2019 11th International Conference on Cyber Conflict:
Silent Battle
T. Minárik, S. Alatalu, S. Biondi,
M. Signoretti, I. Tolga, G. Visky (Eds.)
2019 © NATO CCD COE Publications, Tallinn

Understanding the Strategic Implications of the Weaponization of Artificial Intelligence

Dr Joe Burton[1]
New Zealand Institute for
Security and Crime Science
University of Waikato
New Zealand

Dr Simona R. Soare
Institut d'Etudes Européennes,
Université Saint Louis - Bruxelles
Belgium

Abstract: Artificial Intelligence (AI) is expected to have a revolutionary impact across societies and to create economic displacement and disruption in security and defense. Yet the impact of AI on national security and military affairs has received relatively scant attention. The existing policy-focused literature has concentrated mainly on the technological, ethical or legal limitations of deploying AI and on the risks associated with it. This paper seeks to contribute to the debate by outlining the strategic implications of the weaponization of AI for international security. It explores how and in what ways AI is currently being utilized in the defense sector to enhance offensive and defensive military technologies and operations and assesses the ways in which the incorporation of AI into military platforms will affect war fighting and strategic decision-making. The paper is in four sections. Section one develops a typology of military AI that forms a foundation for the rest of the paper. The second section examines the uses of AI in cyberspace and the relationships between 'cyber weapons' and AI capabilities. The third section examines how the embeddedness of AI-based capabilities across the land, air, naval and space domains may affect combined arms operations. The final section distills the main strategic implications of weaponized AI, which include the speed of decision-making and action as well as enhanced domain situational awareness.

Keywords: *artificial intelligence, weaponization, cyber defense, strategy*

1 This paper was drafted with equal contributions from both authors. The authors would like to thank the CyCon program committee and the two anonymous reviewers for their invaluable feedback on this paper.

1. INTRODUCTION

James Cameron's cult film *The Terminator* depicted a dystopian future in which Skynet, a malevolent Artificial Intelligence (AI), initiates a nuclear war against humans to ensure its own survival. The film was released in 1984, well before the advent of modern forms of AI, but was prescient in foreshadowing some of the concerns that have come to dominate debates about intelligent computer systems. The late renowned scientist Stephen Hawking described AI as the single greatest threat to human civilization,[2] which is not the first time scientific and technological innovation has been perceived as an existential threat,[3] and Henry Kissinger has warned that AI will change human thought and human values.[4] In recent years activists, scientists and governments[5] have sought to place UN-level bans on 'killer robots', including Lethal Autonomous Weapons Systems.[6] The technology that *The Terminator* films depicted is not yet with us, and a form of self-aware artificial intelligence described as 'general AI' is, according to most analysts, some decades away, yet the impact of AI in international security is beginning to receive sustained attention.

By some accounts, an AI arms race is emerging between the great powers, and the US, China and Russia in particular.[7] AI systems are already being incorporated into weapons platforms and military technologies, including missile defense systems, Unmanned Aerial Vehicles (UAVs), Unmanned Underwater Vehicles (UUVs), fighter aircraft and naval platforms.[8] In the realm of cyber security, AI could revolutionize how we protect computer systems from nefarious actors, but could also be used to develop much more sophisticated attack vectors, methods and technologies. The proliferation of AI to non-state actors, the rapid pace of technological change and the growing sophistication of the new technologