state is currently leading in AI development?,* none of the experts believed that Russia was leading in AI development. The experts did not have much information regarding specific AI development in Russia due to the closed nature of the regime. Many of the experts stated that this information was difficult to obtain due to the covert nature of Russia's military programmes. However, many of the experts stated that according to their knowledge, Russia's AI development programmes were not equal to the US's or China's, due to Russia's financial and resource constraints.

Discussion

In examining previous research on AI and our expert interviews, the conclusion is that AI is a technology that will profoundly affect the international balance of power and security competition in the coming years. According to this research and the AI experts we interviewed, AI has the potential to alter many aspects of defence including decision-making, logistics, data storage, data analysis, communications, and cyber and kinetic warfighting capabilities. If AI technology continues to proliferate as many expect, countless areas of security will be impacted by the evolving technology.

Regarding global AI competition, previous researchers and our experts seem divided regarding what major power state is currently leading in AI development. The experts we interviewed were evenly divided regarding what state they consider to be leading in AI development as it pertains to the US and China. This finding highlights the point raised by some of our experts that much remains to be determined regarding global AI competition, and many factors can affect what state ultimately prevails in AI development in the coming years. Thus, a more helpful question may be: *how is each major power state currently developing and applying AI* rather than *what state is currently leading in AI development*.

In considering how the major power states are developing and applying AI, our findings suggest each major power state has certain advantages and weaknesses. China is investing heavily in AI for its military (1.6 billion USD annually)[139] and appears to be undertaking aggressive developments regarding the military applications of AI, and previous research and our experts contend that China may significantly rely on AI within its military in the future for multiple purposes, including combat, due to the composition of its army and the lack of combat experience. China also appears capable of producing numerous AI applications across many sectors (domestic security, surveillance, economic, government, and military) that can be constructed in an expedient manner spanning multiple purposes, and according to our experts, China excels in many technical aspects of AI, including the creation of AI algorithms that are central to AI programmes. In addition, many of our experts highlighted the strengths and weaknesses of China's authoritarian governance style that allows China some advantages, such as greater control over the direction of research, collaboration between state and private entities, and secrecy regarding novel developments due to the less transparent nature of the regime. However, drawbacks to the authoritarian governance style are that it leads to fewer collaboration opportunities with other states and researchers, less creativity and ingenuity in the private sector, and an overreliance of AI technology for domestic surveillance of its population that could potentially lead to discontent at home which could affect development.

Regarding the US, researchers and our experts were mixed regarding the US's level of AI funding commitments. Some contended that the US is investing sufficiently in AI to keep pace with China, or to stay ahead, while others argued that the US's level of investment is modest and is insufficient to keep pace with China. Multiple AI experts highlighted the lack of US funding opportunities for ambitious "blue sky" projects that are often needed for important scientific breakthroughs. Regarding defence, the US appears to be developing and applying AI within its military in numerous areas with potential for AI assistance in multiple defensive sectors. Some researchers and experts contend that one of the US's strengths is its ability to develop AI applications that are more user-friendly that have strong real-world applications. Additional US strengths are in respect to the presence of numerous technology companies and skilled and creative AI researchers that can accelerate AI development, as well as the possibility of international collaboration that often accompanies research in a democratic and transparent regime.

In respect to Russia, many of our experts stated that Russia is significantly trailing the US and China regarding AI development and military applications. Some of our experts did not have much knowledge regarding AI development in Russia due to the closed nature of the regime, and prior research on AI development in Russia indicates that Russia is lacking in AI development and in applications across a diverse range of defensive sectors. According to our experts and research, Russia is directing much of its attention towards information warfare and AI programmes that can assist in information warfare operations. In addition, multiple AI experts noted the lack of AI funding and developmental resources in Russia. Also, multiple AI experts stated that Russia may be trailing not only the US and China regarding AI development, but also other developed states.

In conclusion, previous research and our expert interviews reinforce the pivotal role that AI will play in shaping international security in the near future. As one of our experts stated: "I do not foresee an area or industry that AI will not impact." Currently, the US and China are applying AI across numerous sectors within their economies, societies, and militaries. This trend is likely to accelerate at a rapid pace in the years ahead as AI technology becomes more powerful and efficient. In addition, Russia appears to be trailing the other major power states in developing and applying AI technology. Lastly, whether the incorporation of AI within militaries in the major power states leads to increased conflict or cooperation will largely depend on the manner it is applied and the extent that humans remain involved in important decision-making processes.

Notes

1. CNBC, 'Putin: Leader in Artificial Intelligence will Rule World', Published September 4, 2017, https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html (accessed May 20, 2022).
2. J. Pandya, 'The Weaponization of Artificial Intelligence', *Forbes*, Published January 14, 2019, https://www.forbes.com/sites/cognitiveworld/2019/01/14/the-weaponization-of-artificial-intelligence/?sh=4a26af193686 (accessed May 20, 2022).

3. M. Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: NY, Penguin Random House, 2017), 188.
4. Congressional Research Service (CRS), ‘Renewed Great Power Competition: Implications for Defense – Issues for Congress’, 2022.
5. See note 4: ii.
6. M. C. Horowitz, ‘Artificial Intelligence, Competition, and the Balance of Power’, *Texas National Security Review* 1, no. 3 (2018): 37–57.
7. A. Radin, et al., ‘China-Russia Cooperation. Determining Factors, Future Trajectories, Implications for the United States, RAND Corporation’, 2021.
8. K. Blachford, ‘The Balance of Power and the Power Struggles of the Polis’, *Journal of International Political Theory* 17, no. 3 (2021): 429–447.
9. Ibid.
10. G. Allen and T. Chan, *Artificial Intelligence and National Security* (Cambridge, MA: Belfer Center for Science and International Affairs, 2017).
11. K. Ayoub and K. Payne, ‘Strategy in the Age of Artificial Intelligence’, *Journal of Strategic Studies* 39 no. 5–6 (2016): 793–819.
12. J. Johnson, *Artificial Intelligence and the Future of Warfare: The USA, China, and strategic stability* (Manchester: Manchester University Press, 2020).
13. F. E. Morgan, B. Boudreaux, A. J. Lohn, M. Ashby, C. Curriden, K. Klima, and D. *Grossman, Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World* (Santa Monica, CA: Rand Project Air Force, 2020).
14. Ayoub and Payne, ‘Strategy in the Age of Artificial Intelligence’, 795.
15. See note 2.
16. S. Davis, ‘Artificial Intelligence at the Operational Level of War’. *Defense and Security Analysis* (2021): 5.
17. J. Sung, T. Nguyen, M. Cantos, C. P. Daniel, M. Kronauge, A. P. Mason, J. Ugolini, and M. E. Westerman, ‘AI: Using Standards to Mitigate Risks’. Public-Private Analytic Exchange Program. US Department of Homeland Security, Defense Intelligence Agency, FireEye, Guideposts, Hilliard Heintze, National Black Leadership Commission on Aids, Rand Corporation, Terbium Labs, 2018, 4, https://www.dhs.gov/sites/default/files/publications/2018_AEP_Artificial_Intelligence.pdf.
18. See note 12.
19. C. Tate and G, Allen, ‘Artificial Intelligence and National Security’, Belfer Study Center, 2017.
20. J. Johnson, ‘Deterrence in the Age of Artificial Intelligence & Autonomy: A Paradigm Shift in Nuclear Deterrence Theory and Practice?’ *Defense and Security Analysis* 36 no. 4 (2021): 422–448.
21. It is important to note that aggregated data on AI defence spending are unavailable for many countries and years due to the classified nature of AI defence spending at the aggregate level. In addition, limited information exists regarding AI funding in Russia. Lastly, the data included in the tables are from 2015–2020 because AI defence spending data are less reliable prior to 2015 and are unavailable after 2020.
22. J. Mattis, *Summary of the 2018 National Defense Strategy of the United States of America* (Department of Defense Washington United States, 2018).
23. See note 16: 5.
24. M. Weinger, ‘The Future Military Artificial-Intelligence Complex?’, *Financial Times*, 2015, https://www.ft.com/content/f589cec7-5dae-3dd8-a1ee-a4b2cd270179.
25. D. S. Hoadley and N. J. Lucas, ‘Artificial Intelligence and National Security’, *The Congressional Research Service* (2018): 13.
26. See note 22: 9.
27. C. H. Heller, ‘Near-Term Applications of Artificial Intelligence’, *Naval War College Review* 72, no. 4 (2019): 6. Department of Defense, ‘Defense Budget Overview for Fiscal Year 2021’, Office of the Under Secretary of Defense Comptroller/Chief Financial Officer, 2020.

28. AI Pathfinders 'fuse data from military, commercial and government sensors to create a common operating picture for North American Aerospace Defense Command and U.S. Northern Command' (Strout 2021). Project Maven refers to the sorting through intelligence footage to discover information and patterns (P. Tucker, 'US Navy Turns to Drones, AI to Monitor Rust', *Defense One*, 2020, https://www.defenseone.com/technology/2020/08/us-navy-turns-drones-ai-monitor-rust/168036/).
29. Heller, 'Near-Term Applications of Artificial Intelligence'.
30. Congressional Research Service (CRS), 'The Army's Optionally Manned Fighting Vehicle (OMFV) Program: Background and Issues for Congress', 2021.
31. See note 30.
32. W. Roper, 'Exclusive: AI Just Controlled a Military Plane for the First Time Ever', *Popular Mechanics*, 2020, https://www.popularmechanics.com/military/aviation/a34978872/artificial-intelligence-controls-u2-spy-plane-air-force-exclusive/.
33. Ibid.
34. S. Azeem, 'Autonomous Unmanned Aerial Vehicles: A Technology Warning Assessment', *The George Washington University, Tech Rep.*, 2012.
35. Ibid.
36. Lockheed Martin, 'Long Range Anti-Ship Missile (LRASM)', *Lockheed Martin*, 2020, https://www.lockheedmartin.com/en-us/products/long-range-anti-ship-missile.html.
37. See note 30.
38. Ibid.
39. Ibid.
40. T. Macaulay, 'The US Navy is Developing AI-Powered Submarines That Could Kill Autonomously', *TNW | Neural*, 2020, https://thenextweb.com/news/the-us-navy-is-developing-ai-powered-submarines-that-could-kill-autonomously.
41. Ibid.
42. Tucker, 'US Navy Turns to Drones'.
43. China Arms Control and Disarmament Association, 'Artificial Intelligence and Its Military Implications', Stanley Center Discussion Paper, 2019, https://stanleycenter.org/wp-content/uploads/2020/05/ArtificialIntelligence-ItsMilitaryImplications-China.pdf.
44. See note 13: 54.
45. J. Lee, H. Davari, J. Singh, and V. Pandhare, 'Industrial Artificial Intelligence for Industry 4.0-based Manufacturing Systems', *Manufacturing Letters* 18 (2018): 20–23.
46. Ibid.
47. See note 30.
48. Ibid.
49. M. Konaev, H. Chahal, R. Fedasiuk, T. Huan, and I. Rahovsky, 'US Military Investments in Autonomy and AI A Budgetary Assessment', Center for Security and Emerging Technology, 30, 2020.
50. M. Sisson, 'The Militarization of Artificial Intelligence Stimson Center', *Stimson Center*, 4, 2019, https://www.stimson.org/2020/the-militarization-of-artificial-intelligence/.
51. S. De Spiegeleire, M. Matthijs, and S. Tim, *Artificial Intelligence and the Future of Defense: Strategic Implications for Small-and Medium-sized Force Providers* (The Hague Centre for Strategic Studies, 2017), 90.
52. World Bank, 'Military Expenditure (% of GDP) – United States, China', 2022, https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS?locations=US-CN.
53. M. Konaev, H. Chahal, R. Fedasiuk, T. Huan, and I. Rahovsky, 'US Military Investments in Autonomy and AI. A Budgetary Assessment', Center for Security and Emerging Technology, 2020.
54. Ibid.
55. R. Steff and K. Abbasi, 'Artificial Intelligence and the Military Balance of Power', in *Emerging Technologies and International Security*, ed. R. Steff, J. Burton, S. R. and Soare (Routledge, 2020).
56. See note 22.

57. See note 22: 7.
58. This refers to the numerical calculations that are used in computer algorithms that run many AI programs and systems.
59. See note 16.
60. See note 16.
61. See note 52: 79.
62. Ibid.
63. See note 30: 22.
64. CNA, 'The China and Autonomy Report', 2021, https://www.cna.org/CNA_files/centers/CNA/CIP/China/ai-newsletters/ChinaAI-Autonomy-Report-Issue-1.pdf.
65. See note 13: 60.
66. Ibid.
67. See note 13: 61.
68. B. Stauffer, 'Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control', *Human Rights Watch* (2020): Para. 38.
69. See note 43.
70. See note 52: 79.
71. See note 13.
72. See note 25.
73. See note 13.
74. L. Saalman, 'The Impact of AI on Nuclear Deterrence: China, Russia, and the United States', EastWestCenter.org, 2020.
75. Ibid.
76. Ibid.
77. See note 43.
78. See note 43.
79. See note 43.
80. See note 25: 22.
81. A. Llachinski, 'AI, Robots, and Swarms Issues, Questions, and Recommended Studies', *CNA*, 2017.
82. Staff, 'War at Hyperspeed', *The Economist*, 2018.
83. See note 13.
84. Ibid.
85. See note 13: 61.
86. Xinhua News Agency, 'Scientific and Technological Innovation, a Powerful Engine for a World Class Army – The Fourth Document on the Leadership and Promotion of the Party Central Committee with Comrade Xi Jingping at its Core', Xinhua News Agency, Published September 15, 2017, http://www.gov.cn/xinwen/2017-09/15/content_5225216.html.
87. J. Hall, 'Artificial Intelligence in the South China Sea', *Global Risk Insights*, 2018: para. 3, https://globalriskinsights.com/2018/12/artificial-intelligence-turning-tide-asia-pacific/.
88. Ibid.
89. Ibid.
90. See note 13.
91. L. Saalman, 'China and Its Hybrid Warfare Spectrum', in *Hybrid Warfare: Security and Asymmetric Conflict in International Relations*, ed. M. Weissmann, N. Nilsson, B. Palmertz, and P. Thunholm (London: I.B. Bloomsbury Collections, 2022).
92. See note 7.
93. See note 13: 61.
94. See note 13: 66.
95. See note 13: 67.
96. Ibid.
97. OODA Loops refers to a dynamic decision-making process that is used by militaries and other organizations that involves four parts: observe, orient, decide, act.
98. See note 13.

99. Ibid.
100. T. Heath, ‘China’s Military Has No Combat Experience: Does it Matter?’, RAND Corporation, 2018.
101. See note 52: 80.
102. Ibid.
103. A. H. Cordesman and G. Hwang, ‘Updated Report: Chinese Strategy and Military Forces in 2021’, Center for Strategic International Studies, Published August 3, 2021, https://www.csis.org/analysis/updated-report-chinese-strategy-and-military-forces-2021.
104. H. Huifeng, ‘China Tech Firms Embrace Inward Economic Pivot, But Some Wary of “Technological Isolation”’, *South China Morning Post*, 2020.
105. D. Faggella, ‘AI in China – Recent History, Strengths and Weaknesses of the Ecosystem’, *Emerj*, 2019, https://emerj.com/ai-market-research/ai-in-china-recent-history-strengths-and-weaknesses-of-the-ecosystem/.
106. Ibid.
107. S. Feldstein, ‘The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression’, *Journal of Democracy* 30 no. 1 (2019).
108. See note 106.
109. CNA, ‘Artificial Intelligence and Autonomy in Russia’, 2021, https://www.cna.org/CNA_files/centers/CNA/sppp/rsp/russia-ai/Russia-Artificial-Intelligence-Autonomy-Putin-Military.pdf.
110. See note 25: 12.
111. See note 25: 25.
112. Y. Borisov, ‘The Development of Artificial Intelligence is Essential for the Successful Conduct of Cyberwarfare’, Ministry of Defense of the Russian Federation, 2018.
113. See note 110.
114. R. Norvosti, ‘Putin Spoke About the Latest Types of Russian Weapons’, РИА Новости, 2018, https://ria.ru/20180301/1515566394.html.
115. See note 110.
116. A. Nadibaizde, ‘Russian Perceptions of Military AI, Automation, and Autonomy’, Foreign Policy Research Institute, 2022.
117. Tate and Allen, ‘Artificial Intelligence and National Security’, *Belfer Study Center*, 2017, 21.
118. See note 25.
119. See note 52: 82.
120. Ibid.
121. R. J. Marks and S. Bendett, ‘Russia is Systematically Copying U.S. Military AI Robotics’, Mind Matters, 2020, https://mindmatters.ai/2020/10/russia-is-systematically-copying-u-s-military-ai-robotics/.
122. See note 75: 41.
123. Ibid.
124. B. Wodecki, ‘Russia’s AI Army: Drones, AI-Guided Missiles and Autonomous Tanks’, Internet of Things World Today, 2022.
125. Ibid.
126. Ibid.
127. See note 125: para. 12.
128. See note 75.
129. See note 75: 41.
130. S. Bendett, ‘In AI, Russia is Hustling To Catch Up’, *Defense One*, 2018, https://www.defenseone.com/ideas/2018/04/russia-races-forward-aidevelopment/147178/.
131. M. Peck, ‘Russia Wants To Use AI-Sea Mines To Sink America’s Navy’, *The National Interest*, 2020, https://nationalinterest.org/blog/buzz/russia-wants-use-ai-sea-mines-sink-americas-navy-120951.
132. R. J. Marks and S. Bendett, ‘Russia Aims To Close the Technology Gap with the United States’, Mind Matters, 2020, https://mindmatters.ai/2020/10/russia-aims-to-close-the-technology-gap-with-the-united-states/.

133. Ibid.
134. V. Boulanin, L. Saalman, P. Topychkanov, F. Su, and M. Carlsson, ‘Artificial Intelligence, Strategic Stability and Nuclear Risk’, Stockholm International Peace Research Institute, 2020.
135. See note 110: 38.
136. See note 53.
137. See note 122.
138. Ibid.
139. J. Harper, ‘China Matching Pentagon Spending on AI’, National Defense, 2022.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Lance Y. Hunter, PhD, is an Associate Professor of International Relations in the Department of Social Sciences and Master of Arts in Intelligence and Security Studies program at Augusta University located in Augusta, GA, USA. His expertise is in security studies and democratization. His research focuses on the causes and effects of terrorism and interstate conflict, democratization, and the relationship between evolving technology and conflict.

Dr. Craig D. Albert, PhD, is Professor and Director of the Master of Arts in Intelligence and Security Studies at Augusta University. He received his PhD from the University of Connecticut in 2009. His areas of concentration include international security studies, ethnic conflict, cyberterrorism, and cyberwar.

Christopher Hennigan is a consultant in Deloitte’s Government & Public Services firm. He is a graduate of Augusta University’s Masters in Intelligence and Security Studies program. His expertise is in Six Sigma and PMO, business risks and mitigation, and machine learning process automation. His research focuses on machine learning, cybersecurity, terrorism, and AI game theory.

Josh Rutland is a graduate of Augusta University’s Masters in Intelligence and Security Studies program. He currently works as a researcher in the Augusta University Department of Emergency Medicine and will soon be employed by U.S. Army Cyber Command as an Information Technology Specialist. His research focuses on information warfare, cybersecurity, terrorism, and biosecurity.

ORCID

Craig D. Albert http://orcid.org/0000-0003-3225-9386

Neurocognitive hacking
A new capability in cyber conflict?

John J. Heslen, *Augusta University*

Abstract. **This article presents a discussion of neurocognitive hacking and its potential for use at the strategic, operational, and tactical levels of cyber conflict. Neurocognitive hacking refers to the ability to activate specific neural areas of the brain, via subliminal or supraliminal stimuli, to shape the behavioral outcomes of an adversary. Research suggests that awareness of mortality-related stimuli has neural correlates in the right amygdala and left anterior cingulate cortex and mediates negative behavior toward out-group members, including unconscious discriminatory behavior. Given its in-group/out-group dynamic, the phenomenon could be exploited for use in information operations toward target populations, specifically ones that are multiethnic, multicultural, or multireligious. Although development of the theoretical framework behind neurocognitive hacking is ongoing, mortality-related stimuli are proposed to activate one's unconscious vigilance system to further evaluate the locus and viability of the suspect stimuli. Research suggests that the subsequent discriminatory affective reactions directed toward out-group members are representative of automatic heuristics evolved to protect the organism in the event a stimulus represents a more serious threat to survival. Therefore, presenting mortality-related stimuli over computer networks to targeted audiences may facilitate the ingestion of tailored propaganda or shaping of specific behavioral outcomes within a population, including sowing division in a target community or weakening support for a specific political regime.**

Key words: Subliminal stimuli, terror management theory, psychological operations, information operations, persuasion, cyberwar, propaganda, mortality bias

The use of propaganda in war likely dates to the dawn of civilization. Its methods are constantly being updated and improved to match current advancements in communications technology. As propaganda (considered a type of information operation) has been inextricably linked with war, employment of these capabilities by major world powers will likely increase in what are referred to as "gray zones" as the dawning of the nuclear age has made kinetic warfare between them far too risky. Philip M. Taylor, in his important work *Munitions of the Mind: A History of Propaganda from the Ancient World to the Present Day* (2003), noted that with the advent of nuclear weapons, war between nuclear-armed adversaries increasingly is prosecuted within the information space. The use of propaganda in information wars between great powers has now become "part of the struggle for perceptions in which words attempt to speak as loud as actions, and sometimes even replace the need for action" (Taylor, 2003, p. 8). In fact, one has to look no further than the 2016 U.S. presidential election to get a glimpse of the new world of great-power information conflict. For example, the U.S. Intelligence Community report assessing Russian hacking activities during the presidential election noted that one of Russia's primary goals was to "undermine the US-led liberal democratic order" (Office of the Director of National Intelligence, 2017, p. ii).

More recently, a European Commission report outlined the "sustained" disinformation campaign by the Russian government to depress voter turnout and influence voter preferences during the 2019 European parliamentary elections (European Commission, 2019). As a result, many in the West are now well acquainted with the dangers of propaganda, sometimes colloquially referred to as "fake news." And there is worry among cyberwarfare analysts that in the future, political conflict utilizing information operations may become ubiquitous

doi: 10.1017/pls.2020.3

Correspondence: John J. Heslen, Augusta, Department of Social Sciences, Political Science, 1120 15th Street, Augusta, Georgia, 30904-4562. Email: *jheslen@augusta.edu*

as it offers nation-states the ability to act covertly while cyber deterrence measures remain dangerously underdeveloped (Valeriano & Jensen, 2019). In fact, many nation-state militaries increasingly are developing and utilizing a suite of cognitive tools to influence and persuade target populations referred to as CAMO, or "cognitive aspects of military operations" (Astorino-Courtois, 2017).

Definitions and conceptual issues

The *Oxford English Dictionary* defines propaganda as "the systematic dissemination of information, especially in a biased or misleading way in order to promote a political cause or point of view." However, because the concept was not sufficiently comprehensive to describe the full range of influences involved in persuasive activities (in addition to the negative connotations the word acquired from its use by the Nazi and Soviet regimes), it fell out of favor in much of the West in favor of "persuasion," considered a more comprehensive and less polarizing term (Markova, 2008). Still, the two words have a tendency to be used interchangeably, but scholars have proposed an interesting dichotomy to differentiate between them. Propaganda is conceptualized as a one-way "monologic" communication from a "source" to a "receiver," with the goal of "transform[ing] the heterogeneous thought of individuals into those of a homogenous 'collective mind' of masses, and to lead those masses to a specific action" (Markova, 2008, p. 41). In other words, propaganda can be thought of as a unidirectional communication in which the source of the message holds the predominance of power with regard to the ability to create, change, or normalize the social reality of the receiver (Markova, 2008).

In contrast, persuasion is conceptualized as a two-way "dialogic" communication in which the power between the source of the message and the receiver is more broadly shared, but it is also conceptualized to include one's internal dialogue and unconscious aspects of thought. However, unlike propaganda, in which the more powerful source seeks to "fuse" its reality with that of the receiver, in persuasion, the source's aim is to "convince the other party of one's own case and of the superiority of one's own idea or belief over that of the [receiver]" (Markova, 2008, p. 45). With this in mind, neurocognitive hacking is proposed to support the role of propaganda by making its ingestion more likely and facilitating persuasion by cultivating a neural environment in the receiver more accommodating to the source's narrative, especially when it involves in-group/out-group dynamics.

For the same reasons discussed earlier, "psychological operations" and "information operations" are terms requiring clear distinctions. Both words are typically used in relation to nation-state-sponsored military or civilian intelligence operations, but they have different operational scopes. The U.S. Department of Defense defines psychological operations (PSYOPs) as "planned operations that convey selected information and indicators to foreign target audiences to influence their emotions, motives, objective reasoning, and ultimately, the behavior of foreign governments, groups and individuals" (U.S. Department of the Army, 2003, p. GL-8).

In modern military operations, propaganda is conceptualized as a tool of PSYOPs. However, because of innovations and advancements in technologies complementary to PSYOPs, along with the efficiencies gained by employing them in concert with other supporting capabilities, Western militaries increasingly refer to information operations as

> the integrated employment of the core capabilities of electronic warfare, computer network operations, psychological operations, military deception, and operations security, in concert with specified supporting and related capabilities to influence, disrupt, corrupt or usurp human and automated decision making while protecting our own.
>
> *(Larson et al., 2009, p. xiii).*

Additionally, although brain science is considered by many to be in its infancy, thanks to recent developments in brain imaging technology, the field is advancing rapidly and gaining insights into once-invisible processes (Jorgenson et al., 2015). As artificial intelligence tools are incorporated into these research efforts, the rate of discovery will likely only increase; however, there are many ongoing foundational debates regarding brain function that have yet to be resolved, such as the relationship between one's evaluation (i.e., attitude) of an object or concept and subsequent behavior toward it (Ajzen & Cote, 2008). As neurocognitive hacking proposes the ability to utilize specific stimuli to activate neural structures (e.g., amygdala and anterior cingulate correlate) for the purposes of exploiting affective reactions and shaping targeted *behavioral* outcomes, a quick treatment of the mood/attitude–behavior linkage is offered.

Conceptually, regarding the reactions to mortality-related stimuli, the author agrees with the logic of Tritt and colleagues (2012) and their use of the term "affect" versus "moods" or "emotions," as it avoids unnecessary theoretical and sematic "baggage." For example, there are ongoing debates regarding the extent to which (conscious) cognitive processes are involved in emotion (Winkielman & Berridge, 2004) and attitude formation (Devos, 2008). However, by using the construct of "affect" to describe moods or emotions that are inclusive of behavioral phenomenon occurring below conscious awareness (e.g., mortality-related stimuli) and thus opaque to self-reported measures, these controversies are largely avoided.

In fact, research by Winkielman and colleagues (2004) suggests that affective responses intense enough to influence one's behavior can remain below awareness. Therefore, reactions to mortality-related stimuli (M-rS) will not be conceptualized in terms of an attitude-mediated behavioral construct (e.g., MODE model, theory of planned behavior) but as a biologically construed "mortality bias" composed of unconscious affective reactions (Winkielman et al., 2005), activated by an energized threat detection system (Tritt et al., 2012) for the purposes of reducing false negative (Type II) errors (Haselton & Buss, 2000) and "psychological uncertainty" (Tritt et al., 2012). A more in-depth discussion of this topic will be presented later in the article.

Psychological operations in conflict

Taylor (2003) highlights many fascinating uses of propaganda throughout history, covering operations from ancient Greece to the post–Cold War era. His well-regarded book highlights the evolution of propaganda and explicates how its methods generally mirror the overall advancement of technology in a society. In ancient Greece, for example, architectural marvels like the Acropolis were used to persuade citizens and noncitizens alike of the superiority of Greek culture. More than two millennia later during the Cold War, Americans utilized radio (e.g., Voice of America) in their propaganda efforts against the Soviet bloc. As communications technology in the information age continues to advance, its technologies will provide increasingly rich support for disseminating propaganda in new digital formats to match the current crop of polarizing memes, "fake news" on social media, and "deepfake" manipulations of video clips.

Regardless of the polarizing nature of the concept, Taylor suggests that readers approach propaganda as a morally neutral concept best looked upon as a "process for the sowing, germination, and cultivation of ideas" (2003, p. 2). He notes that the Vatican operationalized this process during the Protestant Revolution in Europe to defend itself against heretics with such success that even today, Catholics and Protestants can still be distrustful of one another. This example alone could attest to the incredible power of propaganda; however, Taylor is quick to note that it alone is not enough to win conflicts but is most effective when integrated with other levers of power, such as diplomatic, military, or economic.

Contemporary history is replete with examples of political and military leaders using propaganda and other broader forms of PSYOPs to their benefit—or their peril, if they failed to incorporate it into their repertoire of operational capabilities. For example, John Nagl's *Learning to Eat Soup with a Knife* (2002) highlights the positive role PSYOPs played during the 1940s–1950s British fight against the communist insurgency in Malaysia. Nagl points out the British Army after World War II was more of a "learning organization" compared with their American counterparts and more willing to experiment with unconventional tactics, including the use of PSYOPs to win "hearts and minds." To a large extent, this resulted from the British Army's experience fighting many "small wars" in remote locations over the previous two centuries and being forced to adapt to myriad enemies and operating environments. However, because the British did not have inexhaustible resources to bring to bear on these limited wars, they focused heavily on understanding (and exploiting) the motivations of their enemies, an effort often made easier by partnering with indigenous civil authorities.

This approach prompted experimentation with various tools of persuasion and influence to help the British exploit and ultimately break the will of their enemies. In Malaysia, for example, British field commanders were given flexibility to employ various PSYOP tactics (including blasting propaganda from loudspeakers mounted on airplanes, placing bounties on the heads of insurgent leaders, and dropping leaflets over enemy territory), an effort that by 1960 largely proved successful (Nagl, 2002).

Nagl highlights how the U.S. Army, in contrast with the British, was less apt to incorporate PSYOPs into military operations as the Americans relied on (and could afford to use) a policy of fielding much larger units, intent on using overwhelming force to exact total destruction on the enemy. He notes that when General William Westmoreland took command in Vietnam, his army was sorely lacking in its understanding of PSYOPs and thus its preparation was "inappropriate to the demands of counterinsurgency warfare in South Vietnam" (Nagl, 2002, p. 174). American deficiencies were compounded by the strength of PSYOPs conducted by their National Liberation Front foes, who, as the scholar Francis Fitzgerald pointed out, were geared toward inculcating the "systematic encouragement of hatred" of the United States. Similarly, having recognized the motivational benefit of generating hatred within the Vietnamese people toward their French occupiers a generation prior, Ho Chi Minh purportedly said, "I have no army, I have no finances, I have no education system, I only have my hatred" (Fitzgerald, 1972, p. 169). Hate is a powerful motivator and the rise of the internet, in combination with other information communication technologies and platforms, has made it easier for both separatists and terrorist groups to generate and exploit it for specific purposes.

One of the more recent and concerning innovations of terrorist organizations is their ability to utilize social media for radicalization and recruitment. Communications scholar Gabriel Weimann (2015) notes that social media has given terrorists groups an enormous advantage because of the "effectively limitless" audience it creates for recruitment and the ease with which propaganda videos can be uploaded in response to dynamic operating environments. Two cases from nearly a decade ago highlight the speed and effectiveness with which PSYOPs can be used in the cyber domain to support the terrorist radicalization process.

In 2011, Arid Uka, an Albanian Muslim immigrant to Germany, admitted to becoming self-radicalized as a result of consuming online jihadist propaganda. In roughly six months, after watching numerous propaganda videos, including one doctored by the Islamic Movement of Uzbekistan falsely depicting U.S. soldiers sexually assaulting Iraqi and Afghan women, Uka drove to nearby Frankfurt Airport and shot and killed two U.S. military personnel transiting from a base in the United Kingdom (Bohleber & Bohleber, 2012).

Similarly, the case of British citizen Roshonara Choudhry (Pearson, 2015) highlights a rare instance of a female terrorist attacker being radicalized to action online. In early 2010, Choudhry, a fairly typical university student working toward completing her degree at King's College, London, admitted to becoming radicalized as a result of spending several months viewing hours of propaganda videos. Many of these videos featured the well-known jihadist Sheikh Abdullah Azzam and the American radical propagandist Anwar al-Awlaki. By the end of the academic year, Choudhry had dropped out of university, become estranged from most of her friends and family, and attempted to stab to death a British member of Parliament. Her case was highly unusual in that most jihadist attacks are committed by men, with the vast majority of Islamic propagandist urging women to participate in support actions only. However, during questioning after her arrest, she noted that she overcame these gender and ideological barriers to action after viewing videos of Sheikh Abdullah Azzam, who decreed "even women" had a duty to engage in jihadist attacks (Windsor, 2018).

Information operations in the digital era

Although the capabilities of terrorist groups to conduct PSYOPs for purposes of recruitment and radicalization present a formidable challenge to global security, they generally lack the resources needed to conduct full-spectrum information operations. The science of information operations in the digital era is advancing rapidly, and it is increasingly characterized by the "integrated employment, during military operations, of information-related capabilities in concert with other lines of operation to influence, disrupt, corrupt, or usurp the decision-making of adversaries and potential adversaries while protecting [one's] own" (U.S. Department of Defense, 2014, p. ix). Modern information operations apply theories developed from the study of persuasion and motivation, and the capability is becoming increasingly sophisticated as it parallels advances in those fields (Jowett & O'Donnell, 1986). These advances are supported by new imaging tools such as functional magnetic resonance imaging (fMRI) and co-registration techniques that combine old and new technologies to produce more comprehensive scans. These new imaging tools have illuminated previously unknown neural processes in the brain responsible for interpreting the social world, including highlighting the importance of neural structures such as the amygdala in evaluating stimuli with emotional value. As many of these processes are known to occur below the level of consciousness (Adolphs et al., 1995) they underscore the susceptibility of neural components like the amygdala to manipulation as their activation may be exploited to shape targeted behavioral outcomes.

In fact, research suggests that *subliminal* mortality-related stimuli (e.g., an image of a dead body) can activate the amygdala and other neural components associated with the processing of threat stimuli (Quirin et al., 2011). Therefore, exploiting these unconscious processes, combined with the ability to commandeer communication networks supporting smartphones and other electronic media devices, is proposed to offer a powerful tool for increasing the ingestion of propaganda and subsequent shaping of perception and behavior of adversaries. It should be noted that although the effects on behavior from neurocognitive hacking are conceptualized to be small, the ability to "nudge" a small group of people in one direction can have enormous strategic consequences, as highlighted by the fact that the 2016 U.S. presidential election was determined by fewer than 80,000 votes in three states (Bump, 2016).

To highlight the growing strategic relevance of information operations in interstate conflict, consider the indictments by the U.S. Office of Special Counsel. It charged Russian agents with interfering in the 2016 presidential election by seeking to "sow discord" within the U.S. political system (U.S. Department of Justice, 2018). As part of their efforts, Russian operatives were accused of creating numerous social media handles geared toward increasing public polarization over several highly contested social issues. According to the indictment, these social media accounts sought to inflame attitudes over issues such as African American and Muslim civil rights, using provocative names, including "Woke Blacks" and "United Muslims of America." Although it is difficult to calculate the extent to which this propaganda influenced the public, consider if the targeted audiences had first been primed with mortality-related stimuli that prompted the activation of neural correlates related to threat perception and response. It is proposed that such a targeted neurological attack on these neural structures would have increased the rate of propaganda ingestion by the targeted audience and subsequent level of societal discord within the United States. The ability to influence or "hack" the perception of adversaries to shape behavioral outcomes is of increasing interest to national security stakeholders in the United States, many of whom assess the United States' current capability in this area to be critically underdeveloped (Astorino-Courtois, 2017).

Why mortality-related stimuli?

Research utilizing priming stimuli is controversial; however, a meta-analysis by Weingarten et al. (2016) and research by Winkielman et al. (2005) suggest they can have a significant effect on behavior. These effects appear even greater when primes incorporate the use of words with negative valence (Nasrallah et al., 2009). In fact, robust research exists suggesting that priming with mortality-related stimuli mediates prejudicial behavior toward out-groups and their respective cultural symbols (Greenberg et al., 1990). After perceiving mortality-related stimuli, individuals have been shown to unconsciously exhibit negative biases toward out-groups, including increased hostility and willingness to engage in avoidance behavior, as well as a greater willingness to actively denigrate out-group cultures (Burke et al., 2010; Greenberg et al., 1990). The preponderance of research utilizing mortality-related primes has traditionally been subsumed under a theoretical construct referred to as terror management theory (TMT).

Motivated by the writings of Ernst Becker, TMT's theoretical framework is controversial, as it asserts that during the course of evolution, humans reached a level of cognitive sophistication enabling an awareness of the inevitability of death (Greenberg et al., 1986). According to TMT, the inexorable nature of death created such maladaptive terror in humans that it prompted the species to generate a suite of psychological coping mechanisms referred to as "cultural anxiety buffers" (Rosenblatt et al., 1989). When mortality-related stimuli are salient in the environment, the buffers are theorized to work by managing the paralyzing effects of death awareness. Therefore, they are purported to facilitate a deeper fusion with one's cultural worldview to give symbolic immortality and mitigate the finality of biological death (Solomon et al., 2004). To test these assertions, TMT researchers utilize a priming manipulation referred to as "mortality salience."

The effects of mortality salience (i.e., awareness of mortality-related stimuli) are explored in research settings by asking subjects to consider the ramifications of their physical death and examining their subsequent behaviors. In general, TMT researchers have found that priming with mortality-related themes tends to facilitate "worldview defense" (a worldview comprising the foundational beliefs one holds to help understand and interpret the world). A worldview defense, therefore, is roughly defined as actions involving the defense of one's culturally based belief system. Examples include exhibiting negative discriminatory behavior toward out-groups (i.e., those holding different worldviews) or positive discrimination toward one's in-groups (Greenberg et al., 1990).

Although experiments using mortality salience manipulations to induce prejudiced reactions are numerous (Greenberg et al., 1990; Rosenblatt et al., 1989), the framework suggested by TMT has been criticized on several theoretical grounds (Fessler & Navarrete, 2005; Holbrook et al., 2011; Kirkpatrick & Navarrete, 2006). For instance, researchers in the field of coalitional psychology assert that reactions to mortality salience (i.e., mortality-related stimuli) result from a series of prosocial behaviors that evolved in humans to better coordinate social interactions within in-groups, especially behaviors that would be adaptive when reacting to crises or threatening situations (Kirkpatrick & Navarrete, 2006). Coalitional psychology offers a cogent explanation for why mortality-related stimuli affect social relations, and it, along with other complimentary theories (Haselton and Buss, 2000; Tritt et al., 2012), provides a well-grounded explanation for the biological foundations of the phenomenon.

Error management theory, an evolutionary perspective on the development of cognitive biases, proposes that under conditions of uncertainly, reactions to adaptively relevant stimuli (e.g., dead animals, potential threats from out-groups, or sexual opportunities) are biased toward false positive (Type I) errors. For example, men have been shown to overestimate the extent to which women have sexual interest in them, as this bias would likely facilitate greater reproductive success over the long term. Similarly, when primed with fear, individuals presented with neutral faces have been shown to attribute more anger to them (Haselton et al., 2009), a bias that is likely to have proved protective during the environment of evolutionary adaptiveness. Given that mortality-related stimuli could resolve into actual existential threats (e.g., finding a dead body in one's immediate environment could represent the presence of a lethal attacker or a lethal disease), response heuristics encoded to avoid making false negative (Type II) errors would likely have been adaptive.

Regarding the biological mechanics of reactions to mortality-related stimuli, Tritt et al. (2012) propose the existence of a "biological anxiety system" activated by states of "psychological uncertainty." When psychological uncertainty results from a "mismatch" between one's actual and expected reality, a component of this biological anxiety system, referred to as the behavioral inhibition system (BIS), is triggered. As Tritt and colleagues (2012) note, the BIS is thought to be integral to both the function of anxiety and approach-avoidance dynamics, as it activates inhibitory neural components located in the right hemisphere of the brain. This explanation comports with findings by Quirin et al. (2011) showing activation of the right amygdala and left anterior cingulate cortex (ACC) subsequent to priming with mortality-related stimuli.

Research suggests that subliminal stimuli can trigger affective reactions without conscious awareness (Winkielman & Berridge, 2004); however, there is controversy over this issue. Addressing this, Custers (2009) notes that most models of human goal pursuit conceptualize a *conscious* mechanism for determining whether a goal will be pursued based on the "expected value" of the goal. However, because reactions to subliminal goal primes can occur outside of conscious awareness, Custers asserts that the most logical mechanism for determining the expected value of a goal outside of conscious awareness "would be one that relies on affective [not conscious] processes" (2009, p. 179). Additionally, based on research suggesting that there are differentials in the amplitude of affective reactions to valenced stimuli (Holbrook et al., 2011), there are likely only a few affective reactions that are as impactful on behavior as those induced by mortality-relevant stimuli.

With the foregoing theoretical framework in mind, the author proposes the term "mortality bias" (Heslen, 2016) to capture the suite of automatic processes related to reactions to mortality-related stimuli.

Role of the amygdala in shaping perception and behavior

Although several neural structures are thought to influence social cognition and human decision-making in concert with the amygdala (Adolphs, 2003, 2009), the structure is of particular interest to the conceptualization of neurocognitive hacking given growing evidence of its involvement in the processing and encoding of emotion, fear, and ambiguity within social contexts (Whalen, 1998).

The amygdala is an almond-shaped structure located deep within the temporal lobes of the brain. Traditionally, it was thought to function solely for evaluating threat-related information; however, it is now assessed to be part of an early vigilance system that works with other neural structures to process the emotional value of stimuli. As such, researchers have surmised that the amygdala may act as a complex area for processing the "social, homeostatic, and survival-related meaning of a class of complex stimuli, such as facial expressions of some emotions" (Adolphs et al., 1995, p. 5889). This view of the amygdala's role in facilitating emotional evaluations is supported by other research suggesting that it is more sensitive to negative *animate* versus negative *inanimate* stimuli in human social contexts. For example, subjects presented with subliminal images of both threatening animals and threatening inanimate objects, such as pointed guns, experience greater amygdala activation in response to the threatening animals (Fang et al., 2016). Again, this finding suggests that the amygdala plays an important role with other neural components in linking emotional valence to specific objects.

Combining evidence suggesting that unconscious perception of mortality-related stimuli activates the right amygdala and the left rostral ACC with findings suggesting that an activated amygdala is correlated with the propagation of unconscious racial stereotypes (Phelps et al., 2000), a logical leap has been made suggesting that the discriminatory behaviors induced by mortality-related stimuli are mediated by the activation of these neural structures. Therefore, by integrating research suggesting that individuals continuously (and automatically) update their social evaluations of others (Wheeler & Fiske, 2005) with evidence suggesting that the amygdala plays a significant role in facilitating these automatic evaluations (Adolphs et al., 1995), neurocognitive hacking proposes the ability to exploit this interplay to shape perceptual and behavioral outcomes of targeted audiences.

Additionally, scholarship suggests that subliminal exposure to mortality-related stimuli mediates behavior in a similar manner as conscious (supraliminal) exposures. For example, in one experiment, American participants subliminally primed with the word "death" were more critical of anti-American essays than participants who were subliminally primed with the word "field." Interestingly, terror management theory researchers have found that reactions (e.g., discrimination toward out-groups) occurring subsequent to *conscious* mortality primes do not manifest until after a short distraction exercise is given. This is not the case with subliminal mortality-related primes, the effects of which can be immediately observed without a distraction (Arndt et al., 1997).

The need for a distraction exercise has been proposed by Heslen (2016) to involve the "System 1" and "System 2" dual-processing cognitive construct suggested by Stanovich and West (2000) and popularized by Daniel Kahneman in his book *Thinking*, *Fast and Slow* (2011). In general, while both systems are believed to work in a complimentary fashion, System 1 is thought to comprise the suite of automatic mechanisms that constantly evaluate and respond to environmental stimuli, whereas System 2 facilitates more conscious, deliberative functions. As such, the distraction exercise following conscious awareness of mortality-related stimuli likely interrupts the conscious appraisals of System 2, thus giving primacy to System 1 dynamics, where automatic behavioral heuristics are generated.

Cognitive aspects of military operations (CAMO)

Several of the United States' strategic adversaries, including Russia, have been involved in researching the cognitive and psychological aspects of information warfare for decades (Thomas, 2004). However, within the U.S. defense establishment, there is growing recognition that the armed forces are behind in this area of research and lack the ability to incorporate knowledge of the "human/cognitive domain" into military operational planning. As opposed to the physical/kinetic domain (e.g., weapons systems and personnel training), where the United States is considered dominant, its utilization of the human-cognitive domain (i.e., the ability to influence "attitudes and behaviors of populations or opponent forces by manipulating information and otherwise preying on human perceptual vulnerabilities"; Astorino-Courtois, 2017, p. 6) is proposed to be lacking. It is a space, however, in which United States' major adversaries are assessed to have invested heavily. Increasingly, this suite of cognitive tools is conceptualized as the cognitive aspects of military operations (CAMO) and incorporates techniques to exploit three key psychological functions: cognition, affect, and conation (Astorino-Courtois, 2017).

In 2016, the Strategic Multilayer Assessment Office of the U.S. Department of Defense published a white paper titled "A Cognitive Capabilities Agenda: A Multi-Step Approach for Closing DoD's Cognitive Capability Gap" (Astorino-Courtois, 2017). The paper made several recommendations for how the United States could close the cognitive capabilities gap proposed to exist with its strategic adversaries, including Russia, China, Iran, and North Korea. In addition to recommending updates to doctrine, it called for the Defense Department to increase funding for "actionable cognitive research" and the development of "analytic tools" to integrate cognitive capabilities with the physical/kinetic aspects of warfighting (Astorino-Courtois, 2017). The ability to operationalize these cognitive capabilities in "gray zones" (characterized as intense areas of military competition that fall short of conventional war; see Votel et al., 2016) is an area in which Western allies, given adequate investment, can significantly increase their effectiveness, in light of recent scientific advancements.

Among the West's strategic competitors, Russia is assessed to be the most advanced with regard to their research efforts and ability to execute cognitive operations. In fact, during the 1960s and 1970s, Russia developed a theory of information warfare referred to as "reflexive control" to maximize advantages in both cognitive and computer-based decision-making processes. Reflexive control has been defined as a method of deception to relay "specially prepared information to incline [adversaries] to voluntarily make [a] predetermined decision desired by the initiator of the action" (Thomas, 2004, p. 237). Russian military theorists assert that during a conflict, the combatant with the greatest understanding of enemies' moral, cognitive, and psychological underpinnings, including those of senior decision makers, tend to be most successful because they are better prepared to induce the enemy into making adverse decisions (Thomas, 2004).

The need to increase understanding of the cognitive tools used to manipulate perception was well illustrated by the 2016 U.S. Intelligence Community report on Russian hacking and the subsequent indictment filed by the U.S. Office of Special Counsel in February 2018. The 2016 report, "Background to 'Assessing Russian Activities and Intentions in Recent U.S. Elections'" claimed that Russia tried to influence the 2016 election to "undermine the U.S.-led liberal democratic order" and "public faith in the U.S. democratic process" (Office of the Director of National Intelligence, 2017, p. ii). The later indictment filed by the Office of Special Counsel highlighted both the organizations involved and the tactics used by Russian operatives to accomplish these goals. For instance, the special counsel charged the Russian Internet Research Agency with election interference in part for its role in creating highly polarized fake social media accounts.

These accounts were geared toward spreading inflammatory and derogatory information for the purposes of creating strife in the U.S. political system. Fake social media profiles such as "Woke Blacks," "Blacktivist," and "United Muslims of America" were established in an attempt to suppress the minority vote during the election. Examples of messages on these sites included memes such as, "Hillary Clinton Doesn't Deserve the Black Vote," "Hillary is a Satan…," and "Donald wants to defeat terrorism…Hillary wants to sponsor it" (U.S. Department of Justice, 2018). Given the strong reinforcing dynamics that group membership can exact on individual behavior (Glassner, 1985; Jost et al., 2016), the efficacy of these messages, when engineered to exploit central characteristics of group identity, should not be underestimated (Hogg et al., 1995).

To understand how a constant stream of socially engineered messages can influence one's emotional state, consider a study conducted by Facebook several years ago. In early 2012, the social media giant initiated an extraordinary experiment aimed at manipulating the emotional states of 700,000 of its newest members. For one week, the site changed these users' newsfeeds to display a preponderance of either happy/positive or sad/negative news stories for the purposes of assessing any effects on their emotional states. At the end of the week, depending on their respective treatment, users did show a propensity to post either positive or negative words, providing evidence of an "emotional contagion" effect (Meyer, 2014).

Consistent with this research, in 2016, the chief executive officer of Cambridge Analytica—the data analytics firm that assisted multiple African elections as well as, later, the Trump presidential campaign—stated, "If you know the personality of the people you're targeting, you can nuance your messaging to resonate more effectively with those key audience groups" (Nyabola, 2019, p. 1). A former employee of Cambridge Analytica turned whistle-blower offered a more succinct and colloquial analysis of his company's mission, noting that the company had succeeded in developing a "psychological warfare mind-fuck tool" (Halpern, 2018).

In fact, during testimony to the U.K. Parliament, the employee accused Cambridge Analytica of specifically developing software tools, known as psychometrics, to target voters in the 2016 U.S. presidential campaign (National Public Radio, 2018). The use of these psychometrics to identify individuals or populations susceptible to socially engineered propaganda, in concert with the ability to manipulate neural areas of the brain to unconsciously shape behavior and disrupt decision-making abilities (i.e., neurocognitive hacking), foreshadows the sophisticated and potentially dangerous future of information operations in cyber conflict.

Neurocognitive hacking

The potential to secretly exploit adversary computer networks to prompt users with stimuli for the express purpose of manipulating neural structures and influencing political behavior is sobering. This capability is even more concerning as it can likely be accomplished using subliminal imagery of which a target audience is unaware. Because of its focus on manipulating actual neural structures in the brain, it goes beyond what has been conceptualized as "cognitive hacking," a phenomenon more concerned with manipulating perception through the use of deception. For example, a common illustration of cognitive hacking involves the case of Mark Jakob, who created a series of false media releases to lower the cost of a specific stock and subsequently realized a significant profit (Cybenko et al., 2002). Although cognitive hacking can be covert and include the manipulation of perceptions, the concept does not address shaping behavior through the targeted activation of neural structures.

A simple way to conceptualize neurocognitive hacking is through the lens of a Russian propaganda sample used prior to the 2016 U.S. presidential election. Below is a highly inflammatory anti–Hillary Clinton advertisement that surfaced in the months leading up to the election. Although there is no way to quantify the amount of influence this or other ads may have had on the undecided electorate, it is possible that enough voters were swayed in critical states to have influenced the outcome of the election. Consider if, prior to viewing such an ad (Figure 2), subliminal, mortality-related images (Figure 1) were presented to activate the amygdalae and its associated neural components (i.e., threat response activation) of potential voters. Based on knowledge of the effects of mortality-related stimuli on in-group/out-group behavior, it is possible the images would have generated even greater emotion resonance and anti-Clinton sentiment on behalf of undecided voters.

PSYOPs personnel have been known to utilize disinformation campaigns (e.g., spreading rumors) to exacerbate underlying levels of societal conflict. And, with the advent of social media platforms, disinformation can spread more quickly and even generate lethal violence against perceived out-groups. For example, in early 2018, at least nine people were killed in India as a result of a rumor being spread in rural communities regarding the existence of a child abduction ring (Dwoskin & Gowen, 2018). Using WhatsApp, tens of thousands of unsuspecting citizens spread video images of a faked child abduction, resulting in enraged local mobs attacking and killing innocent strangers suspected of involvement. Therefore, it is likely that the use of neurocognitive hacking techniques to activate threat-related neural structures of a targeted audience, prior to initiating a disinformation/rumor campaign, would increase both effectiveness and the rate of dissemination.

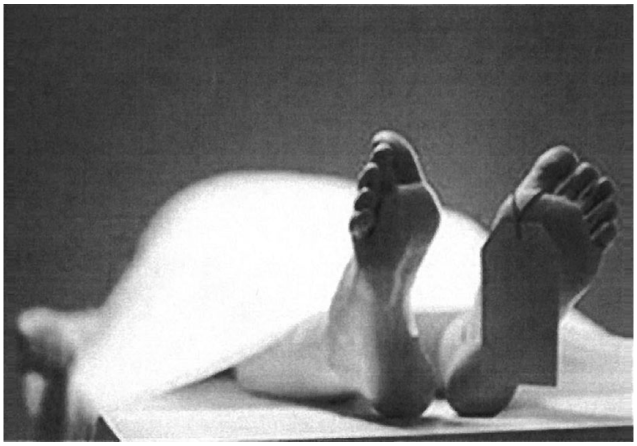
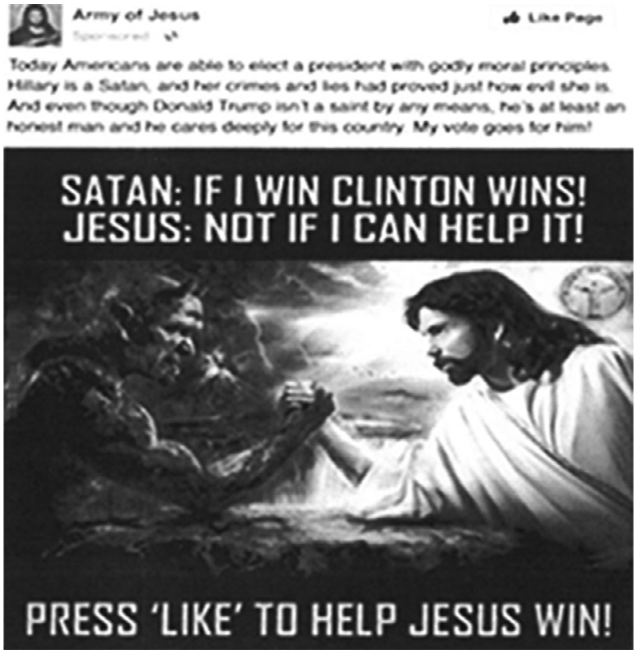


Figure 1. A dead body on a gurney is an example of a mortality-related stimulus. Shutterstock.com, royalty-free photo.



Army of Jesus
Sponsored
Like Page

Today Americans are able to elect a president with godly moral principles. Hillary is a Satan, and her crimes and lies had proved just how evil she is. And even though Donald Trump isn't a saint by any means, he's at least an honest man and he cares deeply for this country. My vote goes for him!

SATAN: IF I WIN CLINTON WINS!
JESUS: NOT IF I CAN HELP IT!
PRESS 'LIKE' TO HELP JESUS WIN!

Figure 2. An example of a Russian-produced propaganda meme found on Facebook prior to the 2016 U.S. Presidential election (Singer & Brooking, 2018).

Additionally, strategically targeting specific members of a community who are more likely to believe conspiracy theories could increase both the rate and success at which the disinformation is spread. For instance, research suggests that political conservatives and less educated individuals are more likely to believe in conspiracies (Douglas et al., 2015; Fessler et al., 2017). Furthermore, there is research suggesting that individuals who are 65 or older are seven times more likely to share disinformation on social media than younger individuals (Guess et al., 2019). Thus, initially targeting older, less educated, and conservative members of a community with neurocognitive hacking attacks, followed by socially engineered, culturally appropriate disinformation, may increase the probability that the disinformation is ingested and disseminated throughout the community.

In remote tactical operations, in which military personnel need access to adversaries' computer systems to disrupt command and control capabilities, neurocognitive hacking techniques could also be operationally useful. For example, subliminally prompting enemy forces with mortality-related stimuli may increase their propensity to open socially engineered emails containing malicious payloads, especially if the subject lines contain culturally resonant references.

In addition to utilizing subliminal mortality-related imagery, the use of subliminal sounds to induce or compound an effect is also a theoretical possibility. Research suggests that aversive sounds at specific acoustic frequencies can activate the human amygdala (Kumar et al., 2012). Although this research did not investigate the effects of sounds in subliminal frequencies, exploring any possible compounding effects on behavior when combining subliminal visual and auditory prompts would be of interest, as the observed effects from visual prompts tend to last longer when given multiple times (Levy et al., 2014). Examples of effective mortality-related auditory prompts may include hissing sounds from venomous snakes or linguistic threats such as the words "death" and "kill."

Future research

In light of revelations regarding Russian interference into U.S. and European Union elections, the ultimate aim of this line of inquiry is to better understand the extent to which behavior can be manipulated in response to subliminal prompts of mortality-related stimuli (M-rS) and explore means by which the effects can be mitigated (i.e., neurocognitive cybersecurity). These opportunities include investigating the effects of M-rS on hostile attribution, voting behavior, and group polarization. Given prior evidence suggesting M-rS produces effects similar to state anxiety (Gauthier, 2011) along with evidence suggesting state anxiety can be compounded (Pederson & Larson, 2016), exploring the relationship between priming frequency and behavior would be of interest (e.g., would prompting with 20 subliminal M-rS prompts every minute be more effective in eliciting targeted behavioral outputs than one every 10 minutes?).

Also, given that state anxiety is known to disrupt cognitive functioning (Eysenk et al., 2007), exploring for decrements in reaction time, attention, and memory resulting from M-rS would be of interest. If negative effects were found, and given susceptibility to email phishing has been attributed to limitations in cognitive ability (Goel et al., 2017), it follows that neurocognitive attacks could disrupt the speed and efficiency with which adversary military coders are able to efficiently respond to malware attacks during time-critical military operations.

Additionally, given previous research suggesting a correlation between the small ballistic eye movements referred to as micro-saccades and threat salience (Laretzaki et al., 2011), along with the strong association between the human visual system and the amygdala (Burra et al., 2013), further exploration of this connection would be of interest. A strong correlation may allow for the development of neurocognitive cybersecurity protocols for cyber operations personnel (i.e., detection of subliminal attacks using M-rS) as well as a biometric for use in medical diagnostics (e.g., Parkinson's or Alzheimer's disease) or security operations (e.g., detecting individuals with lethal ruminations/intent before boarding airplanes or entering military installations or entertainment venues).

An additional research avenue involves exploring the effects of repeated subliminal activations on the amygdala and ACC. In other words, much like the hippocampus of London taxi drivers has been shown to change after years of memory training (Maguire et al., 2006), it is logical to assume the amygdala and other neural components are similarly malleable. If so, determining whether amygdala growth moderates negative emotional reactivity (e.g., hostile attribution and ego threat) would be of interest.

Similarly, there is much research indicating the existence of neural correlates to common emotional experiences such as empathy, beauty, and romantic love (Lane & Nadel, 2002). Therefore, determining any repeated exposure effects to subliminal images known to activate these areas would be of interest and could extend the scope of neurocognitive hacking. For instance, would pairing subliminal images associated with positive emotional states with images of a cultural out-group help generate positive affect toward them? If so, how many subliminal exposures would be required to manipulate the specific neural structures and produce an effect? Hundreds an hour? Thousands a day? Millions over a period of weeks or months?

Finally, as evidence suggests that "mere exposure" to subliminal images of neutral objects can subsequently increase positive affiliation for them (Zajonc, 2001), extending this line of research to neurocognitive hacking techniques may prove useful in civil-military operations. For example, consider a situation whereby the United Nations deployed a peacekeeping force composed of soldiers ethnicity different from the combatants they were ordered to separate. Theoretically, exposing the combatants to subliminal images of the ethnically different peacekeeping troops before the deployment may decrease the amount of suspicion or resistance the indigenous populations initially generate.

Conclusion

In February 2013, Valery Gerasimov, a retired Russian general, published a short essay on the use of information in modern warfare that came to be known as the "Gerasimov Doctrine" (Duncan, 2018). Among other things, the doctrine highlighted the need for Russia to possess the capability to execute sustained information operations for purposes of creating chaos and unrest against adversaries. Russia is assessed to have used the tactics outlined in the Gerasimov Doctrine during its 2014 Ukrainian operations and prior to its annexation of Crimea (Duncan, 2018). Judging by the information in the U.S. Intelligence Community report on Russian social media hacking and the Office of Special Counsel's February 2018 indictment, Russia has now turned its sights on Western liberal democracies. In fact, the director of the U.S. Federal Bureau of Investigation recently warned of the ongoing interference and "significant counterintelligence threat" posed by Russians actors to the 2020 U.S. presidential election (Barnes & Goldman, 2019).

Information warfare theorists predict these types of information operations will be ubiquitous in the future (Polyakova & Boyer, 2018), so further exploration of cognitive tools like those discussed here should be undertaken to give democratic countries defensive as well as offensive leverage. Supporting the development of mitigating strategies against adversaries who may employ these tactics against democratic elections should be a priority. These mitigating strategies would best be conceptualized in terms of providing *neurocognitive cybersecurity* and prioritized for those with important military or national security-related responsibilities.

References

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, 4(3), 165–178. https://doi.org/10.1038/nrn1056

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, 60, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *Journal of Neuroscience*, 15(9): 5879–5891.

Ajzen, I., & Cote, N. G. (2008). Attitudes and the prediction of behavior. In W. D. Crano & R. Prislin (Eds.), *Frontiers of social psychology*: *Attitudes and attitude change* (pp. 289–311). Psychology Press.

Arndt, J., Greenberg, J., Pyszczynski, T., & Solomon, S. (1997). Subliminal exposure to death-related stimuli increases defense of the cultural worldview. *Psychological Science*, 8(5), 379–385. https://doi.org/10.1111/j.1467-9280.1997.tb00429.x

Astorino-Courtois, A. (2017). *A cognitive capabilities agenda: A multi-step approach for closing DoD's cognitive capability gap*. Strategic Multilayer Assessment Office, U.S. Department of Defense. https://nsiteam.com/a-cognitive-capabilities-agenda-a-multi-step-approach-for-closing-dods-cognitive-capability-gap/

Barnes, J., & Goldman, A. (2019, April 26). FBI warns of Russian interference in 2020 race and boosts counterintelligence operations. *New York Times*. https://www.nytimes.com/2019/04/26/us/politics/fbi-russian-election-interference.html

Bohleber, W., & Bohleber, M. (2012). *Processes of political and terrorist radicalization in late adolescences—Some case examples* [Paper presentation]. Annual Freud Conference, Melbourne, Australia. http://www.freudconference.com/downloads/BohleberLeuzinger-BohleberRadicalizationinadolescence1-2.pdf

Bump, P. (2016, December 1). Donald Trump will be president thanks to 80,000 people in three states. *Washington Post*. https://www.washingtonpost.com/news/the-fix/wp/2016/12/01/donald-trump-will-be-president-thanks-to-80000-people-in-three-states/

Burke, B., Martens, A., & Faucher, E. (2010). Two decades of terror management theory: A meta-analysis of mortality salience research. *Personality and Social Psychology Review*, 14(2), 155–195.

Burra, N., Hervais-Adelman, A., Kerzel, D., Tamietto, M., Gelder, B. D., & Pegna, A. J. (2013). Amygdala activation for eye contact despite complete cortical blindness. *Journal of Neuroscience*, 33(25), 10483–10489. https://doi.org/10.1523/jneurosci.3994-12.2013

Custers, R. (2009). How does our unconscious know what we want? The role of affect in goal representations. In G. B.

Moskowitz & H. Grant (Eds.), *The psychology of goals* (pp. 179–202). Guilford Press.

Cybenko, G., Giani, A., & Thompson, P. (2002). Cognitive hacking: A battle for the mind. *Computer*, 35(8), 50–56. https://doi.org/10.1109/mc.2002.1023788

Devos, T. (2008). Implicit attitudes 101: Theoretical and empirical insights. In W. D. Crano & R. Prislin (Eds.), *Attitudes and persuasion* (pp. 61–94). Psychology Press.

Douglas, K. M., Sutton, R. M., Callan, M. J., Dawtry, R. J., & Harvey, A. J. (2015). Someone is pulling the strings: Hypersensitive agency detection and belief in conspiracy theories. *Thinking & Reasoning*, 22(1), 57–77. https://doi.org/10.1080/13546783.2015.1051586

Duncan, A. J. (2018). New hybrid war or old dirty tricks? The Gerasimov debate and Russia's response to the contemporary operating environment. *Canadian Military Journal*, 17(3), 6–16.

Dwoskin, E., & Gowen, A. (2018, July 23). On WhatsApp, fake news is fast and can be fatal. *Washington Post*. https://www.washingtonpost.com/business/economy/on-whatsapp-fake-news-is-fast–and-can-be-fatal/2018/07/23/a2dd7112-8ebf-11e8-bcd5-9d911c784c38_story.html

European Commission. (2019). *Report on the implementation of the action plan against disinformation*. https://eeas.europa.eu/sites/eeas/files/joint_report_on_disinformation.pdf

Eysenk, M., Derakshan, N., Sanos, R., & Calvo, M. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7(2), 336–353.

Fang, Z., Li, H., Chen, G., & Yang, J. (2016). Unconscious processing of negative animals and objects: Role of the amygdala revealed by fMRI. *Frontiers in Human Neuroscience*, 10. https://doi.org/10.3389/fnhum.2016.00146

Fessler, D. M., & Navarrete, C. D. (2005). The effect of age on death disgust: Challenges to terror management perspectives. *Evolutionary Psychology*, 3(1), 147470490500300. https://doi.org/10.1177/147470490500300120

Fessler, D. M., Pisor, A. C., & Holbrook, C. (2017). Political orientation predicts credulity regarding putative hazards. *Psychological Science*, 28(5), 651–660. https://doi.org/10.1177/0956797617692108

Fitzgerald, F. F. (1972). *Fire in the lake*. Atlantic Monthly Press.

Gauthier, C. (2011). *Are we afraid or anxious about death? Clarifying the meaning of "terror" in terror management theory* [Doctoral dissertation]. New School for Social Research.

Glassner, B. (1985). Review of *Social identity and intergroup relations*, by H. Tajfel. *Contemporary Sociology*, 14(4), 520–521. https://doi.org/10.2307/2069233.

Goel, S., Williams, K., & Dincelli, E. (2017). Got phished? Internet security and human vulnerability. *Journal of the Association for Information Systems*, 18(1), 22–44. https://doi.org/10.17705/1jais.00447

Greenberg, J., Pyszczynski, T., & Solomon, S. (1986). The causes and consequences of a need for self-esteem: A terror management theory. In R. F. Baumeister (Ed.), *Public and private self* (pp. 189–212). Springer-Verlag. https://doi.org/10.1007/978-1-4613-9564-5_10

Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology*, 58 (2), 308–318. https://doi.org/10.1037//0022-3514.58.2.308

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1). https://doi.org/10.1126/sciadv.aau4586

Halpern, S. (2018, March 30). Cambridge Analytica and the perils of psychographics. *The New Yorker*. https://www.newyorker.com/news/news-desk/cambridge-analytica-and-the-perils-of-psychographics

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, 27(5), 733–763. https://doi.org/10.1521/soco.2009.27.5.733

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91. https://doi.org/10.1037//0022-3514.78.1.81

Heslen, J. (2016). *Leading a more effective intelligence community: Understanding and managing the cognitive challenges of human intelligence collection in lethal environments* [Doctoral dissertation]. University of Oklahoma.

Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, 58(4), 255–269. https://doi.org/10.2307/2787127

Holbrook, C., Sousa, P., & Hahn-Holbrook, J. (2011). Unconscious vigilance: Worldview defense without adaptations for terror, coalition, or uncertainty management. *Journal of Personality and Social Psychology*, 101(3), 451–466. https://doi.org/10.1037/a0024033

Jorgenson, L. A., Newsome, W. T., Anderson, D. J., Bargmann, C. I., Brown, E. N., Deisseroth, K., … Wingfield, J. C. (2015).

The brain initiative: Developing technology to catalyse neuroscience discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668), 20140164. https://doi.org/10.1098/rstb.2014.0164

Jost, J., Banaji, M., & Nosek, B. (2016). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881–919. https://doi.org/10.31234/osf.io/6ue35

Jowett, G. S., & O'Donnell, V. (1986). *Propaganda and persuasion*. Sage Publications.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kirkpatrick, L., & Navarrete, C. (2006). Reports of my death anxiety have been greatly exaggerated: A critique of terror management theory from an evolutionary perspective. *Psychological Inquiry*, 17(4), 288–298. https://doi.org/10.1080/10478400701366969

Kumar, S., Kriegstein, K. V., Friston, K., & Griffiths, T. D. (2012). Features versus feelings: Dissociable representations of the acoustic features and valence of aversive sounds. *Journal of Neuroscience*, 32(41), 14184–14192. https://doi.org/10.1523/jneurosci.1759-12.2012

Lane, R. D., & Nadel, L. (2002). *Cognitive neuroscience of emotion*. Oxford University Press.

Laretzaki, G., Plainis, S., Vrettos, I., Chrisoulakis, A., Pallikaris, I., & Bitsios, P. (2011). Threat and trait anxiety affect stability of gaze fixation. *Biological Psychology*, 86(3), 330–336. https://doi.org/10.1016/j.biopsycho.2011.01.005

Larson, E., Darilek, R., Kaye, D., Morgan, F., Nichiporuk, B., Durham-Scott, D., Thurston, C., & Leuschner, K. (2009). *Understanding commanders' information needs for influence operations* (Report No W74V8H-06-C-0001). RAND Corporation. https://www.rand.org/content/dam/rand/pubs/monographs/2009/RAND_MG656.pdf

Levy, B. R., Pilver, C., Chung, P. H., & Slade, M. D. (2014). Subliminal strengthening. *Psychological Science*, 25(12), 2127–2135. https://doi.org/10.1177/0956797614551970

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, 16(12), 1091–1101. https://doi.org/10.1002/hipo.20233

Markova, I. (2008). Persuasion and propaganda. *Diogenes*, 55 (1), 37–51. https://doi.org/10.1177/0392192107087916

Meyer, R. (2014, June 28). Everything we know about Facebook's secret mood manipulation experiment. *The Atlantic*. https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

Nagl, J. A. (2002). *Learning to eat soup with a knife: Counterinsurgency lessons from Malaya and Vietnam*. University of Chicago Press.

Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: Enhanced sensitivity to negative word valence. *Emotion*, 9(5), 609–618. https://doi.org/10.1037/a0016305

National Public Radio. (2018, March 27). *They don't care: Whistleblower says Cambridge Analytica aims to undermine democracy*. https://www.npr.org/sections/thetwo-way/2018/03/27/597279596/they-don-t-care-whistleblower-says-cambridge-analytica-seeks-to-undermine-democr

Nyabola, N. (2019, February 15). *The spectre of Cambridge Analytica still haunts African elections*. Al Jazeera. https://www.aljazeera.com/indepth/opinion/nigerian-elections-money-190215080009476.html

Office of the Director of National Intelligence. (2017, January 6). *Background to 'Assessing Russian activities and intentions in recent US elections': The analytic process and cyber incident attribution*. https://www.dni.gov/files/documents/ICA_2017_01.pdf

Pearson, E. (2015). The Case of Roshonara Choudhry: Implications for theory on online radicalization, ISIS women, and the gendered jihad. *Policy & Internet*, 8(1), 5–33. https://doi.org/10.1002/poi3.101

Pedersen, W. S., & Larson, C. L. (2016). State anxiety carried over from prior threat increases late positive potential amplitude during an instructed emotion regulation task. *Emotion*, 16(5), 719–729. https://doi.org/10.1037/emo0000154

Phelps, E. A., Oconnor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12 (5), 729–738. https://doi.org/10.1162/089892900562552

Polyakova, A., & Boyer, S. P. (2018). *The Future of political warfare: Russia, the West, and the coming age of global digital competition*. Brookings Institution. https://www.brookings.edu/wp-content/uploads/2018/03/the-future-of-political-warfare.pdf

Quirin, M., Loktyushin, A., Arndt, J., Küstermann, E., Lo, Y.-Y., Kuhl, J., & Eggert, L. (2011). Existential neuroscience: A functional magnetic resonance imaging investigation of neural responses to reminders of one's mortality. *Social Cognitive and Affective Neuroscience*, 7(2), 193–198. https://doi.org/10.1093/scan/nsq106

Rosenblatt, A., Greenberg J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory I: The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57(4), 681–690. https://doi.org/10.1037//0022-3514.57.4.681

Singer, P.W., Emerson, B.T. (2018). *Like War: The Weaponization of Social Media*. Houghton-Mifflin Harcourt Publishing.

Solomon, S., Greenberg, J., & Pyszczynski, T. (2004). The cultural animal: Twenty years of terror management theory and research. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 13–34). Guilford Press.

Stanovich, K., & West, R. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23(5), 645–726. https://doi.org/10.1017/s0140525x00003435

Taylor, P. M. (2003). *Munitions of the mind: A history of propaganda from the ancient world to the present day*. Manchester University Press.

Thomas, T. (2004). Russia's reflexive control theory and the military. *Journal of Slavic Military Studies*, 17(2), 237–256. https://doi.org/10.1080/13518040490450529

Tritt, S. M., Inzlicht, M., & Harmon-Jones, E. (2012). Toward a biological understanding of mortality salience (and other threat compensation processes). *Social Cognition*, 30(6), 715–733. https://doi.org/10.1521/soco.2012.30.6.715

U.S. Department of the Army. (2003). *Psychological operations tactics*, *techniques*, *and procedures* (FM 33-1-1). https://fas.org/irp/doddir/army/fm3-05-301.pdf

U.S. Department of Justice. (2018). *United States of America v. Internet Research Agency LLC*. https://www.justice.gov/file/1035477/download

U.S. Department of Defense (2014). JP-3-13, Information Operations (JP 3-13). https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_13.pdf

Valeriano, B., & Jensen, B. (2019). *The myth of the cyber defense: The case for restraint* (Policy Analysis No. 862). CATO Institute. https://www.cato.org/publications/policy-analysis/myth-cyber-offense-case-restraint

Votel, J., Clevland, C., Connett, C., & Irwin, W. (2016). Unconventional warfare in the gray zone. *Joint Forces Quarterly*, 80, 101–109.

Weimann, G. (2015). Terrorist migration to social media. *Georgetown Journal of International Affairs*, 16(1), 180–187.

Weingarten, E., Chen, Q., Mcadams, M., Yi, J., Hepler, J., & Albarracin, D. (2016). On priming action: Conclusions from a meta-analysis of the behavioral effects of incidentally-presented words. *Current Opinion in Psychology*, 12, 53–57. https://doi.org/10.1016/j.copsyc.2016.04.015

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7(6), 177–188. https://doi.org/10.1111/1467-8721.ep10836912

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice. *Psychological Science*, 16(1), 56–63. https://doi.org/10.1111/j.0956-7976.2005.00780.x.

Windsor, L. (2018). The language of radicalization: Female internet recruitment to participation in ISIS activities. *Terrorism and Political Violence. Advance online publication*. https://doi.org/10.1080/09546553.2017.1385457

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychology*, 13(3), 120–123.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31(1), 121–135. https://doi.org/10.1177/0146167204271309

Zajonc, R. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10(6), 224–228. https://doi.org/10.1111/1467-8721.00154

AI-based MultiModal to Identify State-linked Social Media Accounts in the Middle East: A Study on Twitter

Abdullah Melhem
School of Computer and Cyber Sciences
Augusta University
Augusta, USA
abanimelhem@augusta.edu

Ahmed Aleroud
School of Computer and Cyber Sciences
Augusta University
Augusta, USA
aaleroud@augusta.edu

Zain Halloush
School of Computer and Cyber Sciences
Augusta University
Augusta, USA
zhaloush@augusta.edu

2023 IEEE International Conference on Intelligence and Security Informatics (ISI) | 979-8-3503-3773-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISI58743.2023.10297235

***Abstract*—State-linked propaganda on Social Media poses a new challenge at the geopolitical level for the United States and other countries. The widespread of social media platforms makes it easier for adversaries to spread disinformation, conspiracy theories and social-cyber attacks at a scale that was not possible without such networks. Not only English content on social media represents such a challenge since some of those cyber-mediated attacks are initiated by agents who target other languages. In this paper, we proposed a Multimodal AI approach to detect state-linked accounts on twitter. As opposed to previous efforts, we focus our research on the Middle East and the Anti-USA content on Twitter. We trained AI Multimodal approaches on data with categorical, textual and numerical features. The study utilized experimental Twitter data connected to numerous suspected state-linked accounts on the platform. Additionally, we collected data to represent the negative samples. The findings indicate that the significance of textual modalities and AI language models in identifying state-linked accounts was limited. Our study demonstrated the crucial significance of account metadata and other modalities to detect state-linked propaganda and the associated accounts effectively.**

***Index Terms*—Multimodal, sentiment analysis, AI, language models, state-linked, disinformation, propaganda.**

I. Introduction

The rapid growth of social media platforms makes them powerful tools for large scale political campaigns. Before social media, governments were using the term "state-linked attacks" to refer to cyber-attacks such as spam, DDoS, phishing, theft of confidential information, click fraud, cyber-sabotage, and cyber-warfare. However, social media discourse leads to other socio-political attack vectors that affect people's opinions and minds, which represent a new significant national security challenge [1]. The majority of existing research that examines state-linked Anti-USA content on social media primarily concentrates on Russia's endeavors to manipulate public opinions, particularly targeting English-speaking communities [2]. However, state-linked activities that target the United States also exists in other low resource languages such as Arabic [3]. Twitter has recently published state-backed information operations datasets. The published data consists of manipulation campaigns originating from 17 countries, which includes Several countries in the Middle East [4].

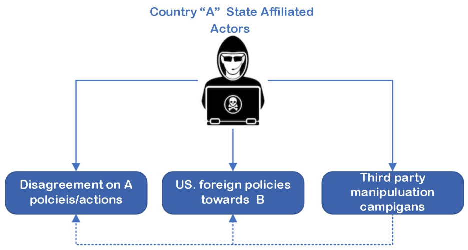

Country "A" State Affiliated Actors
Disagreement on A polcieis/actions
US. foreign policies towards B
Third party manipuluation campigans

Fig. 1. Motivations of state-affiliated Social Media accounts in the Middle East

Detection of state-linked or (state-affiliated) social media accounts that exist in low resource languages, such as Arabic, is a difficult task due to several reasons: First, most of the AI language models were developed to work on high resource languages such as English. Second, the cultural-geopolitical contexts and language dialect makes it difficult to create AI models to identify such accounts and the sentiments of tweets associated with them. Furthermore, most of the existing research on this area is solely computational, which largely ignores contextual information on social media when it contains sensitive political content about the key leaders in the Middle East. Detecting state-linked activities is a complex task that extends beyond the realm of sentiment analysis.

From our initial analysis of Twitter data (2018-2021), we observed three motivations behind state-linked accounts targeting the United States (see Fig. 1). Specifically, if A is a state that sponsors the affiliated account targeting the US, then there three possible motivations:

1) Disagreement between the USA and states' A internal policies or actions, such as democracy or human rights.
2) USA foreign policies and relations with country B, when A and B are in conflict.
3) Third Party campaigns. For example, Russia has reportedly utilized Middle Eastern social media to influence

This material is based upon research supported by the U.S. Office of Naval Research under award number N000142212549

the discourse surrounding Syrian and Yemeni conflicts [3].

Other challenges in this area of research are computational. Given the small amount of available social media training data, such as tweets, relying solely on language models to create Machine Learning(ML) classifiers to identify state-linked tweets and the associated accounts is not straightforward. Even when using regular tabular features such as the account details, or numerical features such as the likes and retweet counts, without appropriately combining these features and use them through an integrated pipeline, it is difficult to predict the behavior of the ML models. For instance, just by encoding both textual tweet features and account features, we are only creating a single feature vector that can be fed into a ML/DL model. However, this approach assumes that categorical text, and numerical features provide complementary information and that encoding them separately and concatenating them is sufficient to capture this information.

Multimodal AI applies a neural network, specifically an MLP, to process categorical and numerical features into learned features [5]. These, when combined with the embedded textual features, form fully connected layers, enabling state-linked content prediction across various dimensions with appropriate modality weightings. This research argues that Multimodal AI models can be used in low resource languages such as Arabic to identify targeted social media state-linked activities and the accounts associated with them. Our method is summarized as follows: we first created a dataset of positive and negative tweets that belong to state-linked and non-statelinked accounts. We then utilized several multimodal AI approaches to classify tweets that correspond to those accounts as originating from state-linked accounts/non-state-linked. We argue that while some tweets may spread text-based or sentiment-based propaganda against the US, they are not necessarily state-linked, which highlights the importance of our Multimodal analysis of the data. To our knowledge, this study represents one of the early research endeavors focused on examining the detection of state-linked activities in the Middle East. We are not investigating trolls or social media bots, but our approach aims at detecting state-linked activities that manipulate public opinions on discourse where the United States is involved.

The rest of the paper is organized as follows: Section II reviews the important related works in the field. Section III gives a glance over the multimodal approach adopted in this research and the dataset creation steps. Section IV present the results of our experiments. Section V concludes the work.

II. Related Work

Research has studied the phenomena of spreading fake news and rumors on social media platforms using ML algorithms. Such research efforts have encompassed various areas, including content analysis [6], studying the characteristics and features of fake news propagation [7], analyzing the features of users' profiles [8], and studying the enormous effect of social media bots to amplify the impact of fake news expansion [9], [10]. Other studies focused on the correlation between social media bots and accounts' fabrication and social spamming [11]. Most of such studies concluded that using Deep Learning models overcomes the problem of manual feature extraction [12]. The study conducted by Sharma et al. in [13] proposed an Attentive Mixture Density Network approach to examine similar activities of malicious accounts that spread fake information. The method captures the latent characteristics of group influence between those accounts based on Temporal Point Processes and Gaussian Mixture Model. A similar study conducted by the authors in [14] presented an unsupervised methodology for detecting synchronized activities among accounts based on behavioral traces and linking suspicious handles between them. The study includes tweets about the U.S. election, Hong Kong protests, Syria civil war, and bitcoin manipulation. Additionally, in [15], the authors examined Twitter and Reddit users identified as Russian and Iranian state-sponsored trolls. The analysis involved various features, such as the number of followers/friends, overall profile persona, and temporal patterns during specific events. Similarly, in [16], the authors analyzed behavioral and linguistic features to identify Russian Trolls using SAGE analysis. The found a high likelihood of these Russian Trolls remaining active and conducting targeted campaigns even after the U.S. election. A study on the activities of the Russia's Internet Research Agency and the state-linked accounts was conducted by [17]. The study shows the significant temporal change in the forcefulness of these accounts and the language variation (Russian and English) based on the change of events. The visual aspects of social media interactions have become a significant concern for the research community, leading to various multimodal frameworks [18]. In [19], a multi-modal variational autoencoder was introduced to detect fake news both textually and visually. The experiments were conducted on datasets from Weibo and Twitter platforms, where the authors combined a binary classifier for text data with a bimodal variational encoder.

Additionally, [20] presented a hybrid CNN-RNN model for detecting fake news and misinformation in textual data. Another multi-modal framework, KMAGCN, was introduced in [21], combining textual information with knowledge concepts and visual data to detect fake news. The authors designed an adaptive graph convolutional network to capture text data dependencies, incorporating an attention mechanism to match visual and textual semantics for more accurate fake news detection.

In the current approaches, there is often a heavy reliance on a single modality, which is typically limited to text-based cues. However, it is crucial to consider dynamics of the content shared as well to enhance the effectiveness of state-linked account detection and overcome inherent limitations. Moreover, the current approach tends to overlook state-linked activities in regions with a multitude of geopolitical events, such as the Middle East. By adopting a more comprehensive multi-modal perspective that takes into account the dynamics of the content shared, we can better capture and analyze state-linked activities in these high-stakes regions.

III. Methodology

The goal of the utilized AI multimodal approach is to leverage the recent trend of using transformers in prediction tasks that not only involve textual features but other types of features. The approach is depicted in figure 2. Our approach is based on the methodology proposed by [22]. Tweets labeled as belonging to state-linked accounts or non-state-linked accounts are the input to the proposed method. The collected data consists of numerical, categorical and textual features. Data preprocessing consists of stop word removal, removing special characters, encoding of categorical attributes and tokenization.

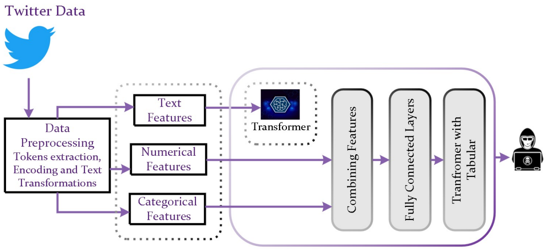


Fig. 2. Research Approach

The existing Hugging Face Transformers were used to process textual features. The features combination module allows a feature-rich training pipeline. Transformers incorporate numerous pre-trained models [23]. There are variety of Multi-lingual models, which also allows the reuse of those transformers with tabular features. The feature combination takes as input $\boldsymbol{x}$, a text features generated by a Transformer model and the preprocessed tabular (categorical ($\boldsymbol{c}$) and numerical ($\boldsymbol{n}$)) features, and outputs a combined multimodal representation m. The parameters of the combining process are aggregated and trained using fully connected layers to predict if the tweet belongs to a state-linked account. The following sections describe each of the feature combination methods utilized in this research, where the uppercase bold letters represent 2D matrices, lowercase bold letters represent 1D vectors. b is a scalar bias, W represents a weight matrix, and $\|$ is the is the concatenation operator.

A. Feature combination

The feature combination methods used to create multimodalities are:

1) Text only: The model will only take tweet text as the only input without using any other modalities. Formally, it can be expressed as follows:

$$\boldsymbol{m} = \boldsymbol{x} \tag{1}$$

2) Concat: In this multimodal, the feature vectors for textual, categorical and numerical features are encoded then combined together using the traditional concatenation process as follows:

$$m = x \| c \| n \tag{2}$$

3) Individual MLPs on categorical and numerical features then concat (MLP + concat): In this multimodal, individual MLPs (multi-layer perceptrons) are used to process categorical and numerical features, before concatenating them with the textual features into a single input representation as follows:

$$\boldsymbol{m} = \boldsymbol{x} \| \operatorname{MLP}(\boldsymbol{c}) \| \operatorname{MLP}(\boldsymbol{n}) \tag{3}$$

4) MLP on concatenated categorical and numerical features then concat (Concat+MLP): In this model, MLP is applied on the concatenated categorical and numerical features. The output is then concatenated with the output of the transformer that processes textual features. Formally, this can be expressed as:

$$\boldsymbol{m} = \boldsymbol{x} \| \operatorname{MLP}(\boldsymbol{c} \| \boldsymbol{n}) \tag{4}$$

5) Attention on categorical and numerical features: In this model categorical and numerical features are combined into a single mode and processed separately by individual MLPs (multilayer perceptrons) to extract relevant information. The MLP outputs are then concatenated, and an attention process is applied to weigh the contribution of each feature to the final output as follows:

$$\boldsymbol{m} = \alpha_{x,x} \boldsymbol{W}_x \boldsymbol{x} + \alpha_{x,c} \boldsymbol{W}_c \boldsymbol{c} + \alpha_{x,n} \boldsymbol{W}_n \boldsymbol{n} \tag{5}$$

$$\alpha_{i,j} = \frac{\exp\left(\operatorname{LeakyReLU}\left(\boldsymbol{a}^T \left[\boldsymbol{W}_i \boldsymbol{x}_i \| \boldsymbol{W}_j \boldsymbol{x}_j\right]\right)\right)}{\sum_{k \in \{x,c,n\}} \exp\left(\operatorname{LeakyReLU}\left(\boldsymbol{a}^T \left[\boldsymbol{W}_i \boldsymbol{x}_i \| \boldsymbol{W}_k \boldsymbol{x}_k\right]\right)\right)} \tag{6}$$

6) Gating on categorical and numerical features and then sum: in this approach, a gating function is applied to the representations of each modality, typically using Sigmoid or Softmax activation functions [24]. This function generates gating values that indicate the relevance or importance of each modality information to be nominated for the final fusion as follows:

$$\boldsymbol{m} = \boldsymbol{x} + \alpha \boldsymbol{h} \tag{7}$$

$$\boldsymbol{h} = \boldsymbol{g}_c \odot (\boldsymbol{W}_c \boldsymbol{c}) + \boldsymbol{g}_n \odot (\boldsymbol{W}_n \boldsymbol{n}) + b_h \tag{8}$$

$$\alpha = \min\left(\frac{\|\boldsymbol{x}\|_2}{\|\boldsymbol{h}\|_2} * \beta, 1\right) \tag{9}$$

$$\boldsymbol{g}_i = \mathrm{R}\left(\boldsymbol{W}_{g_i} \boldsymbol{i} \| \boldsymbol{x} + b_i\right) \tag{10}$$

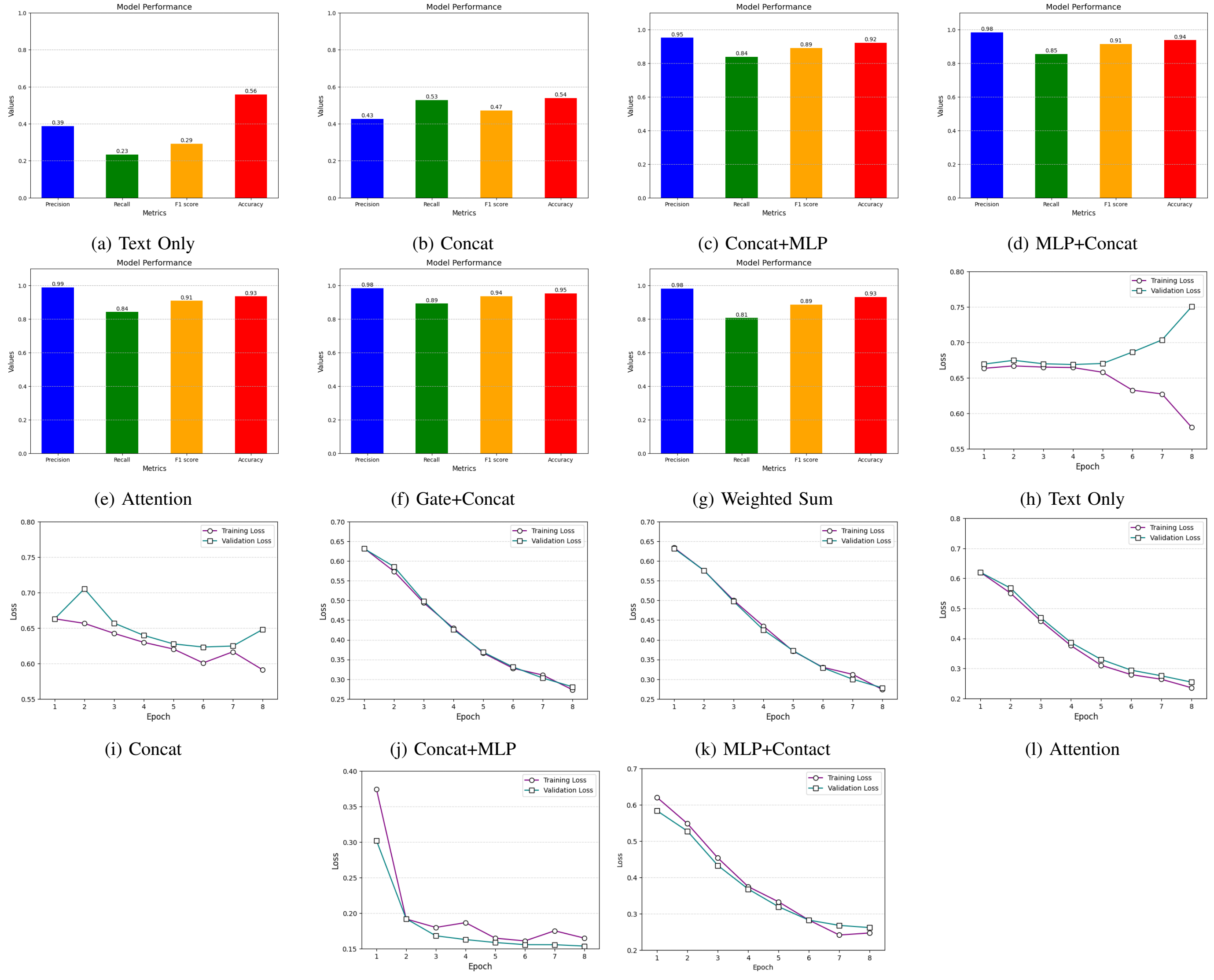
where β is a hyperparameter and R is the activation function.

7) Weighted feature sum for text, categorical, and numerical features (i.e. Weighted Sum): This approach integrates multiple modalities by assigning separate weights to each modality as follows:

$$\boldsymbol{m} = \boldsymbol{x} + \boldsymbol{w}_c \odot \boldsymbol{W}_c \boldsymbol{c} + \boldsymbol{w}_n \odot \boldsymbol{W}_n \boldsymbol{n} \tag{11}$$

IV. Experiments

This section discusses the results of our experiments. First we discuss the how our data is collected, then we present the results of different experiments.



(a) Text Only
(b) Concat
(c) Concat+MLP
(d) MLP+Concat
(e) Attention
(f) Gate+Concat
(g) Weighted Sum
(h) Text Only
(i) Concat
(j) Concat+MLP
(k) MLP+Contact
(l) Attention
(m) Gate+Contact
(n) Weighted Sum

Fig. 3. Results of different modalities

TABLE I. STATE-LINKED DATASET

Time-frame	**Country**	**No. of accounts**
December 2019	Saudi Arabia	5929
September 2020	Saudi Arabia	34
September 2019	UAE& Egypt	4248+ 271

A. Data Collection

Between 2018-2022, Twitter published a comprehensive public archive of data related to state-backed information operations[1]. 37 data sets have been shared as part of this effort [4]. The published datasets consist of manipulation campaigns originating from 17 countries, spanning more than 200 million Tweets.

Since we were interested in targeted state-linked attacks, some of the tweet activities were related to the US and are mainly Anti-US. We selected accounts that are suspected to be related to some countries in the Middle East. It is important to acknowledge that the labeling of state-linked activities is subject to certain inherent limitations and potential biases. Recently, Elon Musk labeled accounts of NPR, etc. as ”state-linked” and led to some big pushbacks. Musk agreed to change the label to “government-funded media” [25]. These incidents highlight the challenges and complexities in accurately categorizing accounts and the potential for misinterpretations. We curated topics that were most likely indicative of government activities. The selected accounts were mainly tweeting propagandic themes on topics related to Muslims brotherhood in the Middle East, Iranians’ role in the Middle East, War in Syria. and Sanctions on Qatar

State-linked accounts were selected from 3 sub-datasets. The number of state-linked accounts per each subset is shown

[1] The code is available at the following link: https://github.com/ahmed-aleroud/multimodal-statelinked

TABLE II. FEATURES

Non-state linked features	State linked features
user_id	userid
username	user_display_name
name	user_screen_name
user_profile_url	user_reported_location
geolocation	user_profile_description
language	user_profile_url
tweet_language	follower_count
tweet_text	following_count
tweet_time	account_creation_date
retweet_count	account_language
like_count	tweet_language
reply_count	tweet_text
urls	tweet_time
hashtags	tweet_client_name
mentions	in_reply_to_userid
is_retweet	in_reply_to_tweetid
retweet_tweetid	quoted_tweet_tweetid
retweet_userid	is_retweet
	retweet_userid
	retweet_tweetid
	quote_count
	reply_count
	like_count
	retweet_count
	hashtags
	urls
	user_mentions

TABLE III. DESCRIPTION OF THE DATASET

	Non state-linked	State-linked	Total
Training	6075 (63%)	3608 (37%)	9683
Validation	774 (64%)	436 (36%)	1210
Testing	758 (63%)	453 (36%)	1211

TABLE IV. PERFORMANCE COMPARISON OF DIFFERENT MODELS ON TESTING DATA

Model name	P	R	F	RT	TSPS	ESPS
Gate+Concat	0.992	0.878	0.932	**3.26**	**371**	**46.626**
Attention	0.995	**0.897**	**0.944**	3.47	349	43.792
Individual mlps	0.576	0.490	0.530	3.54	342	42.882
MLP+concat	**0.980**	0.895	0.936	3.48	348	43.635
Weighted FS	0.767	0.647	0.702	3.58	338	42.457
Concat	0.386	0.165	0.231	3.45	351	44.065
Text Only	0.632	0.027	0.052	3.48	348	43.744

in table I. Accounts were suspected to be linked to Saudi, UAE and Egypt governments. In addition to those positive samples, negative non-state linked tweets were collected using a list of political and non-political keywords. One possible approach to collect those tweets was to make a basic assumption that involves gathering tweets from specific user accounts that publish non-political content and labeled them as non-state linked. However, this simple assumption does not help to create robust AI models that can identify differences between accounts that publish political content that is not state-linked. Indeed, both state and non-state linked accounts publish political and non political content. We followed a more rigorous approach to deem the collected tweets as belonging to non-state linked accounts. If the tweets primarily originate from accounts that mainly tweet non-political topics or on personal interests, the tweets are labeled as originated from a non-state linked accounts. Tweets that heavily publish on the topics related to our positive samples were discarded as those may also indicate state-linked accounts. Finally, not all accounts that publish content against specific countries such as the USA are automatically classified as state linked. There are many reasons why an individual or organization may post content that is against the USA, such as political beliefs, personal grievances, or advocacy for a particular cause. Therefore, when no specific government policies are being discussed in those accounts, the tweets and the associated accounts are classified as non-state linked. The features of state and non-statelinked accounts are shown in table II. Example of numerical features are, Follower count, Following count, Retweet count, Like count and Reply count. Examples of categorical features are Tweet language, Account language and Is retweet. Example of textual features are User profile description, Tweet text, Hashtags and Mentions. We created a joint list of features that are common between state and non-state linked accounts.

B. Results

We tested the different feature combination methods discussed earlier. We used the language model $Davlan/bert-base-multilingual-cased-ner-hrl$ [26] as a transformer model. We used the $AutoTokenizer$ from the transformers module for tokenization. Our experiments ran using 8 epochs, a learning rate of $5e-5$, eval-steps of 2000 and a weight-decay of 0.01. We used regular expressions to clean the ”tweet-text” column. The dataset is split into training, validation, and testing sets as shown in table III. All experiments run on a NVIDIA A100 GPU with 80 GBs of VRAM. The results of different combination approaches are shown in Figure 3. Text only modalities show that the values of training loss and validation loss are consistent, which indicates that the model is not overfitting at the small values of epochs, as the validation loss is relatively close to the training loss (Figure 3h). However, the values of different metrics suggests that the model performs no better than random guessing for Recall and F1 scores (Figure 3a). The values of Accuracy suggest a moderate discriminative ability. Upon concatenation, the model's ability to distinguish between state and non state-linked attacks has significantly improved (Figures 3b and 3i). It is noticed that the Recall values have significant increase compared to text only combination method, suggesting that the model captures a higher proportion of actual state-linked attacks. The results of the MLP model trained on concatenated categorical and numerical features before concatenating them with textual features show that the training loss and validation loss decrease consistently over the epochs, indicating that the model is learning and generalizing well to unseen data (Figure 3j). The model demonstrates good ability to identify state-linked accounts (Figure 3c). The result of training individual MLP models on numerical and categorical attributes and combining their outputs with the textual features (MLP+Contact), lead to a slightly better results compared to applying MLP on the concatenated features (Figure 3d). The results of the attention mechanism has been also effective in identifying state and non-state linked accounts (Figure 3e). The results were close to those achieved by MLP+Concat. Gating on the categorical and numerical features then concact yields significant Recall values (Figure 3f). Since this type of modalities generates gating values that indicate the relevance or importance of each modality's information for the final fusion process, it can captures the optimal weights of each modality, yielding comparable classification results compared to the attention approach. We noticed that the weighted Feature Sum approach leads to relatively lower Recall values compared to other combinations of modalities (Figure 3g). This can be attributed to how the Weighted Feature Sum approach treats modalities by assigning weights. Sometimes incorrect assignment of those weights can lead to diluting the impact of important modalities or amplifying noise from less informative ones. The last experiment tested the capability of different modalities to classify unseen data. For this experiment, we used a testing dataset that consists of 1211 tweets that belong to state and non state-linked accounts. We then applied the 7 fine tuned models on the testing data. Results are shown in table IV. Results on the testing data show significant advantages for the attention combination approach. MLP+Concat also shows a significant advantage in terms of Precision. In addition, Gate+Concat approach was very efficient in terms of runtime, sample processed per second, and the shorter evaluation steps. The results also highlight the limited prediction of the text modality and the simple concatenation approaches as observed in the last two rows in the table.

V. Conclusions

This research investigates the capabilities of AI multimodal approaches to identify state-linked accounts on social media platforms. We limit our study to twitter and examined the activities of such accounts in the Middle East. We showed that relying on tweets' content only, or tweets' text is not necessarily enough to identify state-linked activities. We also found the combining numerical and categorical modalities yields better model performance in terms of identifying those accounts. We think that this study can be extended to consider other language models. One extension to this work is to conduct an AI explainability study on each combination approach and examine the role of different features in the classification process.

References

[1] P. N. Howard, S. Woolley, and R. Calo, "Algorithms, bots, and political communication in the us 2016 election: The challenge of automated political communication for election law and administration," *Journal of information technology & politics*, vol. 15, no. 2, pp. 81–93, 2018.

[2] M. H. Saeed, S. Ali, J. Blackburn, E. De Cristofaro, S. Zannettou, and G. Stringhini, "Trollmagnifier: Detecting state-sponsored troll accounts on reddit," in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 2161–2175.

[3] N. Oweidat. (2023) The russian propaganda in arabic hidden from the west. [Online]. Available: https://www.washingtoninstitute.org/policy-analysis/russian-propaganda-arabic-hidden-west

[4] Twitter moderation research consortium. [Online]. Available: https://transparency.twitter.com/en/reports/moderation-research.html

[5] L.-P. Morency, P. P. Liang, and A. Zadeh, "Tutorial on multimodal machine learning," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Tutorial Abstracts*, 2022, pp. 33–38.

[6] C. Boididou, S. E. Middleton, Z. Jin, S. Papadopoulos, D.-T. Dang-Nguyen, G. Boato, and Y. Kompatsiaris, "Verifying information with multimedia content on twitter: a comparative study of automated approaches," *Multimedia tools and applications*, vol. 77, pp. 15 545–15 571, 2018.

[7] A. Bodaghi and J. Oliveira, "The characteristics of rumor spreaders on twitter: A quantitative analysis on real data," *Computer Communications*, vol. 160, pp. 674–687, 2020.

[8] K. Shu, S. Wang, and H. Liu, "Understanding user profiles on social media for fake news detection," in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2018, pp. 430–435.

[9] C. Cai, L. Li, and D. Zeng, "Detecting social bots by jointly modeling deep behavior and content information," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 1995–1998.

[10] C. Cai, L. Li, and D. Zengi, "Behavior enhanced deep bot detection in social media," in *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 2017, pp. 128–130.

[11] A. Derhab, R. Alawwad, K. Dehwah, N. Tariq, F. A. Khan, and J. Al-Muhtadi, "Tweet-based bot detection using big data analytics," *IEEE Access*, vol. 9, pp. 65 988–66 005, 2021.

[12] S. Rao, A. K. Verma, and T. Bhatia, "A review on social spam detection: Challenges, open issues, and future directions," *Expert Systems with Applications*, vol. 186, p. 115742, 2021.

[13] K. Sharma, Y. Zhang, E. Ferrara, and Y. Liu, "Identifying coordinated accounts on social media through hidden influence and group behaviours," 2021.

[14] D. Pacheco, P.-M. Hui, C. Torres-Lugo, B. Truong, A. Flammini, and F. Menczer, "Uncovering coordinated networks on social media," 01 2020.

[15] S. Zannettou, T. Caulfield, W. Setzer, M. Sirivianos, G. Stringhini, and J. Blackburn, "Who let the trolls out? towards understanding state-sponsored trolls," in *Proceedings of the 10th acm conference on web science*, 2019, pp. 353–362.

[16] J. Im, E. Chandrasekharan, J. Sargent, P. Lighthammer, T. Denby, A. Bhargava, L. Hemphill, D. Jurgens, and E. Gilbert, "Still out there: Modeling and identifying russian troll accounts on twitter," in *12th ACM conference on web Science*, 2020, pp. 1–10.

[17] P. N. Howard, B. Ganesh, D. Liotsiou, J. Kelly, and C. François, "The ira and political polarization in the united states," *Project on Computational Propaganda*, 2018.

[18] S. Hangloo and B. Arora, "Combating multimodal fake news on social media: methods, datasets, and future perspective," *Multimedia Systems*, vol. 28, no. 6, pp. 2391–2422, 2022.

[19] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," 10 2017, pp. 127–138.

[20] J. Nasir, O. Khan, and I. Varlamis, "Fake news detection: A hybrid cnn-rnn based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, p. 100007, 04 2021.

[21] S. Qian, J. Hu, Q. Fang, and C. Xu, "Knowledge-aware multi-modal adaptive graph convolutional networks for fake news detection," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 17, no. 3, jul 2021. [Online]. Available: https://doi.org/10.1145/3451215

[22] K. Gu and A. Budhkar, "A package for learning on tabular and text data with transformers," in *Proceedings of the Third Workshop on Multimodal Artificial Intelligence*, 2021, pp. 69–73.

[23] C. Delangue and J. Chaumond. (2023) Huggingface. [Online]. Available: https://huggingface.co/

[24] W. Rahman, M. K. Hasan, S. Lee, A. Zadeh, C. Mao, L.-P. Morency, and E. Hoque, "Integrating multimodal information in large pretrained transformers," in *Proceedings of the conference. Association for Computational Linguistics. Meeting*, vol. 2020. NIH Public Access, 2020, p. 2359.

[25] M. Masnick. (2023) Npr says enough is enough: Quits twitter. [Online]. Available: https://www.techdirt.com/2023/04/12/npr-says-enough-is-enough-quits-twitter/

[26] D. Adelani. (2023) Huggingface. [Online]. Available: https://huggingface.co/Davlan/bert-base-multilingual-cased-ner-hrl

Article

SM+S
social media + society

Social Media + Society
October-December 2023: 1–22
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051231216964
journals.sagepub.com/home/sms
Sage

Twitter Propaganda Operations: Analyzing Sociopolitical Issues in Saudi Arabia

Craig Douglas Albert, Ahmed Aleroud, Yufan Yang, Abdullah Melhem, and Josh Rutland

Abstract

The purpose of this article is to explore Arabic-language Tweets based out of Saudi Arabia to investigate the social media landscape. Specifically, this article seeks to address the question, "What thematic issues concerning the U.S. sociopolitical landscape are present in Arabic-language Twitter postings?" And, "To what extent can these issues be described as propagandic in nature?" To do so, we propose a machine-learning and artificial intelligence span detection approach to identify propaganda Tweets in Middle Eastern Countries, with a focus on Saudi Arabia. As opposed to previous work, this article maps and investigates different propaganda categories using the BEND Social Cyber Security framework. This article then proceeds to a case study analysis of state-sponsored targeted propaganda from Saudi Arabia and briefly describes the categories of propaganda uncovered. We then relate those categories to the BEND Framework and conclude with policy recommendations and discussion.

Keywords

machine learning, artificial intelligence, social media analysis, BEND framework, the Middle East, Saudi Arabia

Introduction

Social media platforms have become a battlefield for propagandic campaigns both within and across state borders. Research has shown that both party elites and foreign rival countries have been using social media to spread misinformation, disinformation, and propaganda to serve their goals, such as stigmatizing political opponents and interfering with a rival country's domestic politics (Badawy et al., 2018; Benkler et al., 2018). However, studies on propaganda campaigns between rival states largely focus on the United States–Russia and the United States–China, with only a few researchers exploring Middle Eastern countries' propaganda efforts despite the fact that anti-American sentiment widely exists in Middle Eastern countries (Jamal et al., 2015). Nevertheless, anti-American sentiment plays a critical role in great-power politics such as U.S.–Russia (Mendelson & Gerber, 2008) and U.S.–China relations (Weiss, 2013) and thereafter global peace, as well as in transnational terrorist attacks targeted at the United States (Neumayer & Plümper, 2011). Although there is a long way between increasing hostility toward a foreign country to an open war between nation states, research has shown that increasing hostility among rivalries is associated with more terrorist attacks from one to another (Conrad, 2011). Furthermore, various U.S. documents suggest that foreign propaganda threatens the state by undermining national security objectives of the United States and its allies, jeopardizing trust in democracy, fueling political unrest and violence, and destabilizing society (Chernobrov & Briant, 2022). Although there has been a long-standing tension between some Middle Eastern countries and the United States, Arabic influence operation and propaganda strategies targeted at the United States are understudied.

Exploring how Middle Eastern countries frame the United States to their citizens and what propagandic strategies are implemented in the framing process will deepen our understanding of elite behavior and the public sphere regarding the United States in Middle Eastern countries, which is

Augusta University, USA

Corresponding Author:
Craig Douglas Albert, Master of Arts in Intelligence and Security Studies, Augusta University, 2500 Walton Way, Augusta, GA 30904, USA.
Email: calbert@augusta.edu
X: @drcraigdalbert

beneficial to American national security. Our research attempts to provide more evidence on Middle Eastern countries' propaganda efforts and influence operations regarding the United States on social media. Specifically, we ask, "What thematic issues concerning the U.S. socio-political landscape are present in Arabic-language Twitter postings?" And, "To what extent can these issues be described as propagandic and state-sponsored influence operations in nature?" To do so, we analyze state-sponsored Twitter data from Middle Eastern countries, identifying the propaganda techniques with the help of machine learning algorithms, and explore the content of these Tweets using topic modeling. We also add a case study of Saudi Arabia to illustrate our findings. Before delving into the review of literature, it is necessary to conceptualize these terms.

We fit propaganda as a type of influence operation within the larger context of information warfare. In other words, information warfare is an operational objective, aimed at "the deliberate manipulation or use of information by one party on an adversary to influence the choices and decisions the adversary makes in order for military or strategic gain" (Whyte et al., 2020, p. 344). Within this operational environment exists several tactics or tools to achieve a goal. These are the influence operations, which when tied to cyberspace and social media can be referred to as cyber-enabled influence operations (CEIO). As Nakayama (2022) writes, "information operations that leverage means and dynamics unique to cyberspace—with a particular focus on operations targeting social media" (p. 50). Within this realm, propaganda is the spreading of messages through social media channels and is a type of CEIO. As Bastos and Farkas (2019) demonstrate, "Propaganda campaigns are often implemented by state actors with the expectation of causing or enhancing information warfare" (p. 3). Thus, there is a nexus between influence operations and propaganda. In other words, in studying propaganda, this study also fits within the larger context of influence operations and social media information warfare. Thus, to understanding propaganda, one must understand influence operations in general (Cordey, 2019).

There are several understandings of influence operations especially on social media. Perhaps the most significant is Larson et al. (2009), who argue that in the context of U.S. national security, influence operations, "are the coordinated, integrated, and synchronized application of national diplomatic, information, military, economic, and other capabilities in peacetime, crisis, conflict, and postconflict to foster attitudes, behaviors, or decisions by foreign target audiences that further U.S. interests and objectives" (p. 2). Larson et al. (2009) argue that these operations influence a target audience, whether an individual leader, members of a decision-making group, military organizations and personnel, specific population subgroups, or mass publics (p. 2). In general, however, influence operations refer to intelligence operations that interfere in the affairs of another actor (Callanan, 2009; Maschmeyer et al., 2023). Many researchers argue that what is important is that the decision-making capabilities are interfered with through influence operations and information warfare. As Theohary (2018) makes clear, "Whether attacking government agencies, political leadership, or news media to influence public opinion or to complete decisionmakers to take certain actions, ultimately the target of information warfare activities is human cognition" (p. 2). Building on this, Bergh (2020) defines influence operations as a concerted effort by an actor, such as a state or terrorist group, to interfere in the process and meaning making by individuals or groups outside its own legal control through tools and facilities on publicly available social media services (p. 111). An influence operation is, therefore, an umbrella term or, as Cohen and Bar'el (2017) have put it, "a catchall phrase for any action intended to galvanize a target audience—an individual, a prominent group, or a broad audience—to accept approaches and to adopt decisions that mesh with the interest of the instigators of the operation" (p. 13).

As an umbrella term, influence operations include all types of operations in the information domain, not only propaganda operations but also clandestine and intrusive activities such as cyber-espionage and cyberattacks (Brangetto & Veenendaal, 2016). Among all these influence operations tactics, we focus on propaganda in this research because of the following reasons. First, focusing on propaganda allows us to speak to broader audiences and contribute to the rich literature of war and propaganda dating back to Harold D. Laswell, Edward Bernays, and Walter Lippmann. Second, compared with the umbrella terms, such as influence operations and information operations, the concept of propaganda is more operationalizable because studies from both political science and discourse analysis have provided various theoretical frameworks and analytical tools (for example, see Jowett and O'Donnell's, 2018, conceptualization and typology of propaganda and Van Dijk's, 2011, discussion of logical fallacies used in discourse analysis). Finally, compared with other influence operation tactics such as cyberattacks, data on propaganda operations are more available because of social media. Therefore, we focus on state-sponsored propaganda operations on social media, which consists of coordinated accounts exploiting the online space to influence public opinion (Ng et al., 2022) to serve the interest of a state.

Regarding propaganda, we borrow Jowett and O'Donnell's (2018) definition that propaganda is "the deliberate, systematic attempt to shape perceptions, manipulate cognitions, and direct behavior to achieve a response that furthers the desired intent" (p. 2). This definition emphasizes that (1) propaganda is deliberate, meaning that it serves a certain purpose favoring the propagandist's interest; (2) propaganda is systematic, indicating that the propaganda efforts are usually large-scale and persist over time; and (3) the goal of propaganda is to manipulate the audience. Therefore, when a state implements purposeful propagation of an idea or narrative intended to influence a target audience (Theohary, 2018) to shape the audience's perception and manipulate public opinion, it can be understood as a form of influence operations.

The remainder of this article proceeds in the following order. First, we provide a basic review of social media influence operations and propaganda, especially through state-sponsored activities. We then turn to our modeling and research design elements and describe how the technical aspects were implemented. Next, we present our case study analyses and provide case-specific results pertaining to the Kingdom of Saudi Arabia. Finally, based on these results, we provide policy recommendations and discuss holistically why understanding social cybersecurity in the context of Arabic tweets affects the policy realm of U.S. national Security.

Literature Review

Twitter is well known as a domain for the execution of information operations and the dissemination of propaganda to global audiences (Arif et al., 2018; Starbird et al., 2019; Uyheng et al., 2020). The platform's ability to both monitor and amplify information that "trends" or generates significant public interest makes it an ideal breeding ground for infiltrating and directing existing trends (Prier, 2017) or spreading one's own artificially generated trend to a wider audience (Guarino et al., 2020). These operations are often complex in nature and make use of cultivated networks of nodes known to be efficient at communicating a narrative to the largest possible audience in the shortest amount of time (Guarino et al., 2020). They may also be carried out by anyone with internet access, from individuals like Farah Baker (Wolf, 2015) to terrorist groups (Weimann, 2010) and state governments (Kießling et al., 2020). Understanding the success of these operations requires an explanation of the networks that make them possible and how they are employed. This section explores these networks, their creation, and the importance of central nodes or "influencers" and bots. An analysis of successful influence operations carried out on Twitter follows. These examples include terrorist efforts by organizations such as ISIS and al-Shabaab, campaigns carried out by states such as Iran and Russia, and notable independent individual or small group–led efforts such as those of Farah Baker in Palestine and the recent QAnon conspirators.

The networks that make information operations possible, while reaching an extensive and often diverse audience, are often centralized around a specific and small group of individuals who are considered an ideological and informational authority within their immediate community of like-minded followers (Dilley et al., 2022; Guarino et al., 2020). These individuals, or "influencers," use their communal authority to disseminate information to their followers, who share it with their own associates, with the eventual goal being to generate attention around articles of information (Guarino et al., 2020). If done correctly, this information may then generate enough attention that web users on twitter who are outside the typical spheres of influence will be exposed to the information disseminated (Prier, 2017). If those users find the information interesting and begin to interact with and spread that information, it creates what is commonly referred to as a "trend" (Prier, 2017), which will then be broadcast to a global audience by Twitter's front page.

Ideally, trends are intended to be determined by interest of Twitter's human users, but it is possible for these trends to be artificially constructed or "hijacked" as described by Prier (2017). The hijacking as described by Prier (2017) is carried out by directing swarms of bots to generate artificial clicks and abuse the algorithm employed by Twitter to seek out and promote popular tweets, articles, and messages. Prier's (2017) hijacking typically references the steering of existing narratives toward an interested party's perspective of events, but it should be noted that propaganda does not require an existing trend to spread.

Various influencers (Dilley et al., 2022) or "hubs" scattered (Caldarelli et al., 2020, p. 2) across Twitter's userbase command swarms of bots that promote their ideas to disseminate them to wider audiences. These "bot squads" are "groups of bots that follow and retweet the same group of hubs," thus amplifying their messages and making them more likely to trend (Caldarelli et al., 2020, p. 2). The nature of users to associate with "strongly clustered" communities "sharing similar ideas" to their own on Twitter helps boost the effectiveness of this strategy, essentially enabling hubs to generate propaganda trends that spread misinformation and disinformation virally (Caldarelli et al., 2020, p. 2). However, the impact of the message and its resulting influence is not limited to the "true believers" who affiliate themselves with ideologically aligned hubs (Prier, 2017, p. 58). Even those who exist in the "outside network" may be influenced by the propagandic messages, despite "not necessarily subscrib[ing] to the underlying beliefs that support the narrative" (Prier, 2017, p. 58). An infamous example of this phenomenon was seen as a part of the so-called "astroturfing" method used by QAnon conspiracy theorists such as Jason Sullivan, Ron Watkins, and Jim Watkins (Dilley et al., 2022, p. 8), all of whom served as "hubs" for the larger QAnon movement. Despite frequent association with Twitter propaganda and bot usage, it should be emphasized that "hubs" and bot usage are not limited to conspiracy theory style propaganda, any singular ideology, or even any specific actor type. In fact, despite the argument that "such platforms [as Twitter] democratize public discourse, recent years have shown how adversarial actors may employ diverse strategies to manipulate public opinion toward disruptive social and political outcomes" (Uyheng et al., 2020). Independent individuals, states such as Russia, Iran, North Korea, and China (Ferrara, 2020, p. 11) as well as terrorist organizations such as ISIS (Moriarty, 2015) have been known to employ this tactic. The following explore a few examples involving each of these actor types.

As the Westgate mall terror attack of 2013 was occurring, the terrorist organization responsible, al-Shabaab, released over 500 tweets (Mair, 2016). These tweets claimed responsibility and shared continuous updates and live messages to keep the public's focus on them (Mair, 2016). Throughout the 4-day siege, Twitter served as the primary method of communication between the attackers, the government, first responders, and the Kenyan public (Simon et al., 2014). Sullivan's rhetorical analysis of the tweets released during the attack describes the terrorist group's behavior as performative, with the intent clearly being to persuade the local populace to hear their message and support their efforts (Sullivan, 2014). Ultimately, Mair concludes that al-Shabaab's goal was two-fold: maintaining public interest in the attack and controlling the narrative, though their efforts were geographically targeted (Mair, 2016). However, they are not the only group to weaponize twitter to boost their message.

The advent of the Islamic State's (ISIS) rise to power and the coinciding "cyber jihad" aptly demonstrate the effectiveness of Twitter propaganda (Singer & Brooking, 2018). Horror stories emerged across the world of radicalized youth pledging allegiance to ISIS's cyber operations arm, with many engaging in Twitter activism for the group by sharing its ideological messages, videos of beheadings, and fear mongering against the terrorist organization's adversaries (Mitts, 2019; Singer & Brooking, 2018). The radicalization efforts adopted by the terrorist group drew countless headlines and significant scholarly attention, particularly as the group found success among Western audiences that would ordinarily have been unlikely targets and associates of a geographically and culturally foreign movement (Mitts, 2019). The leap from geographically localized terrorist Twitter propaganda such as that seen in the Westgate attack (Mair, 2016) to the larger, globally oriented propaganda efforts seen with ISIS (Singer & Brooking, 2018) demonstrated the potential of Twitter propaganda. In the battle for narrative control, terrorist actors seemed to have evened the playing field for the states they opposed (Singer & Brooking, 2018). However, this is not to say that states have ignored the potential of Twitter and other social media operations.

Alizadeh et al. (2020) demonstrate that it is quite difficult to determine whether social media posts are organic or state-sponsored influence operations. However, the researchers developed a platform-agnostic supervised learning approach to classifying posts as being a part of a coordinated and thus state-sponsored, influence operation or not (Alizadeh et al., 2020, p. 9). They find that content-based features distinguish coordinated influence campaigns and were able to use their supervised learning approach to detect influence operations from Russia, China, and Venezuela across social media platforms. They found that Chinese operations were easier to notice than Russia, and Venezuelan were the easiest to determine state influence at work (Alizadeh et al., 2020, p. 3). In addition, Ng and Carley (2023) analyzed online conversations on Twitter about the Chinese balloon spotted in American airspace in January 2023 and identified that over 46.05% of the Chinese accounts involved in the conversations were bots, which is higher than the average proportion of bot population on social media. They also found that in these conversations, Chinese accounts focused on the shooting and removal of the balloon as well as using narratives that are related to former U.S. president Donald Trump, such as "MAGA" and "SleepyJoe."

Ng et al. analyze image-based influence operations from China, Iran, Russia, and Venezuela. Interestingly, they find three distinct lines of effort for Chinese operations and argue that "The Image-Image network's structure and high clustering coefficient may correlate to high-level coordination and integration of influence operations" (Ng et al., 2022, p. 6). For Iran, they find that Image-based tactics include suppressing political dissidents using political hate speech, vulgar speech, counter-speech, and religion and societal topics (p. 6). Russia's network was highly connected with evidence to drive division within the U.S. political landscape, NATO, and interesting, lifestyle themes of food and travel (p. 6). For Venezuela, their Image–Image network is more decentralized, and they focus on "Breaking News" to describe their own operations as news (p. 6). For all of them except China, there were images and memes of U.S. politicians, showing close correlation between Iran, Russia, and Venezuela with less Chinese connection (p. 6).

Focused on Russia's influence operations, Lukito (2020) argues that the Russian government–supported Internet Research Agency (IRA) produces and disseminates disinformation targeted at the United States across various social media platforms, including Facebook, Twitter, and Reddit, and IRA activities on Twitter are influenced by IRA activities on Reddit. The author further points out that it may be because the IRA is experimenting and trial ballooning on one Reddit to figure out the optimal information to distribute on Twitter. In addition, with the growing tensions between the United States and Russia, Chernobrov and Briant (2022) claim that these two countries have witnessed mutual accusations of disinformation and propaganda campaign targeted at each other, and the threat of disinformation campaign has become an important part in the relationships between the United States and its rivals.

Research on propaganda and information operations in the Arab world largely focused on conventional media and nonstate actors' use of social media. For example, Fahmy et al. (2012) have analyzed how satellite TV may shape public opinion in the Arab world, and Ali and Fahmy (2013) have explored the use of social media by protesters in Iran, Egypt, and Libya during the Arab Spring. Concerning state-sponsored Arabic social media influence operations, there seems to be much less academic literature. There are some key sources to highlight, however. DiResta et al. (2021) analyze Middle Eastern influence operations across social media networks and find a breadth of tactics and narratives, devoted to multiple geopolitical objectives-versus, for instance, Chinese or Russian operations, that tend to focus on singular objectives or targets (p. 99). They conclude that influence operations are gaining steam and becoming more important in the MENA region.

Russia's interference campaign in the 2016 U.S. election is perhaps the most frequently cited, but the state's efforts extend far beyond this. There is evidence to suggest the state's involvement in the Brexit debate in England (Llewellyn et al., 2018), the vaccine debate surrounding COVID-19 in the United States (Broniatowski et al., 2022), and a myriad of other hot button issues at the center of Western politics (Miller, 2019). Beginning in 2022, much of Russia's propagandic focus has shifted to the Ukraine war, although Pierri et al. (2023) interestingly note a distinct drop in "the prevalence of Russian propaganda following the invasion" while also being careful to emphasize that the presence of propaganda "is not negligible" (p. 8).

Beyond Russia, Iran has been observed using Twitter bots and propaganda accounts to attempt to influence the foreign and public policies involving Saudi Arabia, with the state actively promoting biased hashtags and retweeting identifiable propaganda sources (Kießling et al., 2020). Kießling et al. (2020) note that Iran's Twitter campaign activity tends to spike coinciding with major political events, although they also find that the majority of Iran's attempts at influencing foreign policy were unsuccessful. China has also been observed employing Twitter propaganda strategies, notably surrounding public discourse concerning territorial disputes in the South China Sea, in which most of the 19 "important China actors in the #SouthChinaSea conversation" were identified as representing "central-level state news media" (Nip & Sun, 2022, p. 61). While each of these campaigns have varied in effectiveness and scale, their very existence demonstrates the popularity of social media influence operations and Twitter propaganda and among state actors. However, even individuals can employ these methods to create powerful effects.

One notable example belongs to Farah Baker, the widely acclaimed "Anne-Frank of Palestine" (Wolf, 2015). Baker began trending on Twitter in 2014 during the Israeli attacks on Gaza as part of its campaign against Hamas (McNeill, 2019). Baker was only 16 at the time, but her tweets rapidly evolved into one of the most successful narrative campaigns seen on Twitter (Patrikarakos, 2017). The narrative campaign rapidly picked up traction for its perceived authenticity and emotional gravity, allowing Baker to amass a substantial following (Patrikarakos, 2017). This following transformed Baker almost overnight from a humble teenager into an influencer hub with a vast network of like-minded supporters who spread her message to a global audience (Patrikarakos, 2017). This campaign ended up being so influential that it forced the Israeli Defense Force to reevaluate its own public relations and narrative influence strategies (Patrikarakos, 2017). A similar phenomenon occurred more recently with the spread of vaccine-related conspiracy theories on Twitter. About 65% of these conspiracy theories were traced back to a "Disinformation Dozen" of individuals running multiple accounts across the platform (Bond, 2021, para. 8). This handful of individuals managed to build massive community networks that spread propaganda across multiple social networks, including Twitter, to generate millions of retweets and views, eventually reaching many users outside their usual spheres of influence (Nogara et al., 2022). This network was so influential that U.S. Congressional Representatives and multiple state attorney generals repeatedly urged social media platforms, including Twitter, to ban the accounts of the "Disinformation Dozen". While these platforms took efforts to do so, the hubs were resilient in their creation of multiple new accounts to replace their lost ones.

Even within the world of Western democracies, intentional use of bots on social media to serve certain political purposes has been increasing. For example, Howard and Kollanyi (2016) find that political bots have been strategically used in the United Kingdom's 2016 referendum on leaving the European Union. Another European country, Germany, has also witnessed the use of bots by both political and private actors on social media to manipulate the public sphere and thereafter the 2017 German Federal Election (Neudert, 2018). In the United States, Howard et al. (2018) and Howard et al. (2017) have demonstrated the influence of bots, mis/disinformation, junk news, algorithms, and automated political communication in general on both local and federal-level elections. These examples demonstrate the power and prolificity of Twitter propaganda. Effective propaganda can be generated by almost anyone in the world, with the only requirement being an internet connection (Patrikarakos, 2017). Social media in general, but Twitter in particular, are likely to remain an influential component of public discourse surrounding major political events for the foreseeable future. As such, they will also remain an avenue to dispense propaganda and shape the course of narrative battles between states, interest groups, individuals, and even adversaries to the global system like ISIS. Now that a brief review of the literature has been presented, let us turn to the specifics of the current project.

Research Design

Twitter Data

This article addresses a new category of propaganda attacks that are tied to state-linked accounts that spread anti-U.S. propaganda by taking advantage of specific geopolitical crises in the Middle East. We investigated the role of general language models and general training data to detect those forms of targeted propaganda. Our general propaganda data are selected from a public data set. The state-linked data are selected from Twitter Moderation Research Consortium (TMRC), through which Twitter shares large-scale data on information operations to the public since 2018. TMRC has published data sets of state-linked information operations originating from various countries, including Iran, Russia, China, Saudi Arabia, and more. For the purpose of this research, we focus on state-linked Twitter accounts' activities originating from Middle Eastern countries, including Saudi Arabia and Egypt. The following section describes each of both data sets in detail.

Data Collection

We used two distinct data sets to study both general and targeted propaganda. Data Set 1, "General propaganda" is a preexisting labeled data set consisting of a diverse collection of tweets sourced from well-known news outlets in Arab countries, supplemented by international news sources. This data set consists of 3,200 tweets from each source, with an additional 100 random tweets per source for augmentation, resulting in a sample of labeled 930 tweets. Data Set 2, referred to as "Targeted propaganda," was selected from Twitter's publication of a publicly available archive covering state-backed information operations from 2018 to 2022. The present research narrowed its focus within this data set to tweets related to the United States and predominantly characterized as anti-US propaganda. The data selected from this data set were not prelabeled. The labeling process encompassed both binary and multilabel classification, conducted by native Arabic-speaking annotators. The following two subsections describe each data set in detail and the labeling process for Data Set 2.

Data Set 1: General Propaganda

This data set was collected from the top news sources in Middle Eastern states, which include the social media pages for news sources such as Al Arabiya, Sky News Arabia from UAE, Al Jazeera, and Al Sharq from Qatar (Alam et al., 2022). Five international sources were added to those sources, including Al-Hurra News, BBC Arabic, CNN Arabic, and France24. The most recent 3,200 tweets from each source were selected. Another 100 random tweets were also used to augment each source. Then, a sample of 930 tweets for annotation were selected. As this is a multilabel classification problem, a skewed distribution is noticed in this data set as shown in Table 1.

Data Set 2: Targeted Propaganda

Between 2018 and 2022, Twitter published a comprehensive, public archive of data related to state-backed information operations. Thirty-seven data sets have been shared as a part of this effort. The published data sets attributed platform manipulation campaigns originating from 17 countries, spanning more than 200 million Tweets and nine terabytes of media.

As we were interested in targeted propaganda, we only focused on the tweets related to the United States and are mainly anti-United States in nature. The selected accounts were generally tweeting propagandic themes such as

- Muslim Brotherhood in the Middle East
- Iranians' role in the Middle East
- War in Syria
- Sanctions on Qatar

We created a dictionary to select all tweets related to the United States and U.S. institutions to include the U.S. military, the U.S. Army, and U.S. Armed forces in general. We also included keywords related to U.S. leaders who were active during the timeframe of tweets such as Donald Trump and Jared Kushner. The number of state-based tweets per each data set is shown in Table 2. Accounts were suspected to be linked to Saudi Arabia. We followed a rigorous procedure to label the data. As these are targeted propaganda tweets and need to be labeled against the 17 propaganda categories, the labeling process was time-consuming. It started in October 2022 and ended in January 2023. The labeling process consists of binary class classification and multilabel classification. The labeling process is demonstrated in Figure 1.

In the first phase, two annotators who are native Arabic speakers spent some time understanding the propaganda techniques in Arabic. They were given examples on each technique from the first General Propaganda Data Set 1 and many other publicly available examples on the web. The two annotators then conducted a binary classification to classify tweets into no propaganda labels (0) and (1), which is potential propaganda. Potential propaganda tweets were then labeled in the second phase to identify and extract contiguous spans of text that correspond to at least one propaganda technique. Propaganda span detection focused on identifying specific segments or spans of text within each tweet that contained elements of propaganda. Given that a single tweet can encompass multiple types of propaganda as shown in Table 1, span detection techniques allow to highlight and isolate these propagandistic phrases or sentences. This approach enables a more granular analysis, enabling us to better understand the various propaganda tactics employed in a single tweet and help in the development of effective countermeasures against the spread of disinformation and biased content on social media platforms.

In total we labeled about 222 as propagandic tweets in the first and second phase of labeling. Only 28 nonpropagandic tweets were found in the labeled data, which were excluded in the second phase. The two annotators hold graduate degrees in information systems. For the second phase labeling, we consulted a third annotator who is a Jordanian domain expert in the fields of political science, press, and media.

The average disagreement in the labeling processes in the multiclass stage using the Kappa index was 11%, which

Table 1. General Propaganda Data.

Propaganda technique	Description	No. of tweets
Appeal to authority	Stating that a claim is true simply because a valid authority or expert on the issue said it was true, without any other supporting evidence offered	28
Appeal to fear/prejudice	Seeking to build support for an idea by instilling anxiety and/or panic in the population toward an alternative	55
Black-and-white fallacy/ dictatorship	Presenting two alternative options as the only possibilities, when in fact more possibilities exist	3
Causal oversimplification	Assuming a single cause or reason when there are multiple causes for an issue	5
Doubt	Questioning the credibility of someone or something	30
Exaggeration/minimization	Either representing something in an excessive manner: making things larger, better, worse or making something seem less important or smaller than it really is	54
Flag-waving	Playing on strong national feeling (or to any group, for example, race, gender, political preference) to justify or promote an action or idea	7
Glittering generalities (virtue)	These are words or symbols in the value system of the target audience that produce a positive image when attached to a person or issue. Peace, hope, happiness, security, wise leadership, and freedom	32
Loaded language	Using specific words and phrases with strong emotional implications (either positive or negative) to influence an audience.	492
Name calling/labeling.	Generate fear and bias using derogatory terms to form an adverse judgment against a person, a group, ideologies, concepts, or institutions that they want us to condemn	288
Obfuscation, intentional vagueness	Using words which are deliberately not clear so that the audience may have its own interpretations	12
Presenting irrelevant data	Introducing irrelevant material to the issue being discussed, so that everyone's attention is diverted away from the points made	1
Repetition	Repeating the same message repeatedly so that the audience will eventually accept it	11
Slogans	A brief and striking phrase that may include labeling and stereotyping. Slogans tend to act as emotional appeals	45
Smears	A smear is an effort to damage or call into question someone's reputation, by propounding negative propaganda	97
Thought terminating	Words or phrases that discourage critical thought and meaningful discussion about a given topic	7
Whataboutism	A technique that attempts to discredit an opponent's position by charging them with hypocrisy without directly disproving their argument	4

Table 2. Our State Backed Propaganda Labeled Data.

Timeframe	Number of states linked accounts	Number of labeled propaganda tweets
December 2019	Saudi_Arabia_112019 (5929 users))	110
October 2020	Saudi Arabia (qatar_082020) (34 users)	108

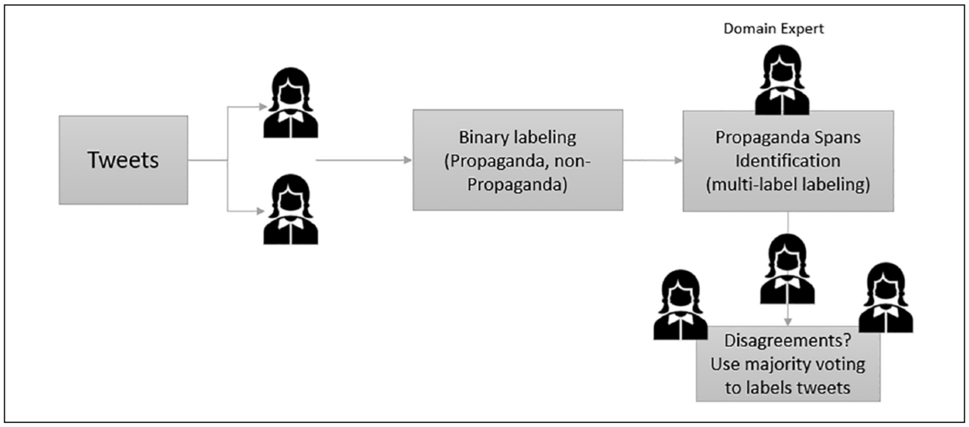

Tweets
Binary labeling (Propaganda, non-Propaganda)
Domain Expert
Propaganda Spans Identification (multi-label labeling)
Disagreements? Use majority voting to labels tweets

Figure 1. Labeling process of the targeted propaganda data set.

Table 3. Sample of General and Contextualized Propaganda.

Propaganda technique	General propaganda (GP)	State-linked/contextualized and targeted propaganda (TP)
Appeal to authority	An Egyptian observer: The 2014 elections were the declaration of liberation, and the current beginning of reconstruction	America knows that if Saudi Arabia gets angry at it, it means that the entire Islamic world is angry, and only Saudi Arabia may absorb the anger
Appeal to fear/prejudice	The pollution of the largest artificial lakes in Lebanon raises the alarm and warns of an environmental disaster	It is not surprising that Israel and America do not hesitate to punish anyone who opposes them, and there is no consideration for any values or covenants if they contradict their interests, unlike the Arabs
Black-and-white fallacy/dictatorship	Dialogue—Talal Abu-Ghazaleh: There is no other solution in Palestine except with the end of the occupation	Iran has two solutions, the sweetest of which is bitter; Either being naked for America or being naked for the Iranian people
Causal oversimplification	A bad future awaits humanity. Artificial intelligence is in the dock	The Arab Spring is an American Spring in origin. America has completely laundered its files in the Middle East. It removed agents and brought new agents
Loaded language	UAE Sheikha Jawaher Al Qasimi criticizes the normalization of education with Israel: "Their curricula recommend killing and usurping Arab land.	They will expel his children / Saudi Arabia is the center of Islam and will not allow the dogs of the "Rafidites," their hypocrisy, their ally Israel, and complicit America to celebrate

Table 4. Propaganda Tweets With Propaganda Spans.

Tweet example	Propaganda spans
The trump administration's sanctions on the International Criminal Court deepen America's failure to deliver justice for the most heinous global crimes	"Loaded Language", "Smears"
Syrian regime oil minister: 90% of Syrian oil is under American influence, and its promises of deliveries are fading	"Loaded Language", "Name calling/Labeling"

consists of cases where one of the annotators labeled a tweet as belonging to a propaganda category and the two others disagree. In those cases, the annotators conducted a second round of labeling to avoid labeling unintentional errors; we then used the majority voting rule to produce the final labeled data set. As with the previous research efforts in this area, the distribution of the propaganda techniques in the newly collected data set was skewed. The sample of our labeled contextualized (targeted propaganda) and general propaganda is shown in Table 3. In Table 4, we also show a sample of propaganda spans in the same sentence.

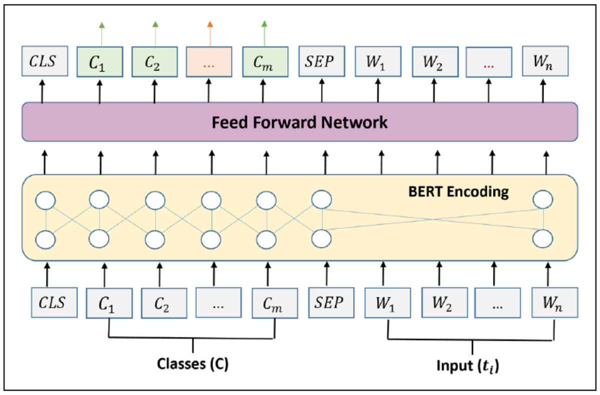


Figure 2. Propaganda detection architecture.

Machine Learning Analysis

We use Alhuzali and Ananiadou's (2021) model, SpanEmo, a machine learning algorithm that can conduct multilabel classification and span detection tasks, to analyze our data[1]. We selected this model for two particular reasons. First, to answer our research question of "To what extent can these issues be described as propagandic in nature?" we need a classification algorithm which can learn from features and labels in the training data and thereafter perform classification tasks on the unseen data (testing data). This task is different from what we commonly see in conventional social science, which is usually causal inference using statistical modeling. Therefore, a machine learning algorithm would be more appropriate to conduct this kind of task. Second, in both of our training and testing data set, each tweet could have more than one label, and this model would allow us to explore the situation where one piece of propaganda uses multiple techniques mentioned before and predict the techniques that are used in the unseen data (testing set). Based on the understanding of multilabeled propagandic tweets, we will be able to distinguish propagandic tweets from nonpropagandic tweets more accurately. Figure 2 summarizes the approach and the deep learning architecture we utilize to detect propaganda categories from the same tweet.

Our propaganda detection approach casts multilabel propaganda classification as span-prediction, which learns associations between labels and words in tweets. Therefore, we utilized the model by Alhuzali and Ananiadou (2021) to classify into 17 categories of propaganda. Let $\{(t_i, y_i)\}_{i=1}^{N}$ be a set of N tweets with the C propaganda classes, where t_i represents the training tweet and $y_i \in \{0,1\}^m$ m are the set of labels. Figure 1 shows how the classes and the training tweets are used as inputs to the model. The training data are processed by a $BERT$ encoding approach developed by Devlin et al. (2018). The inputs to the encoder are the propaganda classes and the training tweets. The hidden layer $(H_i \in R^{T\times D})^2$ containing the training tweets and the set of classes set is obtained as follows: $H_i = Encoder\left(\left[CLS\right]+\left|C\right|+\left[SEP\right]+t_i\right)$. In this formula, $\{[CLS],[SEP]\}$ represents special tokens that are added to the data and $|C|$ denotes the number of propaganda classes. T^2 and D are the length of the input and the dimensionality of data. A feed-forward network (FFN) is utilized, with a nonlinear hidden layer, a Tanh activation $(f_i(H_i))$, and a vector $p_i \in R^D$, which calculates the dot product of f_i and p_i. As our task involved a multilabel propaganda classification, a sigmoid activation is used to determine whether a $class_i$ should be included in the predicted classes as $\hat{y} = sigmoid\left(FFN\left(H_i\right)\right)$. The span-prediction tokens were compared with the ground truth labels as there is a one-to-one mapping with such labels. Using the approach by Yeh et al. (2017). We used the label correlation aware loss as an objective function as $\mathcal{L}_{LCA}\left(y,\hat{y}\right) = \frac{1}{\left|y^0\right||y^1|}\sum_{(p,q)\in y^0\times y^1}\exp\left(\hat{y}_p - \hat{y}_q\right)$. This loss function also fits our training objectives to detect co-occurrence of propaganda because it splits labels into positive and negative pairs based on their co-occurrence. In Formula 3, y^0 denotes the set of negative labels and y^1 denotes the set of positive labels. $\hat{y}_p$ represents the p^{th} element of vector $\hat{y}$. The objective of this loss function is to maximize the distance between the labels based on their co-occurrence. In other words, the model loss increases if it predicts a pair of propaganda labels that should not co-exist for a given tweet. As it was the case in Alhuzali and Ananiadou (2021), the model label-correlation loss is combined with the binary cross-entropy. This aims to help the label-correlation loss to focus on maximizing the distance between co-occurrences while at the same time taking advantage of the binary cross-entropy to maximize the probability of the correct prediction. The overall training objective was computed as follows $\mathcal{L} = \left(1-\alpha\right)\mathcal{L}_{BCE} + \alpha\sum_{i=1}^{M}\mathcal{L}_{LCA}$, where $\alpha \in [0,1]$ denotes the weight used to control of each loss function.

Experiments

Using both data sets, we conducted three types of experiments on our deep learning model:

Table 5. Hyperparameter Settings.

Parameter	Value
Feature dimensions	732
Batch size	6
Early stop patience	50
Number of epochs	50
lr-BERT	2e-5
Optimizer	Adam
Alpha α	0.2
Drop out	0.1

- GP: The training and development data were selected from the General Propaganda (GP) Data Set 1. The testing data are selected from our Targeted Propaganda (TP) Data Set 2.
- TP: The training, development, and testing data were selected from our TP Data Set 2.
- FT: Training data are selected from GP; then, the model is finetuned (FT) on the TP data from Data Set 2.

All our experiments were conducted using the same hyper parameters settings shown in Table 5.

The technical settings of our experiments consist of using the PyTorch implementation (Paszke et al., 2017), HuggingFace implementation of BERTBASE (Wolf et al., 2020), and BERTBASE-ARABIC (Safaya et al., 2020) as a pretrained Arabic-language model. We leave testing our approach on other Arabic-language models such as ARABERT and MARBERT as a future work. Table 6 shows the characteristics of Training (T), Development (T), and Testing (TS) experimental data selected from both GP and TP data sets. Several evaluation measures were used in our experiments as follows:

- The Jaccard Similarity (JaccS): This measure is usually used to quantify the degree of similarity between two sets, such as a set of true labels and a set of predicted labels. It is calculated by dividing the size of the intersection of these two sets by the size of their union. The Jaccard Similarity measures how much overlap or commonality exists between the elements present in the true label set and the predicted label set. A higher Jaccard Similarity score indicates a greater degree of overlap or agreement between the two sets, while a lower score suggests less agreement and more dissimilarity. As we are measuring the reliability of classification for a multilabel classification, which involves assigning multiple labels or categories to each instance, the Jaccard Similarity handles this scenario making it well suited for evaluating the overlap between predicted labels and true labels.
- Macro and Micro F1: Macro-F1 is calculated by averaging the F1 scores for each class individually, while micro-F1 is calculated by averaging the precision and

Table 6. Training, Development, and Testing Data Sets.

Technique	GP			TP			FT		
	T	D	TS	T	D	TS	T	D	TS
Appeal to authority	21	7	2	2	1	2	21	1	2
Appeal to fear/prejudice	48	7	1	3	1	1	48	1	1
Black-and-white fallacy/dictatorship	2	1	1	2	1	1	2	1	1
Causal oversimplification	4	1	1	1	1	1	4	1	1
Doubt	29	1	7	29	18	7	29	18	7
Exaggeration/minimization	44	10	1	8	2	1	44	2	1
Flag-waving	5	2	5	11	7	5	5	7	5
Glittering generalities (virtue)	25	7	2	2	3	2	25	3	2
Loaded language	446	46	6	6	5	6	446	5	6
Name calling/labeling.	244	44	8	3	4	8	244	4	8
Obfuscation, intentional vagueness	9	3	3	7	7	3	9	7	3
Presenting irrelevant data (red herring)	1	0	4	4	3	4	1	3	4
Repetition	9	2	1	1	1	1	9	1	1
Slogans	44	1	1	1	1	1	44	1	1
Smears	85	12	2	6	5	2	85	5	2
Thought-terminating cliché	6	1	2	2	1	2	6	1	2
Whataboutism	3	1	1	4	1	1	3	1	1

Note. GP = general propaganda; TP = targeted propaganda; FT = finetuned; TS = testing.

recall scores for all classes, regardless of their size. Macro-F1 is used when some classes are more important than others, or when we have a data set that is balanced. Micro-F1 is a good metric to use when all classes are equally important, or when we have a data set that is imbalanced. In this work, we used both measures to get a more complete picture of the performance of their model on both balanced and imbalanced data sets.

- Training loss: It is a metric that quantifies how well a machine learning model is performing on its training data during the training process. It represents the error or the difference between the predicted values and the actual target values for the training examples.
- Validation loss: It is a metric used to evaluate a machine learning model's performance on data that it has not seen during training. This separate data set, called the validation set, is distinct from the training data.

Results

The results of our experiments are shown in Figures 3 and 4. Figure 3 shows our experimental results using the Saudi accounts tweets as our TP data. We noticed significantly better metrics when training using TP data compared with training using GP (Figure 3b compared with a). We run our experiments on micro F1-score, macro F1-score, and Jaccard index score. The latter is the size of the intersection divided by the size of the union of the true label set and predicted label set. We observed better F1-Micro, F1-Macro, and JS using training on TP. We also observed that training on GP and then finetuning the model using TP data are the least effective approach to detect targeted propaganda (Figure 3c). This result was validated on the testing data as well, where training using TP led to the best model accuracy in terms of our classification metrics (Figure 3d). It is also noticed that training and validation using TP leads to the most stable model in terms of both training and validation losses (Figure 3b). Results on Qatari accounts tied to the Saudi government were not significantly different compared with Saudi accounts (Figure 4). It is noticed, however, that FT leads to relatively better results compared with Saudi accounts (Figure 4c compared with 3c). We think that this is related to the sources where the GP data set was collected, where it was mainly from some Qatari news networks such as Qatar's Aljazeera news network. As such, similarity between TP of Qatari accounts and GP data leads to better finetuning results.

The results shown in Figure 3a and b highlight the advantages of incorporating targeted propaganda for training purposes. Specifically, as seen in Figure 3b, the F-Micro score achieves a noteworthy increase, reaching 0.63, compared with 0.56 in scenarios employing general propaganda data (Figure 3a). In addition, the Macro F1 scores exhibit significant improvement when the model is trained using the targeted propaganda data set (as demonstrated in Figure 3b). Moreover, training and validation losses highlight the model's robust convergence when trained with the targeted propaganda data set. Figure 3b illustrates a gradual reduction in both training and validation loss values, compared with Figure 3a, where the model's learning capabilities are less effective. It is important to note that an alternative approach,

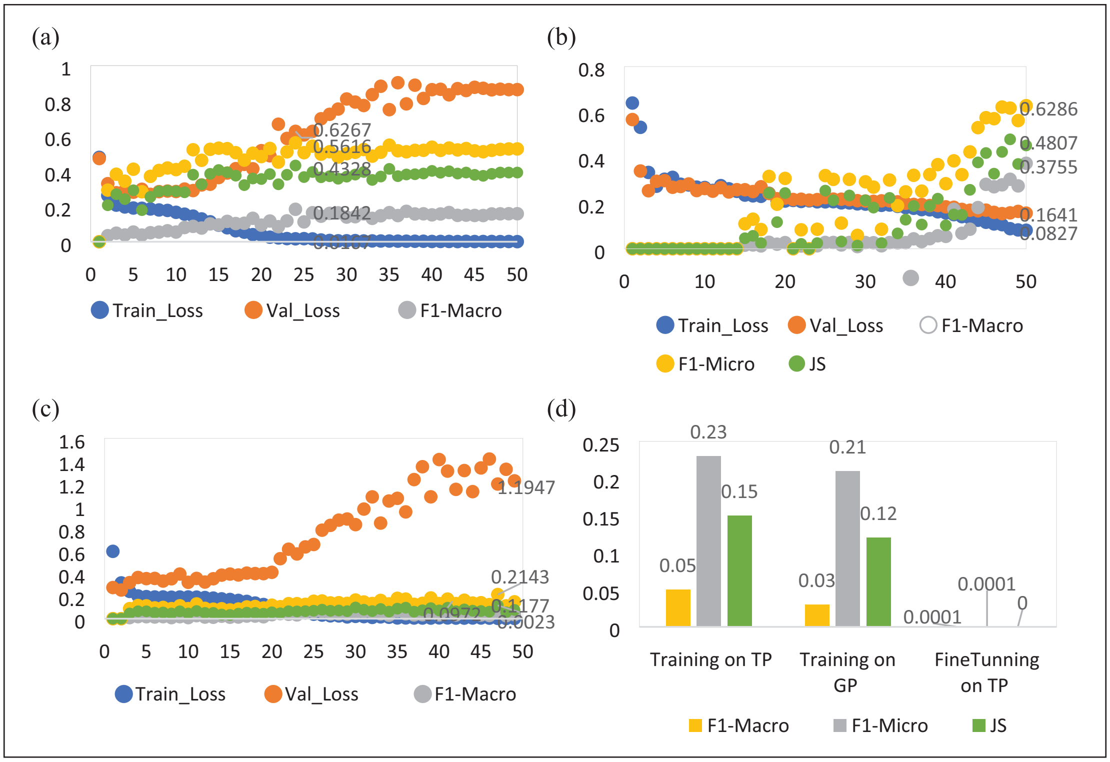


Figure 3. (a) Training using GP. (b) Training using TP. (c) FT on TP. (d) Results on testing data.

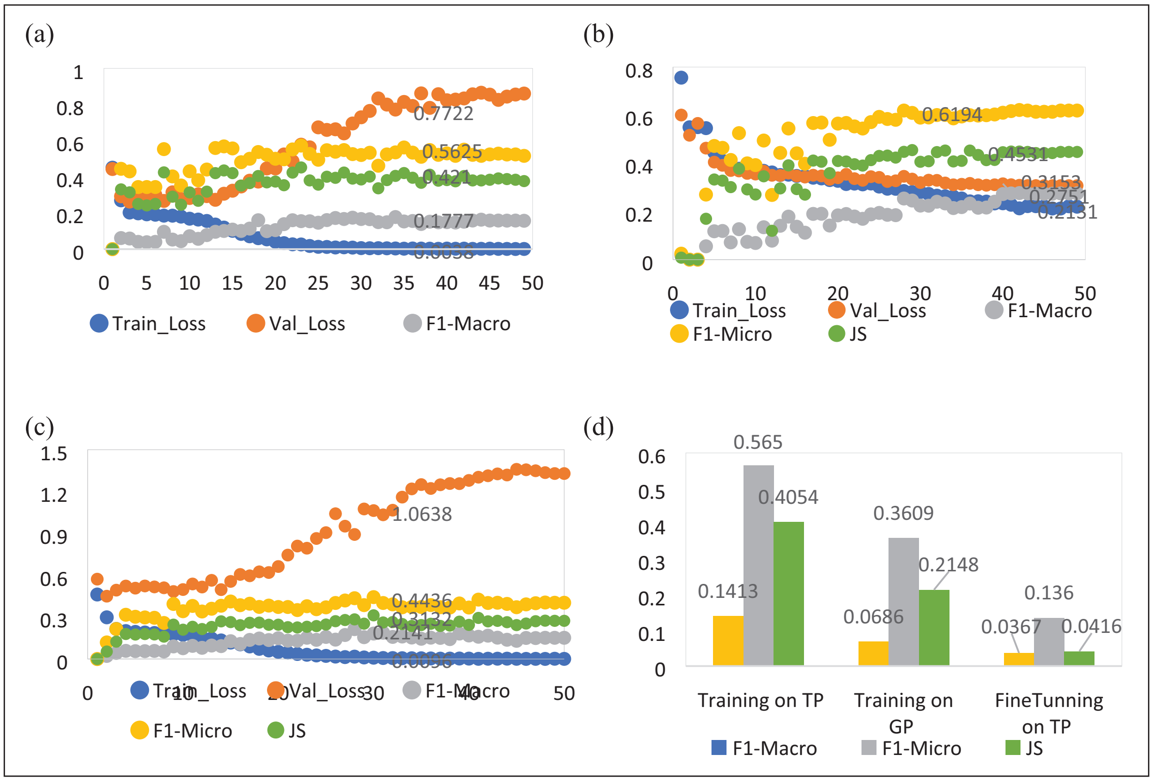


Figure 4. (a) Training using GP. (b) Training using TP. (c) FT on TP. (d) Results on testing data.

involving initial training with general propaganda data followed by finetuning with targeted propaganda data, results in less significant results. Specifically, this approach yields a low JS (Jaccard Similarity) score of approximately 0.11, along with F1-Micro and F1-Macro scores of 0.21 and 0.10, respectively. Furthermore, the consistency of these results on the testing data set, as depicted in Figure 3d, shows the advantages of incorporating targeted propaganda data when predicting targeted propaganda attacks. Figures 4a to d shows almost similar results patterns.

The Case of Saudi Arabia

To illustrate this article's points more specifically and contextually, a single case study is needed to illustrate better the issues discovered. To delve more deeply into the topics, this article presents a case analysis of Saudi Arabia. Before we get to case-specific results on Saudi Arabia, it is important to place the case in the context of this case study. There are several reasons to investigate propaganda activities in Saudi Arabia as opposed to the other countries also investigated in the article, which includes its significant role in geopolitical events. First, it should be noted that Saudi Arabia has one of the highest percentages of social media users in the Middle East (25 million people out of a total population of just under 35 million, and within this, there are an estimated 20 million Twitter users in Saudi Arabia; Shakil et al., 2021). In addition, as a major energy producer Saudi Arabia plays a vital role in the global energy market. In fact, Saudi Arabia provides around 15% of the U.S. crude oil imports and more than 15% of the global crude oil imports yearly between 2000 and 2017 (Beckman & Nigatu, 2021). Furthermore, research shows that negative oil-supply shocks due to sanctions, wars, policies, or natural disasters in Saudi Arabia have an immediate and permanent increase in global oil prices as other oil producers cannot make up for the decrease in Saudi Arabian oil production (Mohaddes & Pesaran, 2016). Saudi Arabia holds a unique religious significance in the Muslim world. Saudi Arabia is the home of Mecca and Medina, the two holiest cities in the Muslim World, and the state "continues to be a center of religious training and its soft power influence in the Sunni World is unmatched" (Ciftci & Tezcür, 2016, p. 6). In addition, although Saudi Arabia may not be willing to openly endanger its relationship with the United States, recent evidence shows that Saudi Arabia has been involved in influence operations against the United States on social media. For example, two Twitter employees were charged with spying for Saudi Arabia on U.S. soil in November 2019 (Barrie & Siegel, 2021), and hashtags such as "Agents of the Embassies" have been used by the Saudi state on social media to resist influence from the Western world (Abrahams & Leber, 2021a). Therefore, we decided to focus on Saudi Arabia as one of, if not the most important great power within the region as it pertains to U.S. interests and great-power competition.

In fact, David Long (2019) argues that there is a unique relationship between Saudi Arabia among the Arab world toward the United States as it holds constant the need for close relations with the United States. In addition, with current tensions rising between Saudi Arabia and Iran, Israel's status in the Middle East (Beck, 2020), and the continuing crisis in Yemen (Darwich, 2020) where Saudi intervention has played a vital role, Saudi Arabia makes a great choice as a case study to examine in this context. It is important to note that as social conditions in Saudi Arabia have been progressing, there have been challenges to the regime; to prevent widespread opposition, the government has significant control over the internet and its censorship powers are steadily increasing (Chaudhry, 2014). Even with censorship, however, there have been movements in the Kingdom's Twittersphere promoting change through hashtag campaigns, for instance, including #women2drive (Chaudhry, 2014).

Abrahams and Leber (2021b) note that Saudi Arabia is one of the most prolific authoritarian regimes in the Middle East that use Twitter as a form of control and as a source of power, to the point that the regime managed to even place spies within the company itself. They note that much of the pro-authoritarian speech on Saudi Twitter is the result of organic activity driven by influential accounts that have built up their followings by toeing the party line—either voluntarily or by regime pressure (Abrahams & Leber, 2021b, p. 1174). Other research concerning Saudi Arabia has focused on emotional analysis of Twitter users living in holy cities compared with more secular metropolitan centers (Shakil et al., 2021), and engagement of Saudi citizens with IO campaigns in general compared with more mainstream news outlets (Barrie & Siegel, 2021). In fact, these authors find that engagement with IO within the Kingdom is not substantial compared with the level of Twitter uses overall, even during significant news events such as the murder of Jamal Khashoggi (Barrie & Siegel, 2021). Now that some context has been provided, we turn to our specific findings.

We analyze propagandic tweets in Data Set 2 from Saudi Arabia using the BEND social cyber security framework proposed by Carley (2020). The BEND framework argues that "influence campaigns are comprised of sets of narrative and structural maneuvers, carried out by one or more actors by engaging others in the cyber environment with the intent of altering topic-oriented communities and the position of actors within these communities" (Carley, 2020, pp. 371–372). In other words, the BEND framework considers both the narratives in the cyber environment and the social networks of cyber actors. This framework explores communication objectives from two dimensions: whether the objective is positive or negative, and whether the objective is aimed at manipulating the narrative (positively or negatively) or manipulating the social networks (positively or negatively). This gives us a two-by-two table in which 16 maneuvers are identified, and the details are shown in Table 6. The BEND framework

Table 7. The BEND Framework.

	Manipulating the narrative		Manipulating the social network	
Positive	Engage	Messages that bring up a related but relevant topic	Back	Actions that increase the importance of the opinion leader or create a new opinion leader
	Explain	Messages that provide details on or elaborate the topic	Build	Actions that create a group or the appearance of a group
	Excite	Messages that elicit a positive emotion such as joy or excitement	Bridge	Actions that build a connection between two or more groups
	Enhance	Messages that encourage the topic-group to continue with the topic	Boost	Actions that grow the size of the group or make it appear that it has grown
Negative	Dismiss	Messages about why the topic is not important	Neutralize	Actions decrease the importance of the opinion leader
	Distort	Messages that alter the main message of the topic	Nuke	Actions that lead to a group being dismantled or breaking up, or appearing to be broken up
	Dismay	Messages that elicit a negative emotion such as sadness or anger	Narrow	Actions that lead to a group becoming sequestered from other groups or marginalized
	Distract	Discussion about a totally different topic and irrelevant	Neglect	Actions that reduce the size of the group or make it appear that the group has grown smaller

Source. Carley (2020).

(Table 7) has been adopted to explore social media content and networks relating to COVID-19 vaccinations (Blane et al., 2022, 2023). Furthermore, Danaditya et al. (2022) have also applied the BEND framework to analyze Indonesian Twitter content, another Muslim community, and by analyzing the maneuvers of narrative and network, they find that a small group of coordinated agents can lead to polarization in the online sphere. Therefore, we believe that the BEND (Carley 2020) framework can provide more insights on the analysis of the Saudi Arabia Twitter sphere.

We conducted an annotation process to map each of the propaganda categories to one or more of the BEND communication objectives. During this phase, we selected three annotators, two of them were media experts who speak Arabic and English. The third expert was an IT professor who has expertise in Natural Language Processing. The annotators, first, reviewed the descriptions of various propaganda techniques. They also reviewed the descriptions of communication effects as defined by the BEND framework. The annotators then proceeded to map each propaganda technique to one or more communication effects. This step involves making associations between techniques and their expected effects. To validate our mapping, we asked annotators to find at least one example from the data set that supports their annotations. This practical validation step is crucial for assessing the real-world applicability and accuracy of the mapping. As the size of the targeted propaganda data is small, the annotators focused partially on a subset of 11 communication objectives that were selected from both manipulation strategies and that are clearly presented in the targeted propaganda data set. While we didn't directly measure the impact on the social network, we rely on the narratives that may lead to such an impact. Table 8 shows one of those examples. For this labeling activity, we asked the annotators to resolve any disagreement between them, through comprehensive discussion among them.

Table 8. Mapping Between Targeted Propaganda and BEND Communication Effects.

Example:
The Arab Spring is an American Spring in origin. America has completely laundered its files in the Middle East. It removed agents and brought new agents.

Strongest Propaganda Span:
Causal Oversimplification
Communication Effect: "Dismiss"

The statement simplifies complex geopolitical events like the Arab Spring by suggesting a direct cause-and-effect relationship between the Arab Spring and American involvement. It implies that the Arab Spring is solely an outcome of American actions, which is an oversimplified view of a multifaceted situation. This oversimplification can lead to the "Dismiss" effect by downplaying other complexities of the Arab Spring.

The mapping results are shown in Figure 5. Some BEND strategies are associated with more than one propaganda technique, for instance, the excite effect is associated with six types of propaganda.

A frequency analysis of mapping the propaganda found in the state-linked data about the Saudi Arabia data set is shown in Figure 6. As most of the analyzed propaganda is aimed at criticizing, undermining, or discrediting the United States, "Dismay" and "Neutralize" strategies were frequently used. "Dismay" strategies aim to elicit fear, worry, or doubt, which can be effective in making an audience question or reject the target of the propaganda. "Neutralize" strategies, on the contrary, aim to diminish or dismiss the impact of an opposing viewpoint, which can be useful in a setting where there are conflicting ideas or narratives. As the effectiveness of a

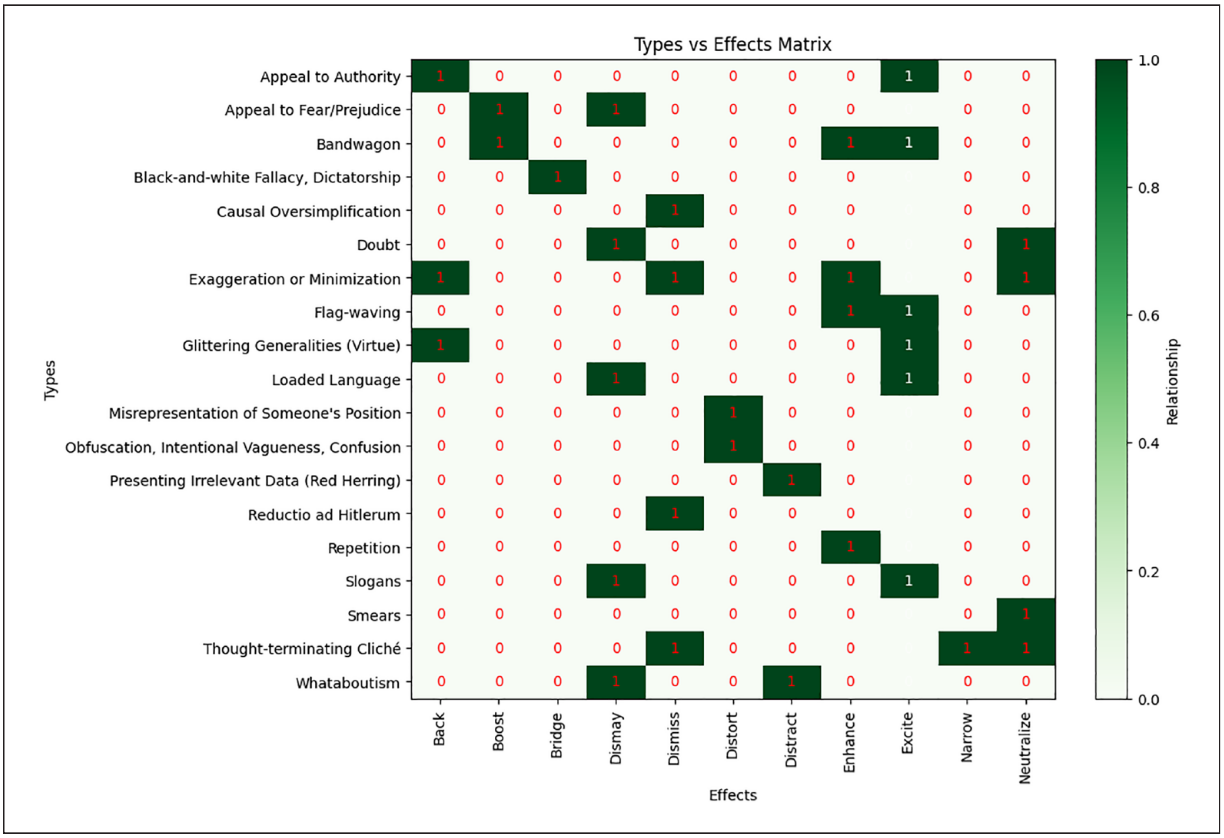


Figure 5. Mapping between propaganda categories and the BEND framework communication objectives.

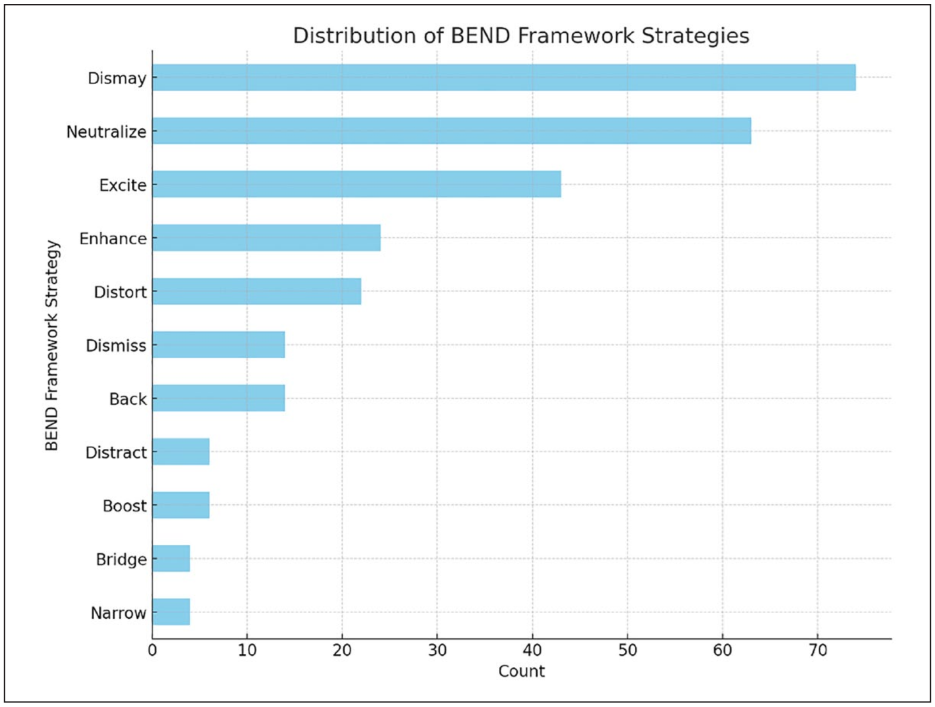
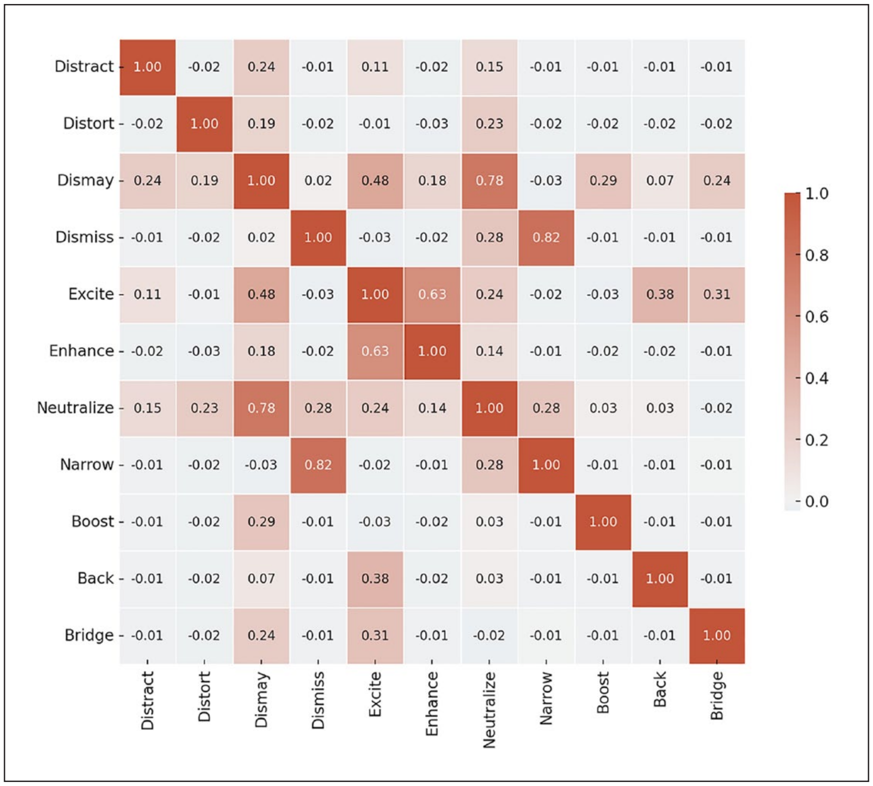


Figure 6. Frequency analysis of BEND communication strategies.



	Distract	Distort	Dismay	Dismiss	Excite	Enhance	Neutralize	Narrow	Boost	Back	Bridge
Distract	1.00	-0.02	0.24	-0.01	0.11	-0.02	0.15	-0.01	-0.01	-0.01	-0.01
Distort	-0.02	1.00	0.19	-0.02	-0.01	-0.03	0.23	-0.02	-0.02	-0.02	-0.02
Dismay	0.24	0.19	1.00	0.02	0.48	0.18	0.78	-0.03	0.29	0.07	0.24
Dismiss	-0.01	-0.02	0.02	1.00	-0.03	-0.02	0.28	0.82	-0.01	-0.01	-0.01
Excite	0.11	-0.01	0.48	-0.03	1.00	0.63	0.24	-0.02	-0.03	0.38	0.31
Enhance	-0.02	-0.03	0.18	-0.02	0.63	1.00	0.14	-0.01	-0.02	-0.02	-0.01
Neutralize	0.15	0.23	0.78	0.28	0.24	0.14	1.00	0.28	0.03	0.03	-0.02
Narrow	-0.01	-0.02	-0.03	0.82	-0.02	-0.01	0.28	1.00	-0.01	-0.01	-0.01
Boost	-0.01	-0.02	0.29	-0.01	-0.03	-0.02	0.03	-0.01	1.00	-0.01	-0.01
Back	-0.01	-0.02	0.07	-0.01	0.38	-0.02	0.03	-0.01	-0.01	1.00	-0.01
Bridge	-0.01	-0.02	0.24	-0.01	0.31	-0.01	-0.02	-0.01	-0.01	-0.01	1.00

Figure 7. Correlation between BEND communication strategies in the Saudi propaganda data.

propaganda technique often depends on how it resonates with its intended audience, "Dismay" and "Neutralize" strategies are particularly effective at influencing the attitudes and behaviors of the audience with respect to U.S. politics. The overall goals of the propaganda campaigns in state-linked data set also influence the choice of strategies. As the goal was to create confusion or stoke fears about U.S. policies, the "Dismay" strategy was a logical choice. The goal to challenge or undermine an opposing viewpoint also justifies the neutralize strategy.

We then conducted a correlation analysis to examine which strategies are used together. The correlation analysis results shown in Figure 7 also support our analysis. We can draw several findings from this correlation study as follows:

- Dismiss and Enhance (correlation: –0.02): The techniques of dismissal and enhancement do not often appear together in the analyzed propaganda data set. This suggests that Saudi state-linked actors dismiss certain information (likely negative information about Saudi Arabia or its policies); they do not typically try to enhance or amplify other information (e.g., positive information about Saudi–U.S. relations) within the same message.
- Dismay and Neutralize (correlation: 0.78): When trying to create a sense of fear or concern among audiences (Dismay), these actors often aim to neutralize some viewpoints. This means that they are trying to induce worry about certain topics while simultaneously reducing the impact or credibility of opposing viewpoints by the United States. Some supporting statements from the data set is "America will soon face its own trials and will perish. We seek refuge from its evil. InshaAllah." This statement expresses dismay toward America, suggesting that it will face negative consequences. The mention of seeking refuge from its evil attempts to neutralize the power or influence of America. Another statement is "The Iraqi government is elected!? I remember Ayad Allawi won the elections, but Iran and America want him destroyed." This example expresses dismay toward the Iraqi government, suggesting that it is not truly elected and controlled by external forces (Iran and America). By highlighting the influence of Iran and America, the statement attempts to neutralize the legitimacy or credibility of the Iraqi government.
- Dismay and Excite (correlation: 0.48): The positive correlation suggests that messages often aim to both arouse alarm and strong emotions or enthusiasm among audiences. This can be seen as strong emotions that favor the Saudi narrative while creating concern or fear about alternative viewpoints or actions by the United States.

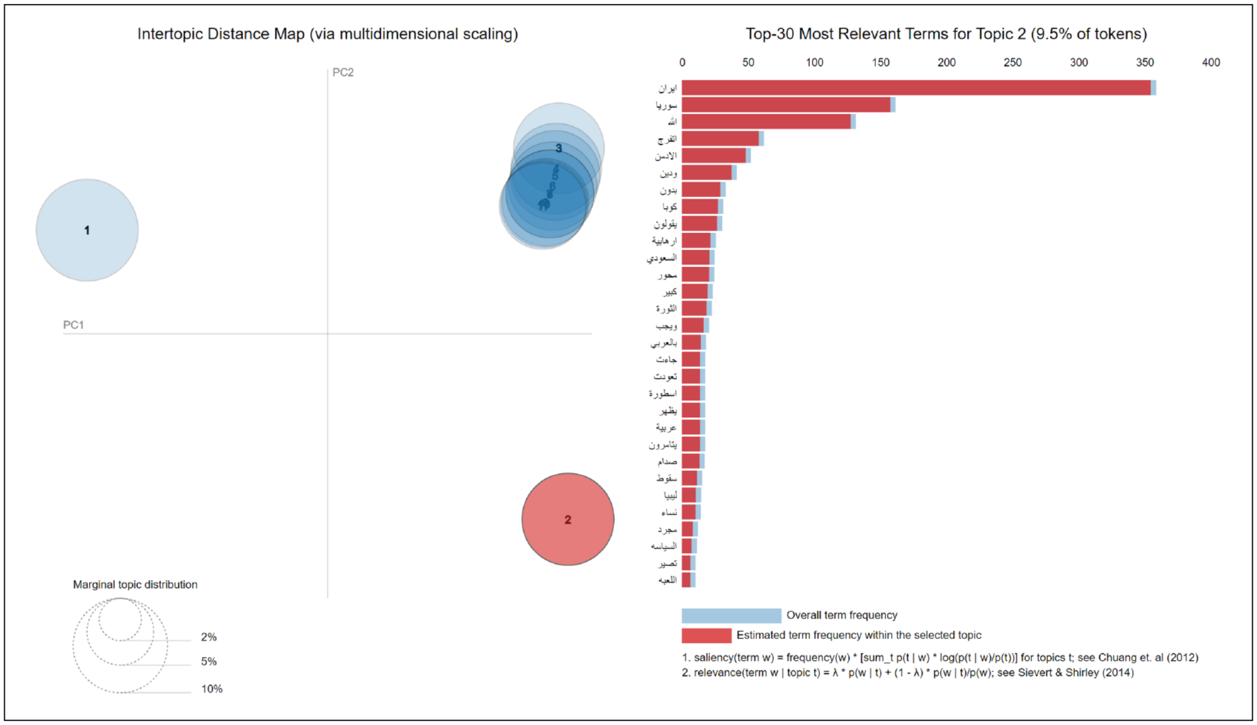


Figure 8. Topic analysis: Thematic Topics Sample 1.

- Excite and Enhance (correlation: 0.63): When Saudi state-linked actors aim to stir up strong emotions, they often try to amplify or enhance certain viewpoints or pieces of information. This could be part of a strategy to create emotional attachment to the messages they want to promote, such as positive perceptions of what Saudis did to the United States.
- Neutralize and Narrow (correlation: 0.28): This correlation indicates that attempts to neutralize certain viewpoints (likely those opposing Saudi interests) often go hand in hand with narrowing the range of debate. This could be a strategy to control the narrative within the United States, neutralizing opposing viewpoints, and limiting the discussion to topics where Saudi state-linked actors can more effectively push their preferred narrative.
- Dismay and Neutralize: Because this article focuses on targeted propaganda operations, our explanation of strong correlation between Dismay and Neutralize is that both techniques aim at creating a sense of doubt, or confusion among the audience. This contrasts with "Excite," which aims to invigorate or energize the audience. Both "Dismay" and "Neutralize" aim to undermine the opponent's message, thereby making them more compatible. Finally, "Dismay" may create an emotional state in the audience that makes them more susceptible to "Neutralize" tactics, which could aim to diminish the credibility or importance of opposing views.

In the course of our research, we implemented a thematic analysis on the provided data sets, utilizing Latent Dirichlet Allocation (LDA) to identify prevalent topics. Our topic analysis results are shown in Figures 8 and 9. As shown in Figure 8, a subset of the anti-U.S. propaganda was notably linked to terms such as "IRAN," "SYRIA," "SHIA," "ALEPPO," and "SADDAM" (translation from English to Arabic is provided in Table 9). Upon further examination of tweets containing these terms, it became evident that the associated user accounts were mainly critiquing U.S. policies in Syria, Iraq, and Yemen. A belief held by some Saudis is that the United States is unfairly favoring Iran-backed Houthi rebels in Yemen. These individuals contend that while the United States has expressed criticism of Saudi Arabia's military intervention, it has not sufficiently held the Houthis responsible for their deeds. Such perceived bias incites unfavorable attitudes toward U.S. politics. For instance, several of the tweets analyzed for this article demonstrate this precise belief. One user commented, "Americans love parasites that spread poison into countries and destroy their stability if you notice that America did not harm the Houthis, Hezbollah, and Muqtada al-Sadr." Another user commented, "The Houthis display slogans against America while it supports them and they support it! Are they America's hand at Saudi Arabia's side so that Saudi Arabia can blackmail them whenever it wants?!" State-linked accounts also have negative views about the U.S. role in Syria. Given the complex regional dynamics, some Saudis may perceive U.S. policies in Syria such as engaging the Iranian government in

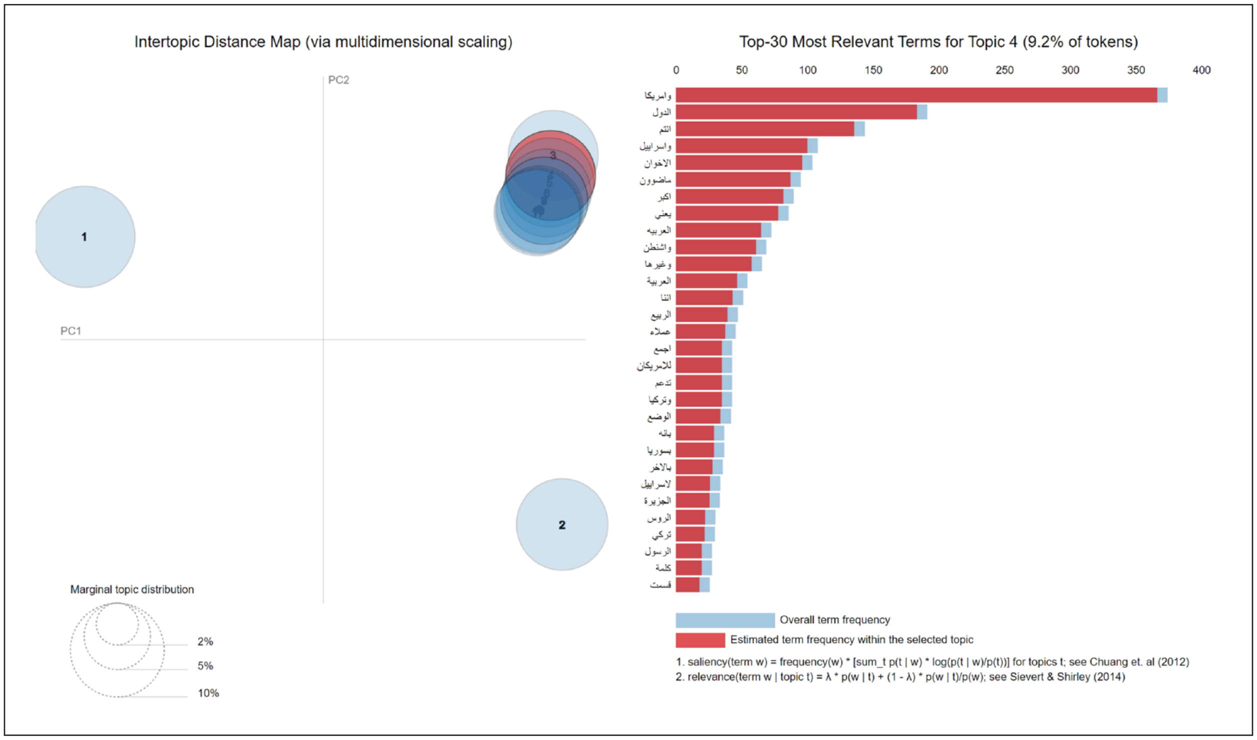


Figure 9. Topic analysis: Thematic Topic Sample 2.

Table 9. Translation of the Most Important Keywords in Figure 8.

Iran	Syria	Allah	terrorist	Saudi	ally	revolution	legend	Arabic	They conspire	Saddam	Fall	Women	Politics	Game
إيران	سوريا	الله	ارهابية	السعودي	محور	ثورة	أسطورة	عربية	يتكمرون	صدام	سقوط	النساء	السياسة	لعبة

Table 10. Translation of the Most Important Keywords in Figure 9.

USA	Israel	Muslim Brotherhood	Washington	Arab Spring	Support	Turkey	Syria	Russian	Prophet	Going on	Divided
امريكا	اسرائيل	الإخوان	واشنطن	الربيع	تدعم	تركيا	سوريا	الروس	الرسول	ماضون	قسمت

regional diplomacy over the Syrian civil war as indirectly benefiting Iran (Gause, 2016). They believe that U.S. actions have allowed Iran to expand.

As shown in Figure 9 and the translation in Table 10, we identified topics pertaining to matters such as the Muslim Brotherhood. Perceived inconsistencies in U.S. policies have led to the belief that the United States endorses the Muslim Brotherhood. For instance, in the context of the Arab Spring uprisings, the United States seemed to initially back certain Islamist groups linked to the Muslim Brotherhood. This has contributed to the impression that the United States holds a favorable stance toward this organization.

Relatedly, most of the discourse about Russia between 2019-2021 focuses on the Russian role in Syria. The role that Russia has played in the Syrian conflict, supporting the Syrian government, has made some Saudis recognize Russia's growing influence in the region and adjusted their discourse, accordingly, acknowledging the need for engagement and dialogue with Russia (Suchkov, 2016). We collected a sample of tweets in 2022 and found that most of the Saudi influencers on Twitter are supporting reforming the Saudi–Russian relationship. Figure 10 shows a sample of a topic analysis of such tweets (translation from English to Arabic is provided in Table 11). Most of the tweets show that influencers focused on "creating a balanced relationship with Russia which should have been invested in decades ago where Saudis sacrificed for the sake of Americans." Many actors believe that if the relationship with Russia had been balanced, Saudi Arabia would have a nuclear program.

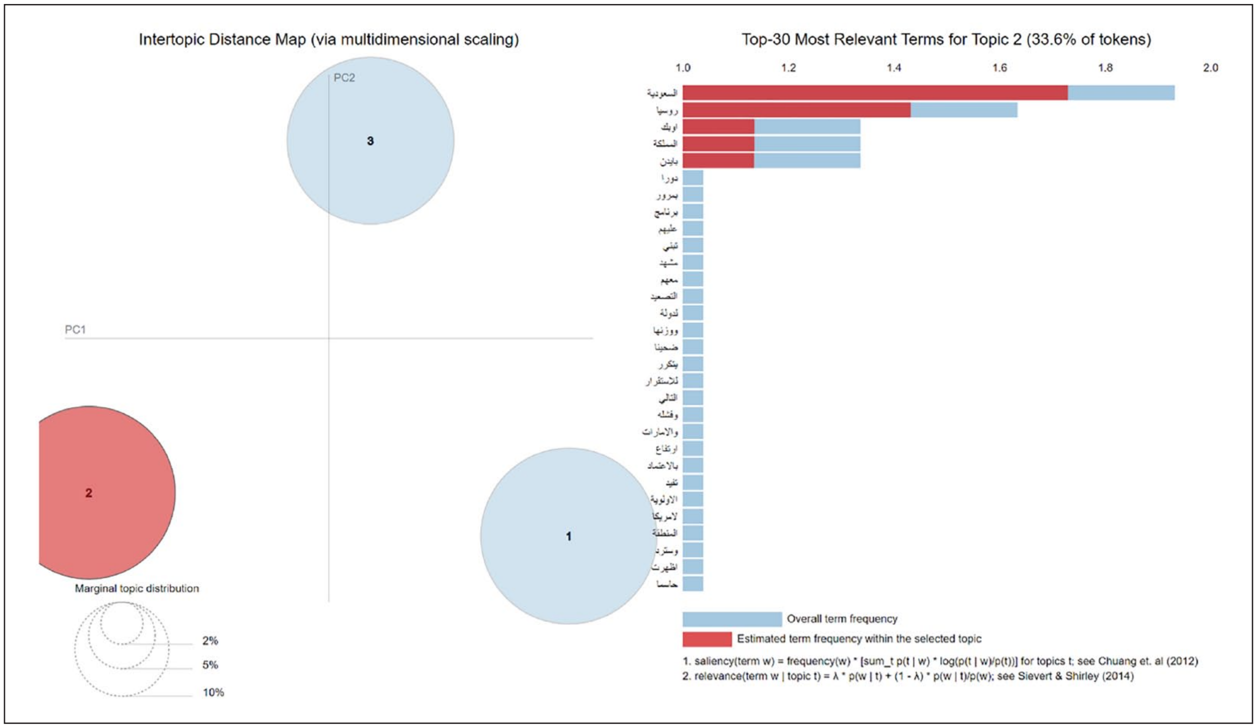


Figure 10. Topic analysis: discourse shifts toward Russia.

Table 11. Translation of the Most Important Keywords in Figure 10.

Saudi Arabia	Russia	OPEC	Biden	Program	USA	UAE	Turkey	Scene	Weight	High	definitive	sacrifice
السعودية	روسيا	اوبك	بايدن	برنامج	امريكا	الامارات	تركيا	مشهد	وزن	ارتفاع	حاسما	تضحية

Note. OPEC = Organization of the Petroleum Exporting Countries.

The analysis above shows that most of the state-linked activities in the analyzed data set are related to state-linked accounts connected to the government in Saudi Arabia. However, that does not exclude efforts by a Russian agent on social media. Those activities often include elements, including propaganda, disinformation, and targeted messaging. Russian media outlets highlight instances where the United States has taken actions contrary to Saudi interests or where policy decisions appear inconsistent. By emphasizing these instances, Russian propaganda attempts to create doubt and erode trust in the U.S.–Saudi alliance, presenting Russia as a more consistent and dependable partner.

U.S. Policy Recommendations

Any attempt from democratic regimes to engage in propaganda or counter propaganda emanating from authoritarian regimes like Saudi Arabia could be seen as dangerous to democracy itself, and it rife with potential problems. To not endanger democracy—a recurring theme right now within the United States' domestic sphere's nexus with national security (Schünemann, 2022), several policies are recommended on how to combat propaganda and influence operations that fall within democratic norms the United States and other democracies can follow. First, as Woolley (2022) notes, "Governments and other institutions working to push back against the cascade of digital falsehood . . . must be clear about the values that drive manipulation campaigns, particularly when autocrats are behind them" (p. 127). In other words, for the United States to effectively combat Saudi propaganda and influence operations more generally, it must first be transparent about what it is doing and why it is doing so.

When engaging in counter-propaganda activities, there must be given clear reasons as to why they are needed; furthermore, the reasons why, in this case, Saudi Arabia targets the United States must be explained within official settings, particularly Congressional hearings, thereby easing the democratic impulse to not engage in such operations as well. It must be made clear to the public, through these types of hearings, what are the mission-set, target vectors, strategy, motivations, and tactics behind Saudi campaigns and how the

United States will respond to them. Of course, for national security purposes, nothing classified needs to be divulged in this context. However, there should be some explanation of these campaigns in perhaps the National Cyber Strategy, or National Defense Policy. Next, and building further on Woolley, social media firms within the United States must do a better job at managing propagandic and influence operations emanating from external nation-state actors; in other words, even if the content is allowed to stay, or accounts are not blocked, it is necessary to publish these events to the public in an attempt at democratic transparency (Woolley, 2022, p. 127). Both Meta and Twitter already do this in some respects, but the data and reports do not seem to be readily available to the general public. Social media firms making these self-studies more accessible to is a step in the right direction toward democratic citizens understanding the online threat derived from propagandic influence operations.

Third, as Albert et al. (2023) argue, for the United States to counter Information Warfare and Influence Operations (IWIO), of which propaganda is a tool, perhaps the most important step is for the United States to have a whole-of-government approach. This could entail creating a new entity that handles all influence operations online, or, if it entails creating something similar to a Joint Operations Command, which would unify all entities within the government concerning doctrine, strategy, and intelligence related to countering propaganda (Albert et al., 2023). What is clear is that the United States lacks a unified doctrine for influence operations, and this puts the United States at a disadvantage, especially against more unified, autocratic regimes, especially those such as Russia and China, but theoretically, Saudi Arabia as well. As it stands, the United States is lagging behind more authoritarian adversaries in the information domain generally because of its "inability to turn data into operational intelligence and its lack of human capital allocation regarding [information warfare] IW" (Albert et al., 2023). The present research demonstrates the parameters in which propaganda operations targeting the United States exist, and thus, help set the operating principles for a whole-of-government approach the United States would need to effectively and ethically counter propaganda, within democratic norms.

Conclusion

Social media threats are becoming increasingly like conventional cyberattacks. In areas such as social engineering, attacks are targeting specific individuals or organizations. This is also true in information operations such as propaganda campaigns that target certain countries, individuals, and organizations. We collected targeted propaganda spans of anti-U.S. texts from different geopolitical contexts and showed that the general frameworks to detect those types of attacks are not effective. We show that social cyber-attack detection models need to be contextualized, meaning that if they target specific groups, countries, or organizations, model finetuning approaches may not be sufficient to identify what propaganda effects attackers are aiming to achieve. This has two implications: Theoretically, it implicates reconsidering the existing behavioral models of the social cyber-attack intentions in low resource languages such as Arabic: for example, considering an extension of the BEND Social-Cyber security framework (Carley, 2020). Practically, our research implicates the need to consider semantics and context to detect those attacks effectively.

We introduced one of the first research attempts to investigate contextualized state-backed social media attacks in the Middle East. We used general training models to detect political propaganda on U.S. personnel and institutions. Our results indicated the limitation of the general propaganda detection models to identify more targeted forms of propaganda. We recommend a possible extension of the existing classification of social cyber threats in other languages such as Arabic. We believe that there is a need for a new sociotechnical framework to detect such attacks, and the authors of this article are currently researching this. Specifically, we will create a revised BEND framework for Arabic social media. As an extension of this work, we believe that we also need to increase the size of the targeted data set as one of the limitations of this work. We also consider studying the emotional reactions of the targeted propaganda attacks. Finally, we will test our method on other Arabic-language models.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This material is based upon research supported by the U.S. Office of Naval Research under award number N000142212549.

ORCID iDs

Craig Douglas Albert https://orcid.org/0000-0003-3225-9386
Ahmed Aleroud https://orcid.org/0000-0003-4337-1488

Note

1. The implementation is an extension of the span detection implementation provided by Alhuzali and Ananiadou (2021), which can be found in the following repository: https://github.com/hasanhuz/SpanEmo. The data sets and code are available upon request to replicate experiments.

References

Abrahams, A., & Leber, A. (2021a). Comparative approaches to mis/disinformation| electronic armies or cyber knights? The sources of pro-authoritarian discourse on Middle East Twitter. *International Journal of Communication*, *15*, Article 27.

Abrahams, A., & Leber, A. (2021b). Electronic armies or cyber knights? The sources of pro-authoritarian discourse on Middle East Twitter. *International Journal of Communication*, *15*, 1173–1199.

Alam, F., Mubarak, H., Zaghouani, W., Martino, G. D. S., & Nakov, P. (2022). *Overview of the WANLP 2022 shared task on propaganda detection in Arabic*. arXiv:*2211*.10057.

Albert, C. D., Mullaney, S., Huitt, J., Hunter, L., & Snider, L. (2023). Weaponizing words: Using technology to proliferate information warfare. *Cyber Defense Review*, *8*(3), 15–31..

Alhuzali, H., & Ananiadou, S. (2021). *SpanEmo: Casting multi-label emotion classification as span-prediction*. arXiv:*2101*.10038.

Ali, S. R., & Fahmy, S. (2013). Gatekeeping and citizen journalism: The use of social media during the recent uprisings in Iran, Egypt, and Libya. *Media, War & Conflict*, *6*(1), 55–69.

Alizadeh, M., Shapiro, J. N., Buntain, C., & Tucker, J. A. (2020). Content-based features predict social media influence operations. *Science Advances*, *6*(30), Article eabb5824.

Arif, A., Stewart, L. G., & Starbird, K. (2018). Acting the part: Examining information operations within# BlackLivesMatter discourse. *Proceedings of the ACM on Human-Computer Interaction*, *2*, 1–27.

Badawy, A., Ferrara, E., & Lerman, K. (2018, August). Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign. In *2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)* (pp. 258–265). IEEE.

Barrie, C., & Siegel, A. (2021). Kingdom of trolls? Influence operations in the Saudi Twittersphere. *Journal of Quantitative Description: Digital Media*, *1*, 1173–1199.

Bastos, M., & Farkas, J. (2019). "Donald Trump is my president!" The Internet Research Agency Propaganda Machine. *Social Media + Society*, *5*(3), 1–13.

Beck, M. (2020). The aggravated struggle for regional power in the Middle East: American allies Saudi Arabia and Israel versus Iran. *Global Policy*, *11*(1), 84–92.

Beckman, J., & Nigatu, G. (2021). Do political factors influence U.S. crude oil imports? *International Journal of Energy Economics and Policy*, *11*(4), 288–297.

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Bergh, A. (2020). Understanding Influence Operations in social media. *Journal of Information Warfare*, *19*(4), 110–131.

Blane, J. T., Bellutta, D., & Carley, K. M. (2022). Social-Cyber maneuvers during the COVID-19 vaccine initial rollout: Content analysis of tweets. *Journal of Medical Internet Research*, *24*(3), Article e34040.

Blane, J. T., Ng, L. H. X., & Carley, K. M. (2023). Analyzing social-cyber maneuvers for spreading COVID-19 pro- and anti-vaccine information. In T. Ginossar, S. F. A. Shah, & D. Weiss (Eds.), *Vaccine communication online: Counteracting misinformation, rumors and lies* (pp. 57–80. Springer.

Bond, S. (2021). Just 12 People are Behind most Hoaxes on Social Media, Research Shows. NPR, https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-twitters-ability-to-curb-vaccine-hoaxes. Accessed November 1, 2023.

Brangetto, P., & Veenendaal, M. A. (2016, May). Influence cyber operations: The use of cyberattacks in support of influence operations. In *2016 8th International Conference on Cyber Conflict (CyCon)* (pp. 113–126). IEEE.

Broniatowski, D. A., Kerchner, D., Farooq, F., Huang, X., Jamison, A. M., Dredze, M., Quinn, S. Q., & Ayers, J. W. (2022). Twitter and Facebook posts about COVID-19 are less likely to spread misinformation compared to other health topics. *PLOS ONE*, *17*(1), Article e0261768.

Caldarelli, G., De Nicola, R., Del Vigna, F., Petrocchi, M., & Saracco, F. (2020). The role of bot squads in the political propaganda on Twitter. *Communications Physics*, *3*(1), Article 81.

Callanan, J. (2009). *Covert action in the Cold War: US Policy, intelligence and CIA operations*. I.B. Tauris.

Carley, K. M. (2020). Social cybersecurity: An emerging science. *Computational and Mathematical Organization Theory*, *26*(4), 365–381.

Chaudhry, I. (2014). Arab Revolutions: Breaking fear|# hashtags for change: Can Twitter generate social progress in Saudi Arabia. *International Journal of Communication*, *8*, 943–961.

Chernobrov, D., & Briant, E. L. (2022). Competing propagandas: How the United States and Russia represent mutual propaganda activities. *Politics*, *42*(3), 393–409.

Ciftci, S., & Tezcür, G. M. (2016). Soft power, religion, and anti-Americanism in the Middle East. *Foreign Policy Analysis*, *12*(3), 374–394.

Cohen, D., & Bar'el, O. (2017, October). The use of cyberwarfare in influence operations. In *Yuval Ne'eman Workshop for Science, Technology and Security*. Tel-Aviv University.

Conrad, J. (2011). Interstate rivalry and terrorism: An unprobed link. *Journal of Conflict Resolution*, *55*(4), 529–555.

Cordey, S. (2019). *Cyber influence operations: An Overview and comparative analysis*. CSS Cyber Defense Reports. https://doi.org/10.3929/ethz-b-000382358

Danaditya, A., Ng, L. H. X., & Carley, K. M. (2022). From curious hashtags to polarized effect: Profiling coordinated actions in Indonesian twitter discourse. *Social Network Analysis and Mining*, *12*(1), Article 105.

Darwich, M. (2020). Escalation in failed military interventions: Saudi and Emirati quagmires in Yemen. *Global Policy*, *11*(1), 103–112. https://doi.org/10.1111/1758-5899.12781

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Dilley, L., Welna, W., & Foster, F. (2022). *QAnon propaganda on Twitter as information warfare: Influencers, networks, and narratives*. arXiv:*2207*.05118.

DiResta, R., Goldstein, J. A., & Grossman, S. (2021). Middle East influence operations: Observations across social media takedowns. *Digital Activism and Authoritarian Adaptation in the Middle East*, 91. https://pomeps.org/wp-content/uploads/2021/08/POMEPS_Studies_43_Web.pdf#page=92

Fahmy, S., Wanta, W., & Nisbet, E. C. (2012). Mediated public diplomacy: Satellite TV news in the Arab world and perception effects. *International Communication Gazette*, *74*(8), 728–749.

Ferrara, E. (2020). *What types of COVID-19 conspiracies are populated by Twitter bots?* arXiv preprint arXiv:2004.09531.

Gause, F. G. (2016). The Future of US-Saudi relations: The kingdom and the power. *Foreign Affairs*, *95*(4), 114–126.

Guarino, S., Trino, N., Celestini, A., Chessa, A., & Riotta, G. (2020). Characterizing networks of propaganda on twitter: A case study. *Applied Network Science*, *5*(1), 1–22.

Howard, P. N., Bolsover, G., Kollanyi, B., Bradshaw, S., & Neudert, L. M. (2017). *Junk news and bots during the U.S. Election: What were Michigan voters sharing over Twitter?* (Data Memo 2017.1). Project on Computational Propaganda. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/03/What-Were-Michigan-Voters-Sharing-Over-Twitter-v2.pdf

Howard, P. N., & Kollanyi, B. (2016). *Bots, #strongerin, and #brexit: Computational propaganda during the UK-EU referendum*. arXiv:160606356.

Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics*, *15*(2), 81–93.

Jamal, A. A., Keohane, R. O., Romney, D., & Tingley, D. (2015). Anti-Americanism and anti-interventionism in Arabic Twitter discourses. *Perspectives on Politics*, *13*(1), 55–73.

Jowett, G. S., & O'Donnell, V. (2018). *Propaganda & persuasion*. Sage.

Kießling, B., Homburg, J., Drozdzynski, T., & Burkhardt, S. (2020). State propaganda on Twitter: How Iranian propaganda accounts have tried to influence the international discourse on Saudi Arabia. In C. Grimme, M. Preuss, F. W. Takes, & A. Waldherr (Eds.), *Disinformation in open online media: First multidisciplinary international symposium* (pp. 182–197). Springer.

Larson, E. V., Darilek, R. E., Gibran, D., Nichiporuk, B., Richardson, A., Schwartz, L. H., & Thurston, C. Q. (2009). *Foundations of effective influence operations: A framework for enhancing army capabilities* (Vol. 654). Rand Corporation.

Llewellyn, C., Cram, L., Favero, A., & Hill, R. L. (2018, May). Russian troll hunting in a Brexit Twitter archive. In *Proceedings of the 18th ACM/IEEE on joint conference on digital libraries* (pp. 361–362). ACM.

Long, D. E. (2019). *The United States and Saudi Arabia: Ambivalent allies*. Routledge.

Lukito, J. (2020). Coordinating a multi-platform disinformation campaign: Internet research agency activity on three U.S. social media platforms, 2015 to 2017. *Political Communication*, *37*(2), 238–255.

Mair, D. (2016). # Westgate: A case study–how Al-Shabaab used Twitter during an ongoing attack." *Studies in Conflict & Terrorism*, *40*(1), 24–43.

Maschmeyer, L., Abrahams, A., Pomerantsev, P., & Yermolenko, V. (2023). Donetsk don't tell–'hybrid war'in Ukraine and the limits of social media influence operations. *Journal of Information Technology & Politics*, 1–16. https://doi.org/10.1080/19331681.2023.2211969

McNeill, L. (2019). Assumed Identity: Writing and reading testimony through and as Anne Frank 1. In J. R. Resina (Ed.), *Inscribed identities* (pp. 61–74). Routledge.

Mendelson, S. E., & Gerber, T. P. (2008). Us and them: Anti-American views of the Putin generation. *Washington Quarterly*, *31*(2), 131–150.

Miller, D. T. (2019). Topics and emotions in Russian Twitter propaganda. *First Monday*, *24*(5). https://doi.org/10.5210/fm.v24i5.9638

Mitts, T. (2019). From isolation to radicalization: Anti-Muslim hostility and support for ISIS in the West. *American Political Science Review*, *113*(1), 173–194.

Mohaddes, K., & Pesaran, M. H. (2016). Country-specific oil supply shocks and the global economy: A counterfactual analysis. *Energy Economics*, *59*, 382–399.

Moriarty, B. (2015). Defeating ISIS on Twitter. *Technology Science*. https://techscience.org/a/2015092904/

Nakayama, B. (2022). Democracies and the Future of Offensive (Cyber-Enabled). *Information Operations*, *7*(3), 49–65.

Neudert, L. M. N. (2018) Germany: A cautionary tale. In S. C. Woolley & P. N. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.

Neumayer, E., & Plümper, T. (2011). Foreign terror on Americans. *Journal of Peace Research*, *48*(1), 3–17.

Ng, L. H. X., & Carley, K. M. (2023). Popping the hood on Chinese balloons: Examining the discourse between U.S. and China-geotagged accounts. *First Monday*, *28*(8). https://doi.org/10.5210/fm.v28i8.13159

Ng, L. H. X., Moffitt, J. D., & Carley, K. M. (2022). *Coordinated through a web of images: Analysis of image-based influence operations from China, Iran, Russia, and Venezuela*. arXiv:2206.03576.

Nip, J., & Sun, C. (2022). Public diplomacy, propaganda, or what? China's communication practices in the South China Sea dispute on Twitter. *Journal of Public Diplomacy*, *2*(1), 43–68. https://doi.org/10.23045/jpd.2022.2.1.4

Nogara, G., Vishnuprasad, P. S., Cardoso, F., Ayoub, O., Giordano, S., & Luceri, L. (2022, June). The disinformation dozen: An exploratory analysis of covid-19 disinformation proliferation on twitter. In *Proceedings of the 14th ACM Web Science Conference 2022* (pp. 348–358). ACM.

Paszke, A., Gross, S., Chintala, S., Chanan, G., & Yang, E. (2017). *Automatic differentiation in PyTorch* [Conference session]. 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, United States.

Patrikarakos, D. (2017). *War in 140 characters: How social media is reshaping conflict in the twenty-first century*. Hachette UK.

Pierri, F., Luceri, L., Jindal, N., & Ferrara, E. (2023, April). Propaganda and misinformation on Facebook and Twitter during the Russian invasion of Ukraine. In *Proceedings of the 15Th ACM Web Science Conference, 2023* (pp. 65–74). ACM.

Prier, J. (2017). Commanding the trend: Social media as information warfare. *Strategic Studies Quarterly*, *11*(4), 50–85.

Safaya, A., Abdullatif, M., & Yuret, D. (2020). Kuisail at Semeval-2020 task 12: BERT-CNN for offensive speech identification in social media. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 2054–2059). International Committee for Computational Linguistics.

Schünemann, W. J. (2022). A threat to democracies? An overview of theoretical approaches and empirical measurements for studying the effects of disinformation. In M. D. Cavelty & A. Wenger (Eds.), *Cyber security politics* (pp. 32–47). Routledge.

Shakil, K. A., Tabassum, K., Alqahtani, F. S., & Wani, M. A. (2021). Analyzing user digital emotions from a holy versus non-pilgrimage city in Saudi Arabia on twitter platform. *Applied Sciences*, *11*(15), 6846.

Simon, T., Goldberg, A., Aharonson-Daniel, L., Leykin, D., & Adini, B. (2014). Twitter in the cross fire—The use of social media in the Westgate Mall terror attack in Kenya. *PLOS ONE*, *9*(8), Article e104136.

Singer, P. W., & Brooking, E. T. (2018) *LikeWar: The weaponization of social media*. Eamon Dolan Books.

Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction*, *3*, 1–26.

Suchkov, M. (2016). Contemporary Russia-Saudi relations: Building a bridge of cooperation over the abyss of discrepancies. *Iran and the Caucasus*, *20*(2), 237–251.

Sullivan, R. (2014). Live-tweeting terror: A rhetorical analysis of@ HSMPress_ Twitter updates during the 2013 Nairobi hostage crisis. *Critical Studies on Terrorism*, *7*(3), 422–433.

Theohary, C. A. (2018). Information Warfare: Issues for Congress. *Congressional Research Service Report*.

Uyheng, J., Magelinski, T., Villa-Cox, R., Sowa, C., & Carley, K. M. (2020). Interoperable pipelines for social cyber-security: Assessing Twitter information operations during NATO Trident Juncture 2018. *Computational and Mathematical Organization Theory*, *26*, 465–483.

Van Dijk, T. A. (Ed.) (2011). *Discourse studies: A multidisciplinary introduction*. Sage.

Weimann, G. (2010). Terror on Facebook, Twitter, and YouTube. *The Brown Journal of World Affairs*, *16*(2), 45–54.

Weiss, J. C. (2013). Authoritarian signaling, mass audiences, and nationalist protest in China. *International Organization*, *67*(1), 1–35.

Whyte, C., Thrall, A. T., & Mazanec, B. M. (Eds). (2020). *Information warfare in the age of cyber conflict*. Routledge Press.

Wolf, H. (2015). Paper is patient': Tweets from the '# AnneFrank of Palestine. *Textual Practice*, *29*(7), 1355–1374.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., & Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations* (pp. 38–45).

Woolley, S. C. (2022). Digital propaganda: The power of influencers. *Journal of Democracy*, *33*(3), 115–129.

Yeh, C. K., Wu, W. C., Ko, W. J., & Wang, Y. C. F. (2017, February). Learning deep latent space for multi-label classification. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1, pp. 2838–2844). AAAI Press.

Author Biographies

Craig Douglas Albert, PhD, is professor and director of the Master of Arts in Intelligence and Security Studies at Augusta University. His main areas of research include strategic cybersecurity, artificial intelligence, influence operations and information warfare, and Ethnic Conflict, with a specific interest in Chechnya. He has published in *Cyber Defense Review*, *Journal of Cyber Policy, Politics and the Life Sciences*, *Intelligence and National Security*, *Defense & Security Analysis*, and *Global Society*. His work has been supported by NSF and Office of Naval Research.

Ahmed Aleroud, PhD, is an associate professor in the School of Computer and Cyber Sciences. His research work focuses on machine learning computational approaches and social approaches to combat security and privacy attacks, including social engineering attacks, computer network intrusions, disinformation, social media propaganda, state-linked social media attacks, online extremism, and re-identification of private information on individuals or devices. He is also working on adversarial attacks on security systems. His research has been published in journals such as *IEEE Transactions*, *ACM Digital Threats Research & Practice*, *Computers & Security*, *Information Security and Applications*, *Information Systems Frontiers*, *Knowledge and Information Systems*. His work has been supported by NSF, Office of Naval Research, European IP Networks (RIPE), and the Department of Energy.

Yufan Yang, PhD, is an assistant professor of political science at Augusta University. Her research interests include international security, political violence, political regimes, propaganda, information technology, statistical modeling, machine learning, and computational social science in general.

Abdullah Melhem is a graduate research assistant at Augusta University's School of Computer and Cyber Sciences. With a master's degree in data science from Jordan University of Science and Technology, Abd has a strong background in advanced data analytics. Previously, he worked as a data scientist at Jordan Design and Development Bureau (JODDB) in Jordan.

Josh Rutland is a graduate of Augusta University's Masters in Intelligence and Security Studies program. His research focuses on information warfare, cybersecurity, terrorism, and biosecurity. His work has appeared in journals such as *Politics and the Life Sciences*, *Politics & Policy*, *Behavioral Sciences of Terrorism and Political Aggression*, *Journal of Cyber Policy*, *The Cyber Defense Review*, *and PLOS Global Public Health.*

A Military of Influencers: The U.S. Army Social Media, and Winning Narrative Conflicts

Lieutenant Colonel Robert J. Ross, Ph.D.
Josh Rutland

ABSTRACT

In the interconnected era of the Internet, the military must confront the new face of an old threat: narrative conflict. Where states once maintained nearly absolute domestic control of the narratives surrounding their military engagements, social media have created a wide array of perspectives, arguments, and disinformation campaigns that constantly affect both the civilian and military populations. These campaigns encourage the questioning of state objectives and threaten the identity of the individual and the collective ontological identity of the society, making it more difficult for states to maintain momentum and support for their military endeavors. Without that support, military campaigns can collapse, regardless of the skill or preparedness of warfighters. This research explores three topics relevant to the U.S. Army in hopes of helping it better equip itself to succeed in narrative conflicts: the strategic impacts of commander's decisions on the battlefield, the need to control signals emissions, and the consequences of bulk internet data sales. It then concludes by providing brief policy suggestions for mitigating these issues.

INTRODUCTION

When the Gutenberg printing press emerged in the late 15th century, it rocked the foundations of societal order in Europe by establishing the first networked era.[1] The ensuing mass production of pamphlets made them accessible to the common person.[2] As the masses of common Europeans began to study religious texts for themselves, new perspectives emerged to challenge the church's

Lieutenant Colonel Robert J. Ross is the Strategic Initiatives Group Chief for the Commanding General of U.S. Army Cyber Command, Fort Gordon, GA. Lieutenant Colonel Ross advises the ARCYBER Commanding General on cybersecurity, information-age conflict, and information warfare strategy. He is a former assistant professor in the Electrical Engineering and Computer Science Department at the U.S. Military Academy at West Point, NY. He is a former Chief Research Scientist for the Army Cyber Institute, a position in which he served as the Information Warfare Team Lead. He has a B.S. degree in Computer Science from Rowan University, an M.S. in Computer Science from Monmouth University, and a Ph.D. in Information Science from the Naval Postgraduate School. Lieutenant Colonel Ross is a cyberwarfare officer and former artilleryman with two combat deployments to Iraq. His research interests are organizational science, strategic foresight, information warfare, 21st century conflict, and financial technology.

authority.[3] Ultimately, the increasingly rapid dissemination of information through advancing technology caused the questioning, undermining, and weakening of the authority of the Roman Catholic Church, which had dominated religious narratives in Europe for more than 1,000 years.

Technology has continued to grow in modern times, with mobile phones and the Internet creating a network of instantaneous communications much, much larger in scope than that of Gutenberg's printing press. The growing technology has amplified impacts on society, with conventional authorities facing unprecedented challenges to their leadership. The time has passed for state control over the information flow across and within its borders using traditional media methods, and official narratives that shaped public opinion in support of the state. Political and ideological dissonance quickly and ubiquitously pours across the borderless Internet from which the global audience drinks.[4] Blog posts, cell phone footage, podcasts, drone recordings, and myriad other content forms are deemed valid regardless of merit or origin.[5] Collectively, they form the new narratives consumed and further propagated by the masses on social media. The result is, once again, a questioning of conventional authority and the degradation of that authority's power at an unprecedented rate. The walls of Westphalia have fallen again.

These developments have troubling implications for contemporary warfighting scenarios, which require a motivated military and citizenry for victory. While traditional military conflict continues, as in the Russian invasion of Ukraine, and remains a critical component of warfare, the importance of narrative conflict has never been greater. The Internet, mobile phones, and social media offer an opportunity for states to infiltrate the minds of their adversaries' citizenry through

Josh Rutland is a graduate of Augusta University's Master in Intelligence and Security Studies program. He currently works as a researcher in the Augusta University Department of Emergency Medicine and will soon be employed by ARCYBER as an Information Technology Specialist. His research focuses on information warfare, cybersecurity, terrorism, and biosecurity. His work has appeared in such journals as *Politics and the Life Sciences, Politics & Policy, Behavioral Sciences of Terrorism and Political Aggression, Journal of Cyber Policy,* and *PLOS Global Public Health*.

widespread, tailored propaganda efforts. These efforts may be designed to facilitate a variety of outcomes, including diminished support for a war. Demoralization on such a wide scale threatens to "rob an army of its spirit and a commander of his courage," which Sun Tzu described as the key to victory, destroying an adversary's will to fight without so much as a single battle.[6] The Islamic State of Iraq and Syria's (ISIS) victory over the Iraqi Army at Mosul provides a potent example of the power of narrative: The 10,000 troops present in Mosul had mostly abandoned their posts out of fear spawned by ISIS terror campaigns that streamed across the Internet long before ISIS forces arrived in the city.[7] The result was an easy victory for ISIS forces. Though the Iraqi force was larger and better armed, its fear of ISIS ultimately ensured its defeat.[8] Even the US has fallen prey to the effects of narrative defeat during the wars in Vietnam, Iraq, and Afghanistan. Wars can be won or lost based on their surrounding narratives despite overwhelming tactical victory in every engagement using traditional military force.

This article argues that winning modern narrative conflicts will demand doctrinal change within the Army and other services in some key areas relevant to information operations, public affairs, psychological operations, and cyber space operations. The study focuses on three important issue areas: the strategic impacts of soldiers' decision-making, vulnerabilities related to signature management, and the threats posed by bulk data collection and sales conducted by third party social media platforms. To demonstrate this point, the article proceeds in two sections. First, we briefly analyze the three focal issue areas using existing literature that highlights their importance and details the security issues. Then, we provide relevant policy suggestions, based on modifications of former and existing Army doctrine generated from researching this topic.

THE STRATEGIC CORPORAL

U.S. Marine General Charles Krulak conceptualized the strategic actions at the lowest tactical level in his 1999 essay titled "The Strategic Corporal." Krulak argued that "success or failure will rest, increasingly, with the rifleman and with his ability to make the *right* decision at the *right* time at the point of contact" with both the enemy and the local population.[9] In addition to the pressures of high stress environments where lives are at stake; the soldier in the field also bears the burden of overcoming two major obstacles: a general hostility and weariness on the part of the local population and the mutually perceived cultural divisions between one's own "ingroup" and the "outgroup" that inhibit communication and personal bonding.[10] While this places additional demand on warfighters, their ability to understand and adopt relevant customs and behaviors of the indigenous populations with which they interact will shape their own personal relationships within that society and the general disposition of that society toward other warfighters with whom they interact in the future.[11]

As such, the ability of Army warfighters also to function and be perceived as "cultural mediators" and community members when interacting with a foreign populace is a critical tool that must be maintained like any other piece of equipment in a soldier's toolkit.[12] This has led to calls for redesigned professional military education processes that highlight the importance of language training, cultural education, and "educational and experiential cross-fertilization between the military and other government agencies" or humanitarian organizations relevant to future operational fields.[13] Major Linda Liddy of the Australian Army also argues that the modern soldier will need to be academically savvy in topics such as "military law and leadership, military history, and current affairs and ethics" in order to prepare fully for their role as warfighters and influencers expected to carry out complex operations with military and humanitarian ambitions.[14]

The omnipresence of cell phones with cameras and Internet connectivity further ensures that tactical-level actions, positive or negative, will ripple across the societies with which they interact and extend beyond their immediate communities.[15] Strategic adversaries could coopt footage depicting cultural insensitivity, whether accidental or deliberate, to fuel terrorist recruitment[16] or turn large populations and Internet communities against the U.S. Army. This could diminish its security, morale, and chances of operational success.[17] Warfighters must do everything in their power to set themselves apart in the minds of those with whom they interact in operational theaters to establish mutual respect, cooperation, and beneficence.[18] A warfighter has a personal presence in the minds of those with whom they interact. This means that the warfighter ceases to be simply an American or a soldier to become a friend or community member, which can be critical in environments such as the Middle East where cultural and familial bonds mean far more than shared regional or territorial residency. In short, impressions and reputations are critical; they can make or break an operation tactically and narratively.[19] Warfighters will need to be able to shape their reputations in a positive way to ensure operational success.

SIGNATURE MANAGEMENT VULNERABILITIES

Signature management vulnerabilities are those associated with the impacts on battlefield events or troop deployments of signals emitted and received from electronic devices.[20] While myriad strategic vulnerabilities exists with respect to signature management and narrative conflict, two significant threats stem from physical infrastructure and "digital exhaust,"[21] which is described by Harper Reed as "a constant trail of activities, behaviors, preferences, signatures, and connections" left behind by every digital device that is tied to both that device and its user.[22] Both have the potential to contribute to adversaries' interception of sensitive information regarding Army units, ultimately resulting in "the design and development of adversary systems, tactics, training, and force preparations capable of countering Army unit capabilities, activities, and intentions."[23] As such, new considerations must be accounted for to limit public knowledge of Army units and their deployments successfully.

Control of physical infrastructure means control over and access to any signals that pass through it.[24] Army units thus cannot be sure their communications are secured when operating in a foreign theater where critical infrastructure is built, owned, and operated by potentially adversarial forces.[25] As the world transitions to 5G technology, this becomes an even greater risk, as 5G infrastructure is being built primarily by China across parts of Asia, Africa, and Europe.[26] This gives China "access to the private data of billions of people" which may include "individuals' medical histories, spending habits, political views, personal details expressed on social media, physical location, financial situation" and much other data the state could adapt to "gain a commercial or technical advantage in data-driven markets, target key individuals for recruitment by intelligence operations, or compromise political figures."[27] Civilians are not the only potential target of this type of data collection. Anyone using the network is vulnerable.[28] As such, it is imperative that the Army anticipate this battlefield vulnerability and develop alternatives to using foreign infrastructure, such as establishing its own permanent infrastructure in contested regions of influence.[29]

While physical infrastructure poses a significant vulnerability, digital exhaust may represent the most significant threat associated with signature management. Digital exhaust refers to the impact on the virtual realm resulting from military movements and engagements.[30] Adversaries could use this information to determine troop movements before they are made public, putting warfighters in harm's way, risking operational failure, and presenting adversaries with an opportunity to humiliate or propagandize against their opponent. The Bellingcat Study, the Second Nagorno-Karabakh War, and recent events in Ukraine all represent examples of how dangerous digital exhaust can be in the wrong hands.

During the Bellingcat Study, a handful of amateur Internet sleuths crowdsourced information largely comprised of the Russian military's digital exhaust to provide decisive evidence that Russian forces had shot down Malaysian Airlines Flight MH17 in July 2014.[31] The Bellingcat group, led by Eliot Higgins, used online videos and photographs to identify the specific Buk anti-aircraft missile that had shot down MH17.[32] It then collected a number of videos and photographs of the Buk that enabled it to plot successfully a timeline and geographic trail of its movements from Russia into Ukraine which proved Russia's culpability in MH17's destruction.[33] The discovery forced Russia into a losing battle with the Bellingcat group to control the narrative surrounding the MH17 incident that ultimately resulted in the Russian government's embarrassment.[34] The Bellingcat group harnessed the power of social media to expose a global power and its army.[35] Anyone with a vested interest, state-or civilian-sponsored, could employ Bellingcat's methods against any army, should that army fail to account for its troops' digital exhaust.

The Second Nagorno-Karabakh War between Armenia and Azerbaijan represents a more direct example of digital exhaust exploitation by one state against another.[36] Using Turkish Bayraktar TB2 drones and Israeli HAROP Loitering Munitions (LM) , Azerbaijani forces devastated the Russian-supported Armenian ground forces through the nearly exclusive use of unmanned strikes.[37] The cameras inside these drones captured live footage of the bombing and the destruction from the strikes, which was then broadcast to both sides by the Azerbaijanis for propaganda purposes.[38] The result was an invigorated war effort by Azerbaijan and a gravely deteriorating Armenian will to fight through the constant reliving of events and fear of unexpected future drone strikes.[39] The kinetic effects of drone strikes are lost lives and destroyed equipment, already damaging to the morale of a targeted belligerent. However, the ability of the drones to capture live full-motion video (FMV) and immediately broadcast this footage to online social media forums create powerful synergies between the kinetic and cognitive effects of unmanned aerial systems. Effects from these unmanned aerial systems cause both physical and psychological deterioration of their intended prey. Azerbaijan used FMV footage to amplify wisely what could be classified as the highly survivable kinetic effects of these weapons. Eventually, the Armenian war effort was crippled after a series of defeats displayed TB2 drones "literally flying circles near three S-300 sites while waiting to strike their targets before doing damage assessment and flying away," forcing the Armenian Army to capitulate rapidly.[40]

The Russian-backed Armenian Army was powerless to counter the effects of these Turkish and Israeli unmanned aerial systems (UAS). Ukraine, with which Russia has been in direct conflict since 2014, noticed this.[41] In September 2021, Ukraine acquired 24 TB2 drones from Turkey to bolster its own efforts against Russia after observing their effectiveness in the Second Nagorno-Karabakh War.[42] The following month, the Ukrainians deployed the TB2s against Russian-backed separatists in Crimea for the first time, damaging a 122mm D-30 howitzer in the Donbass region that had previously injured one Ukrainian soldier and killed another.[43] The Ukrainians followed the Azerbaijan example by using the onboard camera systems to collect and distribute footage of the air strike online.[44] The Ukrainian Army employed this capability with continuously devastating effect after the Russian invasion in February 2022. While this is the most recent example of lessons from the Second Nagorno-Karabakh War's proliferation, it likely represents an early look at how future wars may be fought.[45] This deadly combination of conventional weaponry and narrative shaping tools represents a dangerous threat for states that fail to develop methods for controlling digital exhaust such as drone footage of engagements, especially battlefield losses.

BULK DATA COLLECTION AND SALES

Bulk data collection refers to the mass collection of personal data gathered by social media companies and other website managers.[46] As users browse websites and services that require them to accept "informed consent" agreements coupled with the proliferation of Internet of things (IoT) devices when creating or linking personal accounts, the providers and creators of these services collect bulk data from their browsing patterns.[47] Two types of research typically employ these data: academic and marketing.[48] Marketing research practices in particular represent the greatest threat from bulk data collection, as this type of research usually involves the construction of personalized profiles of each individual user to monitor and record that person's likes, dislikes, interests, purchases, media preferences, and a variety of other traits.[49] While almost all web browsing generates bulk data, social media websites represent the prime collection ground for these data as they offer a look into not only a person's preferences, but also who they associate with, social movements with which they identify , and their personal beliefs.

This process, defined as "microtargeting" by MAJ Jessica Dawson,[50] represents a gold mine from a marketing perspective, as companies can use these data to construct carefully tailored advertisement intended to lure consumers into viewing and purchasing their products. However, from a security standpoint, microtargeting represents a potential narrative nightmare, as it offers anyone with access to this detailed profile information a roadmap for how best to propagandize messages in a way that will convince its target audience to adopt a desired perspective.[51] The Cambridge Analytica case demonstrates the potential for influencing operations based on the "digital exhaust" of users in the form of bulk data intentionally used to microtarget for the purpose of influencing "likely voter" decisions.[52] Both civilians and military personnel are vulnerable to microtargeting practices regardless of their social media use because, "even if an individual does not have a Facebook account, Facebook has a shadow account for them, collected from friends' phones, contacts lists, and emails as well as data Facebook itself purchases."[53] Usually, the only significant barrier to accessing these data is a licensing fee, meaning foreign adversaries can easily acquire them for nefarious purposes.

The same adversaries may also be able to amplify their microtargeted messages to a large audience of military personnel and civilians using "a relatively novel and increasingly dangerous means of persuasion within social media," which Lt Col Jarred Prier calls "commanding the trend."[54] This method involves using bot-driven, falsified swarms of activity or "views" to manipulate the algorithms that social media sites use to "analyze words, phrases, or hashtags to create a list of topics sorted in order of popularity."[55] This activity swarm increases a page's visibility and its likelihood of being clicked and shared by convincing social media algorithms that a topic is growing in popularity, prompting the algorithm to promote it on trending pages.[56] Algorithms do not verify the authenticity of stories before promoting them, nor do they verify the credibility of the users who share them. While some companies have begun modifying their algorithms and attempting to find countermeasures to bot swarms, the reality remains that by the time a topic has reached the trending page it has already spread beyond containment.[57] Narratives promoted in this manner that are harmful to Army interests could prove dangerous and impossible to control.

POLICY SUGGESTIONS

As Joint Chiefs of Staff Chairman General Mark Milley has argued, strategic competitors' increasing capability to "fight the US through multiple layers of stand-off in all domains" means that a "doctrinal evolution" of the American way of war is necessary.[58] The lessons demonstrated in conflicts in Ukraine, Iraq, and Armenia suggest that narrative victory is growing in importance and a continual trend in the future.[59] The doctrinal adjustments necessary for the U.S. Army to fortify itself properly for this changing dynamic of warfare will likely be complex and take time to implement, but they will be essential to victory in future conflicts. The Army is probably the greatest modern conventional warfighting force, but it will need to bolster its ability to shape narratives surrounding conflicts in which it becomes involved to ensure that its conventional victories translate into strategic success.

The modern soldier must become conscious of his or her role as General Krulak's "strategic corporal," straddling the line between warfighter and diplomat.[60] In addition to combat capabilities, a soldier must be well-trained for decision making, problem solving, and positive cultural interaction.[61] Soldiers must be prepared for the eyes of the world on social media to scrutinize any and every action they take. The fate of Army morale and its reputation in the global court of public opinion hinges on the individual warfighter's ability to project a positive image of the Army to further the nation's strategic objectives. It is worth emphasizing that this does not represent a call for any lowering in priority of traditional combat skills and training; it is rather a call to elevate the importance of cultural and linguistic training as well as social media literacy.[62] Basing warfighter evaluations on both combat ability and social skills represents one way of honing these skills among Army personnel.

Signature management vulnerabilities present significant risks to operational security (OPSEC). Improving signals management strategies has been identified as a crucial step in advancing the U.S. Marine Corps' (USMC) contemporary warfighting capabilities for future conflicts.[63] The Army should afford signature management the same importance. Addressing these risks demands a prompt solution to the problems of physical infrastructure and the digital exhaust of personnel. The primary threat in the physical domain comes from China's 5G-infrastructure proliferation through its Belt and Road Initiative.[64] Using NATO as a "forum for collaboration" and expansion of US owned and operated 5G infrastructure is an optimal potential solution.[65] While this initiative will likely require significant investment in 5G-technology development and construction, the US could employ these technologies and their distribution as a diplomatic tool for strengthening relationships with existing allies or building new relationships with potential strategic partners. The Army and NATO operations in allied regions would also enjoy the benefits of US owned 5G systems: safe, trusted, and secure communications technology that would fully support the OPSEC of US joint and coalition forces.

Digital exhaust control may be more difficult to accomplish. The Army's ban on the use of personal communication devices on the battlefield is a constructive step, as it helps prevent the possibility of telecommunications interception, movement tracking using mobile device signals, and exposure to enemy disinformation that might demoralize or misinform soldiers.[66] Taking steps to mask deployment information such as supply purchases that may leave physical or digital paper trails should also be a priority. Purchasing supplies through a third-party or "middle-man," buying supplies in smaller quantities rather than in bulk and sending supplies to deployment zones with warfighters rather than shipping them directly there separately all represent potential solutions. To address online propaganda campaigns such as those seen in the second Nagorno-Karabakh War and Ukraine, the Army might consider using trend hijacking techniques such as bot swarming, as detailed by Prier, to bury adversaries' social media campaigns.[67] The Army needs to develop tactics, techniques, and procedures (TTPs) that mirror the informational effects demonstrated by the TB2 Bayraktar's successes in both the second Nagorno-Karabakh and Ukraine conflicts. TTPs that enhance the synergies between powerful kinetic and psychological effects stemming from these platforms. Furthermore, worth considering is the recruitment of existing social media influencers to help promote the Army's narratives, encouraging warfighters who demonstrate social media proficiency to become a new breed of battlefield correspondent, or the establishment of a U.S. Information Agency similar to the one created by President Eisenhower in 1953 to address US influence strategy during the Cold War. Israel's efforts to recruit young, tech-savvy, female social media operatives from existing Israeli Defense Force (IDF) units represents a notable success in this area.[68]

The Army's approach to bulk data sales and collection must respect the limitations put in place by the Fourth Amendment to the U.S. Constitution. For this reason, direct collection of data on American citizens for the purpose of microtargeted narrative construction is not a possibility. Rather, as MAJ Dawson[69] suggests, it may be useful for the Army to establish limits on data collection through cooperation with social media companies. The prevention of data collection from accounts owned by service members and their families represents a good starting point.[70] The encouragement of more stringent limits on obtaining these data from social media companies and the permitted uses of the data also represents a potential point of collaboration between the Army and social media corporations.

CONCLUSION

As the U.S. Army prepares for future conflicts, it becomes increasingly critical to consider the demonstrations of narrative power from the past and those unfolding in the present day. Winning future conflicts will mean winning narrative conflicts. To do that, the Army needs to adopt appropriate doctrinal changes related to information operations, public affairs, and cyber space operations. Tactical actions will shape strategic success, which emphasizes the need to train and equip warfighters as ambassadors of the Army's intentions and good will. Words, tweets, TikToks, Instagram posts, drone recordings, and any other microtarget-enabling media deemed "view-worthy" are the weapons of narrative conflicts. The Army must learn to leverage these weapons and deny them to strategic adversaries. This means limiting digital exhaust, cooperating with social media companies to undermine adversaries' ability to target US warfighters and citizens, and establishing a comprehensive public relations arm of the Army to promote its narratives on the ideological battleground. As conflict evolves, so too must the warfighter. It is time to forge an Army of influencers.

DISCLAIMER

The views and opinions expressed in this paper are those of the author alone and do not reflect the official policy or position of the U.S. Department of Defense (DoD), U.S. Cyber Command, or any agency of the U.S. Government. Any appearance of DoD visual information or reference to its entities herein does not imply or constitute DoD endorsement of this authored work, means of delivery, publication, transmission, or broadcast.

NOTES

1. Niall Ferguson, *The Square and the Tower: Networks and Power, from the Freemasons to Facebook*, Penguin Books (2019).
2. Ibid.
3. Ibid.
4. David Patrikarakos, *War in 140 Characters: How Social Media is Reshaping Conflict in the Twenty-First Century*, Hachette UK, (2017).
5. Ibid.
6. Sun Tzu, *The Art of War,* translated by Sammeul Griffith, Duncan Baird (2005, original work published 5th century BC), 108.
7. Peter warren Singer and Emerson T. Brooking, *LikeWar: The Weaponization of Social Media*, Eamon Dolan Books (2018).
8. Ibid.
9. General Charles C. Krulak, T*he Strategic Corporal: Leadership in the Three Block War*, Center for Army Lessons Learned Fort Leavenworth KS Virtual Research Library (1999), 5.
10. Stephen Bochner, “The Social Psychology of Cross-Cultural Relations,” in *Culture in Contact: Studies in Cross-Cultural Interaction*, edited by Stephen Bochner, Volume 1, Oxford: Pergamon (1982), 14.
11. Bochner, “The Social Psychology of Cross-Cultural Relations,” 13; Krulak, *The Strategic Corporal.*
12. Bochner, “The Social Psychology of Cross-Cultural Relations,” 15.
13. Kevin D. Stringer, “Educating the Strategic Corporal: A Paradigm Shift,” *Joint, Interagency, Intergovernmental, and Multinational Newsletter* (2011), 65; Major Linda Liddy, “The Strategic Corporal: Some Requirements in Training and Education,” *Australian Army Journal* 2, no. 2 (2004), 139-148.
14. Liddy, “The Strategic Corporal: Some Requirements in Training and Education,” 142.
15. Patrikarakos, *War in 140 Characters.*
16. Charlie Winter, *Media Jihad: Islamic State’s Doctrine for Information Warfare*, London, UK, International Centre for the Study of Radicalisation and Political Violence (2017).
17. Patrikarakos, *War in 140 Characters.*
18. Bochner, “The Social Psychology of Cross-Cultural Relations,” 13; LTC Robert J. Ross, Creating White Space: Interaction and the Adaptation of Team Social Identity in Complex Environments, Naval Postgraduate School Monterey, Monterey CA (2019), 19.
19. Krulak, The Strategic Corporal.
20. Brett van Niekerk and M. S. Maharaj, “Mobile Devices and the Military: Useful Tool or Significant Threat?,” *Journal of Information Warfare* 11, no. 2 (2012), 1-11.
21. Josh Rutland “A Military of Influencers: The U.S. Army, Social Media, and Winning Narrative Conflicts” [unpublished master’s thesis], Augusta University.
22. Brian David Johnson, Alida Draudt, Jason C. Brown, LTC Robert J. Ross, Ph. D., “Information Warfare and the Future of Conflict,” produced by Cyndi Coon, The 2019 Threatcasting Workshop, Arizona State University (2019), 68.
23. U.S. Army, *U.S. Army Techniques Publication 3-13.3: Army Operations Security for Division and Below*, Headquarters, Department of the Army (2019), 1-1.
24. Luiz A. Dasilva, Jeffrey H. Reed, Sachin Shetty, Jerry Park, Duminda Wijeskera, and Haining Wang, “Securing 5G: NATO’s Role in Collaborative Risk Assessment and Mitigation,” *Cyber Threats and NATO 2030: Horizon Scanning and Analysis* (2020), 74-87.
25. Carolyn Bartholomew, “China and 5G,” *Issues in Science and Technology* 36, no. 3 (2020), 50-57.
26. Ibid.
27. Ibid., 52-52.
28. Dasilva et al., “Securing 5G”.
29. Ibid., 85.
30. Johnson et al., “Information Warfare and the Future of Conflict”; Rutland, “A Military of Influencers.”
31. Matt Sienkewicz, "Open BUK: Digital labor, media investigation and the downing of MH17,” *Critical Studies in Media Communication* 32, no. 3 (2015), 208-223; Patrikarakos, *War in 140 Characters.*

NOTES

32. Patrikarakos, War in 140 Characters, 181.
33. Ibid.
34. Ibid. Eliot Higgins, We Are Bellingcat: Global Crime, Online Sleuths, And The Bold Future Of News, Bloomsbury Publishing (2021).
35. Higgins, We Are Bellingcat: Global Crime, Online Sleuths, And The Bold Future Of News.
36. John Antal, "The First War Won Primarily with Unmanned Systems: Ten Lessons from the Second Nagorno-Karabakh War" (2021), https://www.socom.mil/.
37. Ibid.
38. Ibid.
39. Ibid.; Stijin Mitzer and Joost Oliemans, "Aftermath: Lessons of The Nagorno-Karabakh War Are Paraded Through the Streets of Baku," Oryx (2021), https://www.oryxspioenkop.com/; Rutland, "A Military of Influencers."
40. Mitzer and Oliemans, "Aftermath: Lessons of The Nagorno-Karabakh War Are Paraded Through the Streets of Baku," para. 22.
41. Burak Ege Bekdil, "Ukraine is set to buy 24 Turkish drones. So why hasn't Russia pushed back?" *Defense News* (September 29, 2021), https://www.defensenews.com/unmanned/2021/09/29/ukraine-is-set-to-buy-24-turkish-drones-so-why-hasnt-russia-pushed-back/.
42. Ibid.
43. Joseph Trevithick, "Ukraine Strikes Russian-Backed Forces Using Turkish-Made TB2 Drones For The First Time," *The Drive* (October 27, 2021), https://www.thedrive.com/the-war-zone/42894/ukraine-strikes-russian-backed-forces-using-turkish-made-tb2-drones-for-the-first-time.
44. Antal, "The First War Won Primarily with Unmanned Systems;" Trevithick, "Ukraine Strikes Russian-Backed Forces Using Turkish-Made TB2 Drones For The First Time."
45. Stephen Witt, "The Turkish Drone That Changed The Nature Of Warfare," *The New Yorker* (May 9, 2022), https://www.newyorker.com/magazine/2022/05/16/the-turkish-drone-that-changed-the-nature-of-warfare.
46. Rutland, "A Military of Influencers."
47. Trang Tran, "Personalized ads on Facebook: An effective marketing tool for online marketers," *Journal of Retailing and Consumer Services* 39 (2017), 230-242.
48. Ralph Schroeder, "Big Data and the brave new world of social media research," *Big Data & Society* 1, no. 2 (2014), 2.
49. Tran, "Personalized ads on Facebook."
50. Major Jessica Dawson, "Microtargeting as Information Warfare," *The Cyber Defense Review* 6, no. 1 (2021), 63-80.
51. Ibid.
52. Jim Isaak and Mina J. Hanna, "User data privacy: Facebook, Cambridge Analytica, and Privacy Protection," *Computer* 51, no. 8 (2018), 56-59; Dawson, "Microtargeting as Information Warfare."
53. Dawson, "Microtargeting as Information Warfare," 72.
54. Lt Col Jarred Prier, "Commanding the Trend: Social Media as Information Warfare," *Strategic Studies Quarterly* 11, no. 4 (2017), 51.
55. Ibid., 52.
56. Ibid.
57. Rutland "A Military of Influencers."
58. "The U.S. Army in multi-domain operations 2028," Fort Monroe, VA: Army Training and Doctrine Command (2018), 3.
59. Irina Khaldrova and Mervi Pantti, "Fake News: The Narrative Battle Over the Ukrainian Conflict," *Journalism Practice* 10, no. 7 (2016), 891-901; Singer and Brooking, *LikeWar;* Antal, "The First War Won Primarily with Unmanned Systems".
60. Krulak, *The Strategic Corporal.*
61. Ibid.
62. Stringer, "Educating the Strategic Corporal".
63. Capt. Luke Klena, "Technical Signature management for Small Units," *Marine Corps Gazette,* May 2021 (2021).

NOTES

64. Dasilva et al., “Securing 5G.”
65. Ibid., 85.
66. Singer and Brooking, *LikeWar;* Van Niekirk and Maharaj, “Mobile Devices and the Military.”
67. Prier, "Commanding the Trend.”
68. Patrikarakos, *War in 140 Characters.*
69. Dawson, “Microtargeting as Information Warfare.”
70. Ibid.

Weaponizing Words: Using Technology to Proliferate Information Warfare

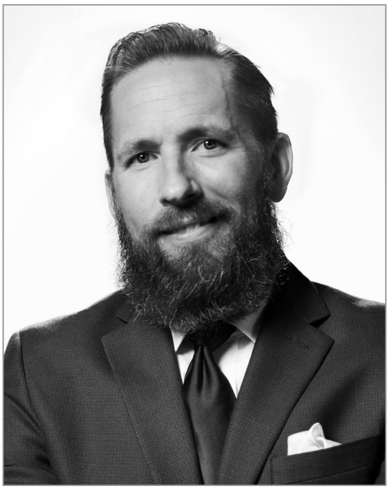
Craig Douglas Albert, Ph.D.
Samantha Mullaney
Lieutenant Colonel Joseph Huitt
Lance Y. Hunter, Ph.D.
Lydia Snider

ABSTRACT

The United States risks losing its information advantage over its near-peer competitors, specifically China. One reason behind this possibility is that the U.S. lacks a coherent doctrine of information warfare, which has put the U.S. at a disadvantage. Considering the Russian interference in elections of several North Atlantic Treaty Organization (NATO) states and allies, including Ukraine, Germany, and, the United States, most stunningly in the 2016 presidential election, this article addresses the question: What is to be done? Before delving into possible solutions, the exact nature of the complex problem must be explored. The purpose of this article is to investigate the ways the U.S. could improve in information warfare, specifically against one of its top near-peer competitors, China. First, this article summarizes how China compares with the United States concerning information warfare and influence operations. Second, it delves into some of the definitional chaos in which the U.S. is mired. Thirdly, the article illustrates the doctrinal and data policies of the U.S. Department of Defense. Finally, it concludes with policy recommendations.

INTRODUCTION

This article asserts that the United States (U.S.) could perform better in the realm of information advantage against its near-peer competitors. Specifically, we examine China's IW (Information Warfare) as it is an increasingly DoD-recognized threat and its growing technological development in the realm of artificial intelligence poses unique threats to the U.S.[1] We demonstrate that the key reason for the current



Craig Douglas Albert, Ph.D., is Professor of Political Science and the Graduate Director of the Master of Arts in Intelligence and Security Studies at Augusta University. His areas of concentration include international relations and security studies, ethnic conflict, cyberterrorism, cyberwar, information operations, and epidemic intelligence. He is widely published, including articles in the *Defense and Security Analysis*; *Iran and the Caucasus*; *Politics*; *East European Politics*; *Chicago-Kent Law Review*; *Politics and Religion Journal*; *Politics & the Life Sciences*; *Cyber Defense Review; Journal of Cyber Policy; Global Society*; and *Intelligence and National Security*. Dr. Albert has testified before a U.S. Congress' joint sub-committee of the House Foreign Affairs Committee. You can follow him on Facebook/Instagram/Twitter @DrCraigDAlbert.

predicament is that the U.S. lacks a coherent doctrine of IW, which puts the U.S. at a disadvantage. China's current advantage is due not to its superior capability, but to the U.S.' lack of clear definition of terms, lack of unified approach, and lack of effective use of data. Thus, the U.S. has the capacity and capability to improve and to regain strategic superiority in this realm. We acknowledge that "information warfare" is not a term currently endorsed and widely used by the U.S. government. In fact, As Ross denotes, the U.S. Army is moving toward a new terminology, contained within the Information Advantage (IA) and Decision Dominance (DD) doctrinal framework.[2] Information Warfare is one of the tasks associated with the IA & DD framework, but we chose to focus on IW to examine an adversary's point of view, and the Chinese Communist party (CCP) is waging information warfare against the U.S.. Also, it is a term commonly used outside the U.S. government and within academia, but we also seek to acknowledge the future of IA & DD in DoD.

As recently as 2018, Seth Jones noted that the U.S. abandoned most of its information capabilities, choosing to focus on lethal rather than political or information operations.[3] Historically, the U.S. has been surprised by its strategic adversaries' sophistication and offensive capability, including non-state actors such as the Islamic State of Iraq and the Levant (ISIS). The Institute for the Study of War acknowledged this in 2016, stating that tactics such as ISIS's virtual caliphate, posed a distinct threat to the U.S. as long as they did not have a clear, government-wide IW strategy.[4] Today, the CCP wields specific information warfare tactics and poses a similar threat.

The U.S.'s IW deficit stems from a lack of a common definition. At times, different units within the U.S. military work against each other, rather than with each other, producing a "silo effect" of data, information, and ultimately intelligence collection and analysis. There is considerable movement within the service branches

Samantha Mullaney is a graduate of Augusta University's Master in Intelligence and Security Studies Program in Augusta, Georgia. The focus of her research is on information warfare forms, tactics, and implications. Her capstone included completing an information warfare internship at the Georgia Cyber Center, where she researched Russian information warfare forms and tactics in Ukraine and the US between 2014 and 2020. Samantha has a BA in History from Fairfield University, an MA in Elementary Education from Boston University, and spent nearly a decade teaching children in Djibouti, Yemen, Jordan, and the UAE. She also speaks intermediate Arabic and basic German. She has been published in *The Cyber Defense Review*.

to adopt and update the language from information warfare and information operations in favor of the term "information advantage." However, many branches are still suffering from a historical lack of common parlance. For instance, when President Clinton established the Broadcasting Board of Governors (BBG), it did not synchronize other elements of public diplomacy or strategic communications, and thus the Departments of State and Defense disseminated different public messaging.[5] Some of this may stem from the fact the Department of State-led Global Engagement Center (which has a vital role supporting information operations) seems to be understaffed, undersourced, and plagued by internal problems that have affected proper messaging in this realm.[6] In fact, Kiesler notes, "There is no recognized leadership to task, direct, resource, or guide policy in the highly complex, disparate field of information operations."[7] LTG Stephen Fogarty and COL (Ret.) Bryan Sparling recently wrote, "The stunning social media-powered rise of ISIS in 2015, Iran's increasing digital belligerence, and China's disinformation surrounding the COVID-19 pandemic" are all examples of information warfare challenges that have begun "a conversation across the defense establishment regarding appropriate roles for the uniformed armed services in this environment of unprecedented information warfare."[8] The above instances of information warfare and information operations (IWIO), as well as Russian interference in several NATO states and allies since at least 2018, begs the question: What is to be done?[9]

Of course, before delving into possible solutions, the exact nature of the complex problem must be explored. The purpose of this article is to investigate how the U.S. is fairing in information warfare, specifically against one of its top near-peer competitors, China. It also seeks to deliver recommendations on how it could do better concluding with specific policies meant to create discussion within the community and mitigate the problems. Before proceeding, however, it is important

LTC Joseph Huitt is a Cyber Warfare Operations Officer, currently serving as Cyber Warfare Deputy Director, Talent Management U.S. Army Cyber Command and also as a senior fellow at West Point's Center for Junior Officers. He is a graduate of the College of Naval Command and Staff, U.S. Naval War College, and a Distinguished Military Graduate of Augusta University. LTC Huitt holds master's degrees in Defense and Strategic Studies, and Intelligence Studies. LTC Huitt has served in leadership positions from the tactical to strategic levels, over his 23-plus years of service. He has gained invaluable experience serving with Special Operations Command in West Africa, USAFRICOM in England, NATO-ISAF in Afghanistan, 66th Military Intelligence Brigade in Germany, USARCENT in Saudi Arabia, 2nd Infantry Division in South Korea, and various stateside units.

to provide some conceptualization of terms that are used throughout this article.

Information warfare (IW) refers to the deliberate use of any element of information to influence the decision making of the adversary and achieve a strategic goal.[10] IW takes place within the information environment, which refers to the physical, informational, and cognitive dimensions that interact with information.[11] Information operations (IO) refer to the specific tactical undertakings in the pursuit of information warfare. The goal of IW is to act in a manner that aids in manipulating the adversary "to win strategic victories and bend the wills of their adversaries without ever engaging in physical combat."[12] It is important to note that IW is used at all stages of warfare, including in kinetic operations. We now turn to a brief illustration of how China dominates the narrative and achieves an advantage across the information environment.

LEFT BEHIND AND OUTMANEUVERED

Malicious actors have benefited from access to modern technology, such as social media platforms, AdTech, and vast troves of stolen data, enabling IW to become one of the cheapest, easiest, and least restrictive types of warfare.[13] The quest to disrupt the decision-making process by using and misusing information is incredibly destabilizing to open societies since IOs target the cognitive domain of individuals and the citizenry as a whole.[14] IW seeks to sow confusion and polarization, thereby destroying the bonds that provide for stability within a society.[15] The U.S.'s historical emphasis on tactical and kinetic activities has placed it at a distinct disadvantage during the current period of conflict between major competing powers, specifically with China.[16] Competing nation-states seek to undermine the U.S.'s democratic norms and stability by utilizing information operations.[17]

Lance Y. Hunter, Ph.D., is Professor of International Relations in the Department of Social Sciences at Augusta University. Dr. Hunter's expertise is in security studies and democratization. Hisresearch focuses on the causes and effects of terrorism and the relationship between evolving technology and conflict. His work has appeared in *Journal of Peace Research*; *Terrorism and Political Violence*; *Party Politics*; *Studies in Conflict and Terrorism*; *Armed Forces and Society; Conflict, Security and Development*; *European Political Science*; *Global Policy*; *Cyber Defense Review*; and *World Affairs.*

China's Strategic Advantage

China possesses a comprehensive doctrine and advanced physical IW assets.[18] This is possibly due in large part to the nature of the totalitarian state, which has more comprehensive control over the information infrastructure than the U.S. and therefore greater strategic advantage.[19] Limited in scope, but strategically long-term, IW measures are consistently implemented, creating a cumulative effect. Chinese IW emphasizes "limited objectives in a limited theatre of operations, conducted away from its borders, higher in tempo, shorter in duration, but highly decisive in nature."[20] By combining the thinking of Sun Tzu and Mao Zedong, Chinese IW is heavily focused on psychology and is used as a weapon in and of itself rather than as a support tool.[21] Most Western scholars define Chinese IW as encompassing China's "three warfares," which include legal, psychological, and media operations. These "warfares" attempt to demoralize the adversary, influence public opinion, and manipulate international law.[22] Most noticeable is China's willingness to use highly integrated IW preemptively, illustrated by its IO campaign against Taiwan.[23] Wortzel explains that China combines electronic warfare, precision strikes, cyber warfare, and attacks on space systems to paralyze an adversary's information capabilities.[24]

Strategically, China adheres to Mao's concept of the "People's War" when waging cyber-enhanced IW. This means utilizing a high volume of cyberattacks or dissemination of disinformation through cyber means. Watts explains the content across platforms is uniform.[25] Furthermore, as a totalitarian state, the CCP can coerce numerous Chinese citizens to do their part and espouse a narrative on behalf of the state, as illustrated by the "50 Cent Party."[26] One advantage is the sheer number of people the CCP has working in this arena. They have the ability to direct vast numbers of actual users to execute bot-like operations. Unlike actual bots, however, these are immune to platform bot

Lydia Snider works for the U.S. Department of the Army as a specialist in foreign malign influence.

violation rules, because behind the accounts are real people. While thousands may be posting at the behest of the CCP, even copying and pasting the same response, when the platform AI studies the event, it sees numerous real accounts, not a bot network. IW is at the forefront of China's revolution in military affairs and is viewed as the critical weapon rather than a support for other military endeavors.[27] China recognizes that it cannot compete with U.S. defense spending and instead, starting in the 1950s, has institutionalized IW which has developed into a Strategic Support Force (SSF), the current central element of China's IW capabilities.[28]

China has created entire institutions to develop IW capabilities, including the Academy of Military Sciences Military Strategy Research Centre, the PLA Academy of Electronic Technologies, and the Xian Politics Academy that trains psychological warfare officers.[29] Additionally, the PLA has utilized simulation training for IW for more than a decade.[30] Psychological warfare units are dispersed throughout the PLA following initial training, providing a common language and doctrine across departments. Additionally, Elsa Kania and John Costello, as well as Larry Wortzel note that China's view of IW subsumes cyber warfare.[31] Given the totalitarian control the CCP needs over the domestic population, this sort of integration of cyber and information capabilities in the international arena would not be out of character. In fact, the control over information and therefore ideology, whether through cyber-mediated elements or not, "may allow for better planning, acquisition, and operations while enabling the creation of a more flexible cadre of personnel tailored toward new paradigms of information operations."[32] China's global network of influencers illustrates this strategy.[33] In this strategy, videos of mostly young Chinese women speaking in the language of the target audience speak of their respect for the target country and its culture and of China as a good friend. These videos appear in over a hundred different languages with almost the same script.[34] Each of these academies and centers gives China a probable advanatage over the U.S. in that they are steadily increasing their understanding of TTPs in the realm of information warfare and have wide dispersion capabilities as well. The resulting strategy allows for more flexibility and fluidity in its offensive operations.

The last two decades have seen China attempt to move from confrontational IW to the appearance of cooperation.[35] However, the facade has grown very thin in recent years with the development of the "wolf warrior diplomacy" strategy, which vigorously targets the U.S. and other Western nations and institutions.[36] China builds the facade through a proliferation of Confucius Institutes, hosting new journalists from Africa in training workshops, and promoting tourism and events for foreign elites.[37] The Belt and Road Initiative is presented as economic cooperation for the betterment of developing states, but large-scale Chinese investment in Africa has led to negative consequences. The CCP's infrastructure investment, a core element of the Belt and Road Initiative, is directly linked to undercutting local construction companies, operating on a profit margin of less than 10 percent, and is often tied to selection and use of Chinese contractors.[38] In addition, these single-source projects often are launched without feasibility studies or may include a clause to allow for a loan's cancellation and immediate repayment.[39] Although the Initiative is presented as a cooperative endeavor, one is reminded that it is indeed another form of Chinese propaganda, aimed at promoting the overall aims of the CCP.

The Chinese strategy focuses on weakening the institutions that stabilize American society by co-opting human networks inside these institutions. Other CCP-backed groups include the Chinese Students and Scholars Association and the China Association for International Friendly Contact. The former is a network across universities that receives funding from the CCP and distributes propaganda targeted at universities where there may be negative narratives about China.[40] The latter organization specifically targets business people and veterans and seeks to shape messaging through invitations to tour China.[41] When China faces an inability to create a façade of cooperation, it relies on different elements of the three warfares to coerce or manipulate adversaries. This is most adeptly seen in China's activity in the South China Sea.[42] China's aptitude in IW is clear. Its fleet of spy ships, SIGINT stations located as far afield as Cuba, its own dedicated SIGINT/EW aircraft, and dispersed human asset network allow it to carry out IW simultaneously along multiple fronts.[43] In terms of media warfare, China has adroitly co-opted media outlets around the world through its front organization, Xinhua News. In Africa especially, this co-option of local journalists has weakened any concerted critique of China's Belt and Road Initiative and extractive policies, helping China wage a psychological war and also enabling the manipulation of Africa's legal structures.

There is little distinction between foreign and domestic media control by the Chinese Communist Party. For example, the Central Propaganda Department controls China National Radio, China Radio International, and CCTV. Consolidating media control is a deliberate attempt to unify domestic and international propaganda narratives.[44] The United Front, Confucius Institutes, and wealthy Chinese working on behalf of the CCP have co-opted universities, professors, think tanks, multinational corporations, and researchers to convey to the public crafted messages on behalf of the CCP, funnel research to China, and censor scientific or academic research that would negatively affect China's reputation.[45] This use of messaging encourages Americans to trust the espoused narrative because it comes from traditionally venerated U.S. institutions, such as universities and think tanks. This creates a unified front within China, where the domestic and international narrative focuses on Chinese supremacy, posing a threat in itself to the effectiveness and longevity of democratic states worldwide. The more people "believe" in China's regime, the more a threat is posed to democratic institutions worldwide, in the long run. This is yet another angle China uses in its information war against the U.S.

Chinese IW is also present on social media platforms. Scott Harold, Nathan Beauchamp-Mustafaga, and Jeffrey Hornung posit that China's use of social media helps it destroy an adversary's command authority through the demonization of a leader and the demoralization of the public.[46] Like Russia's, Chinese IW sees chaos and division as a product of successful psychological warfare, whether waged on social media platforms or through strategically placed individuals parroting a Chinese narrative. Given the totalitarian nature of the CCP, any and every business or actor inside of or connected to China can and may be used for the benefit of the state. One advantage the CCP maintains over the U.S. is its willingness to exert state control over social media platforms, through its censorship of internal conversation and with state control over the now internationally used platform, TikTok. With TikTok, the CCP has a platform that both collects data on users and over which it has complete control of what content is delivered to users.

Currently, China's use of cyber for IW is coupled with a powerful and far-reaching network of human agents cultivated through organizations such as the United Front that help execute highly complex and integrated influence operations.[47] This vast network of human assets in multiple arenas enables China to alter public perception and portray messages favorable to the CCP. Specifically, China targets personnel and institutions with financial incentives to dampen negative publicity.[48] The CCP's response to the COVID pandemic is illustrative of its IW capabilities and its strong coordination between overt and covert IW.[49] Now that a brief case analysis of China's use of IW and IO has been illustrated, it is necessary to understand how and in what ways the U.S. lags behind China in the IW/IO competition. Ultimately, the U.S. cannot replicate the CCP's power over the PLA and utilize its IW forms and tactics without demolishing national and international war standards. That does not mean the U.S. cannot find a way to counter these tactics and maintain democratic norms.

DEFINITIONAL CHAOS

The U.S.'s competitors and near-peer competitors have institutions devoted to the successful utilization of information operations and achievement of strategic advantage in this domain. They also have broad, but useful, definitions of IW. Largely, the U.S.' adversaries define IW as conflict in the information space that forces a specific decision by undermining political, information, social, or economic systems, often using mass psychological tactics to destabilize society by targeting a population.[50] The goal of modern IW against the U.S. is to erode trust in authority and institutions, thereby undermining shared values.[51]

The U.S. government does not have a consistent definition of what IWIOs are, and lacks a dedicated institution or agency with which to wage IWIOs effectively for strategic advantage.[52] IW is divided across multiple agencies in the U.S., such as the Department of State's Global Engagement Center, the CIA, U.S. Cyber Command (USCYBERCOM), and other elements of the military.[53] Essentially, the U.S. uses the "same terms differently in different contexts," which creates confusion and a lack of strategic capability.[54] Scholars such as Whyte define IW as the use or abuse of information to influence the decision-making options and processes of the adversary to achieve military or strategic gains.[55] This is a broad definition that encompasses many tactics within the military and non-military realms. The Army's definition is somewhat similar, noting that IW is a simultaneous effort directed at creating a specific effect in the information environment and is a battle "*of* information," rather than just a battle *for* information.[56] However, a 2012 joint publication from the Joint Chiefs of Staff confined IOs to military operations.[57] In fact, Alicia Wanless and James Pammet note that the U.S. interprets IW/IO in largely military terms and tries to delineate between acceptable and unacceptable actions within these parameters.[58] There is no such distinction for foreign adversaries given their different governing structures.

It is understandable that the military focuses on command and control and how IW targets critical military elements necessary to gain a military strategic advantage. However, the IW waged against the U.S. is far broader than this focused definition. IOs target the cognitive domain of individuals and the citizenry as a whole.[59] China utilizes persistent narratives that cause members of the target society to question themselves, and China seeks to disrupt the decision-making process of a state by using and misusing information. The U.S. government requires a common definition of IW which can be disseminated to national security agencies, the military, and public relations elements. These terms should be clearly defined and the parameters demarcated. The U.S. cannot wage an effective defensive information war without a consistent definition of IW.[60] This article now proceeds to a discussion regarding how the DoD understands and effectuates IW. After detailing this, this article proceeds to set-forth policy recommendations that seek to bolster the U.S.'s IW/IO.

DATA AND THE DEPARTMENT OF DEFENSE

To better understand the impact of information warfare and the U.S. Government's (USG) approach to counter adversary actions, it is imperative to review the existing doctrine and policy that guide it. This article highlights the current guidance from the DoD and some of the challenges of wading through the vast data, directives, and policies which reference decades-old policy, include conflicting guidance, and lack of a common lexicon. To set some common ground, the authors first discuss what DoD defines as data and how this is used to generate information and intelligence. Armed with the understanding that U.S. adversaries and competitors are waging IW, this section outlines the basics of how DoD processes data.

DoD highlights in its Data Strategy that "data is a high-interest commodity and must be leveraged in a way that brings both immediate and lasting military advantage."[61] Joint Publication (JP) 2-0 highlights that raw data must be collected and by itself may not be relevant or useful. As JP 2-0 further illustrates that information consumed solely by itself may be utilized by a commander, but is not of much use for decision dominance. When related to the operating environment and considered in the light of past experience, however, it gives rise to a new understanding of the information, which may be termed *intelligence*."[62] The intelligence directorate enriches information by collecting national tactical means to answer a commander's requirements, enabling decision dominance. DoD made information the seventh joint function in 2017 based on 2016 guidance first established in Joint Publication 1, "Operations in the Information Environment (IE)."[63]

Publicly available information (PAI) is information available on the open Internet and it plays an important role in IW/IO. DOD Directive 3115.18, "DoD Access to and Use of PAI," issued in 2019, outlines the lawful and appropriate access to "obtain, and use PAI to plan, inform, enable, execute, and support the full spectrum of DoD missions."[64] While new directives are important, old directives have not always been updated, causing confusion and gaps in strategy implementation. The U.S. Intelligence Community (IC) is flush with data; however, it is generally just white noise. Because of the definitional chaos of IO,[65] and the silo effect of data, different U.S. agencies approach IWIO differently and are often at odds with one another. Different units across all branches of the military often look at the same data for different issues and do not share the information across the DoD. In many instances, different military organizations are buying the same data from companies under different contracts for each organization. In other words, there is such a disunity of approach in data collection because DoD has not created a data governance entity to manage data acquisition from private industry and make it available across the force. DoD has put the onus on components to develop and implement their own data acquisition plans.[66] Furthermore, if DoD had a data lake that housed curated, publicly and commercially available information, which was available to its components, it would drastically reduce redundant data as a service contracts. This situation is one of the reasons the U.S. is behind the curve relative to China concerning the information domain and battlespace. This strategic adversary has clear conceptual approaches to influence operations, and has a more centralized or unified approach to information warfare and intelligence collection than does the U.S..[67] Thus, a more unified approach will help connect the dots with the U.S.' collected data. It should be noted that the IC has the data at hand but does not always efficiently utilize the data to achieve its ends. As the U.S. plans for future data acquisition it needs to follow its adversaries' lead in tracking narratives in the languages in which they are communicating and bringing on language and cultural experts who understand the nuances of those narratives.

LACK OF A UNIFIED APPROACH

DoD understands the challenges of IW and has developed numerous policies to attempt to address them with the end state of achieving information advantage.[68] However, these new policies failed to provide guidance that would benefit DoD organizations and military branches in the twenty-first century. Despite the existing elements of known national power, diplomacy, information, military, and the economy (DIME), and the aforementioned new policies for DoD, the military branches have developed their own approaches that are not synchronized. The term "IW" is also a point of contention—DoD prefers the term (IO), which encompasses a host of information-related capabilities (IRCs).

DoD has published Joint Publication (JP) 3-13, "Information Operations," in 2012 and updated it again in 2014. The definition of IO outlined in JP 3-13 is "the integrated employment, during military operations, of information-related capabilities in concert with other lines of operation to influence, disrupt, corrupt, or usurp the decision-making of adversaries and potential adversaries while protecting our own."[69] JP 3-13 further discusses that, after analyzing a target audience, desired effects can be accomplished through various means, including DIME actions. Here, the lack of a unified approach becomes apparent as these IRCs are managed separately at the joint level and across all military branches. For context, IRC capabilities can include but are not limited to personnel from the electronic warfare (EW), cyberspace operations (CO), military information support operations (MISO), civil-military operations (CMO), military deception (MILDEC), intelligence, and public affairs (PA) communities.[70]

All the communities mentioned above have developed their own guidance over time, executed it with various authorities, and achieved varying degrees of success. Some of these capabilities are nascent (i.e., cyber), and others have a long tradition (i.e., MILDEC). Historically, it is challenging for DoD to synchronize all these capabilities beyond incorporating them for a specific operation. However, the U.S. Congress has noticed that the environment has changed and identified gaps in its understanding of combating the *shaping operations* U.S. adversaries are conducting within the information environment.

To summarize, the U.S. is behind its strategic near-peer competitors, specifically China, due to the lack of a clearly implemented and unified approach, definitional chaos within the information environment, and inefficient utilization of evolving data and information into intelligence. With the understanding of Chinese influence operations and an illustration of the precise reasons the U.S. is behind its strategic adversaries based on DoD doctrine and implementation, what is to be done?

POLICY RECOMMENDATIONS AND DISCUSSION

The first and most obvious policy recommendation is that the U.S. needs to form a centralized, unified approach dedicated to data, intelligence, and IW. This has already been achieved by the CCP. Although there are some in the U.S. who may oppose the creation of such a plan, this article demonstrates why it is a strategic necessity. The U.S. is losing because of its inability to turn data into operational intelligence and its lack of human capital allocation regarding IW. This gives its adversaries the strategic advantage. What is not necessarily needed is a centralized entity to develop a unified approach. Rather, it is a unified approach based on policy across departments and within a unified command structure.

Existing institutions may provide the backbone from which to consolidate and then disseminate a unified approach to IW. Sue Gordon and Eric Rosenbach argue in *Foreign Affairs* that the Cybersecurity and Infrastructure Security Agency should become the center of gravity for domestic cybersecurity operations.[71] Additionally, they argue that USCYBERCOM ought to be realigned and re-envisioned into something approaching the U.S. Special Operations Command (USSOCOM). In a similar vein, Lieutenant General Timothy D. Haugh, Lieutenant Colonel Nicholas J. Hall, and Major Eugene H. Fan argue that this new information environment requires "tight partnerships among all elements of the DoD, the interagency, and our coalition partners, driving a shift in the weight of effort from preparing for conflict to competing now." They continue, "We do not need a new approach to command and control, but a new framework that both materially creates the awareness among, and organizes the horizontal coordination of, organizations across the continuum of cooperation, competition, and conflict."[72] Regardless of whether a unified approach creates a new entity or reenergizes current entities with new authorizations to handle all aspects of IWIO, this is the first step to help the U.S. counter IWIO by adversaries. It is currently unclear if the upcoming redefinition of terms by the Army, and its switch to using information advantage rather than information operations, as recently noted by Ross, will help or hinder the operational chaos produced by the terminology.[73]

Secondly, once a unified approach is defined, the U.S. needs to develop clear operationalizations and definitions for its information operations and strategic approaches. These concepts need to be clearly codified and implemented across the board, intra- and interagency. Once this is done, it may be necessary to go on the IW offensive. The U.S. needs to set the narrative in several key areas in an assertive way, using digital and social media in a fashion similar to how Radio Free Europe was used in the Cold War to communicate pro-democratic and anti-communist messages to thousands of individuals living behind the Iron Curtain.[74] The advantages and strengths of democracy, democratic participation, and respect for human rights need to lead the agenda-setting program of the U.S..

Currently, the U.S. is playing defense concerning the democratic narrative and, in fact, is generally reactive in response to disinformation and propaganda. There is almost no chance of winning the influence war within the Chinese space if the U.S. does not utilize successful tactics. Justin Sherman explains in a prior article for CDR that the Chinese have built out "variously undemocratic practices, such as online censorship, using digital technologies."[75] However, he also notes that digital authoritarianism affects the international arena, and U.S. national security directly, by allowing authoritarian regimes to consolidate power, encouraging the global diffusion of digital surveillance and propagating the idea of Internet sovereignty, thereby potentially avoiding U.S. deterrence strategies.[76] Thus, authoritarian spaces control the information environment and, conversely, the information environment helps proffer authoritarianism.[77] Playing constant defense is a poor strategy and has been largely unsuccessful for the U.S. Our near-peer competitors' sophistication demands the return to strategic offense in the information environment.

Furthermore, the U.S. must strengthen its defenses. For instance, U.S. policy typically does not allow for individuals within the IC or IWIO domains to engage with fake accounts, bots, or organized campaigns aimed at the U.S. citizenry. In fact, according to Major Jessica Dawson, "The result of this is that there is no agency within the Army charged with understanding the ways in which U.S. adversaries can manipulate the domestic information warfare space... [T] he U.S. Army is unable to assess or respond to threats in the social media space."[78] Although it may draw more attention to these accounts and issues, which the U.S. typically discourages, counter-attacking or taking the offensive may surprise Chinese information operatives. If done in a sophisticated manner, U.S. intervention into these spaces may quickly throw its adversaries into an emotive state, which could derail their policy. The action would also signal a policy and strategic culture shift in the U.S., which could help reassert U.S. dominance in this information space, forcing adversaries to play its game, rather than vice versa.

Additionally, U.S. near-peer competitors use popular influencers to their strategic and cultural advantage.[79] China pushes out influencers targeting its own population, and it hires Western influencers to target the West. In fact, China targets its own population through data-driven analytics to exert domestic control.[80] The U.S. could use a similar methodology against China and foreign adversaries as well, without violating U.S. law, military norms, or democratic codes of conduct. Instead of shutting down DoD military influencers, the U.S. could help them expand to combat Chinese IW/IO. Military members not on TikTok could be used to counter CCP efforts stateside by explaining why they are not on the platform. Active social media influence by exceptionally talented individuals could act as an IWIO deterrence. As Morin states, domestic IIOs would be targeted toward adversarial IIOs and seek to reduce "the viewing of an adversary's IIO content."[81]

As the digital age progresses and the information environment becomes a clearinghouse for great power conflict, the U.S. needs to engage this domain strategically and tactically. It can do so by setting its own agenda in this space, while also remaining dedicated to liberal democracy.[82] As noted earlier, Chinese IWIO strategies focus on active offense at all times; there is no difference in their peacetime versus conflict strategies. To compete within this space, the U.S. needs to choose wisely which elements of IW should be used offensively. David Morin explains that incorporating Information Influence Operations (IIOs) into USCYBERCOM tactics "would allow [the U.S.] to effectively guide perception and even shape the targeted population's perception of reality, if effectively conducted."[83]

The U.S. should also consider its strategic use of the Internet in multiple areas. In terms of web presence on the domestic front, all government sites should be technologically savvy and well-integrated with social media platforms to help bolster government legitimacy among generations that are increasingly technologically-oriented. Additionally, the government should consider policies and procedures that would enable the exclusion of bad foreign actors, companies, and advertisement funding.[84] If the U.S. were to disrupt and deny foreign actors' abilities to disseminate influence operations actively through U.S. companies and Internet platforms, it would begin the process of active defense.

Due to China's regime structure, the U.S. and China are playing two separate games with separate rule books. China is directly targeting U.S. civilian interests, has deep pockets to spread its message, and has control of its own media. It can even pay U.S. companies for advertising space, whereas the U.S. denotes limited funds to IW/IO and does not focus on the same targets. The U.S. should utilize the Internet in a manner that aggressively goes on the offensive on behalf of American citizens. This will likely encourage China to complain that the U.S. has caused offense on the international stage. However, it is long past time for the U.S. to demonstrate clearly its IWIO capabilities and impose costs on its adversaries in their attempts to disrupt American society.

For this to be effective, the U.S. must engage in IWIO through a whole-of-society approach, but one that plays out much differently than the centrally directed, coercive manner of authoritarian regimes. Although this article argues that DoD needs a centralized division and strategy for IW/IO to compete with China, it also needs a decentralized environment which allows for all sectors of U.S. society to engage in the game by their own initiative. This would include defense, entertainment, schools, and the citizenry, as imagined by researchers Cristina-Elena Ivan, Irena Chiru, and Rubén Arcos.[85] The U.S. needs an overarching message to disseminate and, to be effective, it has to come from multiple segments of society. As a part of this whole-of-society approach, U.S. companies will need to play an active role. As Dawson notes, technology companies such as Facebook and Google are ungoverned, unrestricted spaces; as such, they pose a significant security risk for the United States, especially concerning data and intelligence for IOs.[86]

The focus of technology platforms should be to prevent U.S. adversaries from co-opting the platform to wage a disinformation campaign against the U.S. citizenry. Most especially, as Major Dawson insists, "The U.S. must recognize the current advertising economy as enabling and profiting from information warfare being waged on its citizens and address the threat."[87] While we must address the fight the adversaries put in front of us, we win, not by trying to play their game, but by playing ours effectively.

DISCLAIMER

The views presented are those of the author(s) and do not necessarily represent the views of DoD or its components.

NOTES

1. Lily Hay Newman, "It's Time to Get Real About TikTok's Risks," *Wired Magazine* (September 6, 2022), https://www.wired.com/story/tiktok-nationa-security-threat-why/.
2. Lieutenant Colonel Robert J. Ross, PhD, "Information Advantage Activities: A Concept for the Application of Capabilities and Operational Art during Multi-Domain Operations," *The Cyber Defense Review* (Fall 2021): 63.
3. Seth G. Jones, "Going on the Offensive: A U.S. Strategy to Combat Russian Information Warfare," CSIS (October 1, 2018) https://www.csis.org/analysis/going-offensive-us-strategy-combat-russian-information-warfare.
4. Harleen Gambhir, "The Virtual Caliphate: ISIS's Information Warfare," *Institute for the Study of War* (December 2016), https://understandingwar.org/sites/default/files/ISW%20The%20Virtual%20Caliphate%20Gambhir%202016.pdf
5. Lieutenant Colonel Gregory M. Tomlin, PhD, "The Case for an Information Warfighting Function," *Military Review* (September-October 2021): 91, https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/SO-21/tomlin-info/tomlin.pdf.
6. P Jack Kiesler, "A Next Generation National Information Operations Strategy and Architecture," Paper, Belfer Center for Science and International Affairs, Harvard Kennedy School (September 13, 2021), https://www.belfercenter.org/publication/next-generation-national-information-operations-strategy-and-architecture.
7. Kiesler, A Next Generation, 8/36.
8. Lieutenant General Stephen G. Fogarty and Colonel (Ret.) Bryan N. Sparling, "Enabling the Army in an Era of Information Warfare," *The Cyber Defense Review* (Summer 2020): 18.
9. Connor Cunningham, "A Russian Federation Information Warfare Primer," The Henry M. Jackson School of International Studies, University of Washington (November 12, 2020), https://jsis.washington.edu/news/a-russian-federation-information-warfare-primer/.
10. Christopher Whyte, A. Trevor Thrall, and Brian M. Mazanec, "Introduction" in *Information Warfare in the Age of Cyber Conflict*, edited by Christopher Whyte, A. Trevor Thrall, and Brian M. Mazanec, 1-11.
11. See *DOD Dictionary of Military and Associated Terms,* (as of January 2021), and DOD Directive 3600.01, *Information Operations (IO),* (May 2, 2013, incorporating Change 1, May 4, 2017), and GAO, *Information Operations: DOD Should Improve Leadership and Integration Efforts,* GAO-20-51SU (Washington, D.C., October 18, 2019), and Congressional Research Service, *Information Warfare: Issues for Congress,* R45142 (updated March 5, 2018).
12. Scott D. McDonald, Brock Jones, and Jason M. Frazee, "PHASE ZERO: How China Exploits It, Why the United States Does Not," *U.S. Naval War College Review* 65, 3 (Summer 2012) https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1471&context=nwc-review.
13. Justin Sherman, "Digital Authoritarianism and Implications for US National Security," *The Cyber Defense Review* (Winter 2021): 108-109.
14. Christopher Whyte, A. Trevor Thrall, and Brian M. Mazanec, eds., *Information Warfare in the Age of Cyber Conflict* (New York: Routledge, 2021), 237.
15. Catherine A. Theohary, "Information Warfare: Issues for Congress," *Congressional Research Service Report* (2018): 1.
16. Mark Pomerleau,"Why is the United States losing the information war?" *C4ISRNet* (October 2020): 5, https://www.c4isrnet.com/information-warfare/2020/10/05/why-is-the-united-states-losing-the-information-war/.
17. Pomerleau, "Why is the United States?"
18. Manuel Cereigo, "China and Cuba and Information Warfare (IW): Signals Intelligence (SIGINT), Electronic Warfare (EW) and Cyberwarfare," www.lanuevacuba.com/archivo/manuel-cereijo-123.htm.
19. Theohary, *Information Warfare.*
20. J. Yin and P.M. Taylor, "Information Operations from an Asian Perspective: A Comparative Analysis," *Journal of Information Warfare* 7, no. 1 (2008), 3.
21. Yin and Taylor, "Information Operations, 3.
22. Larry M. Wortzel, "The Chinese People's Liberation Army and Information Warfare," *US Army War College, Monographs, Books, and Publications* (2014): 30, https://press.armywarcollege.edu/monographs/506.
23. J. You, "China's Emerging National Defence Strategy," *China Brief* 4, no. 23 (2004): 5-7, http://jamestown.org/china_brief/article.php?issue_id=3152.
24. Wortzel, "The Chinese People's," 13-15.
25. Watts, "China's Propaganda."https://www.wired.com/2016/12/obama-russia-hacking-sanctions-diplomats/.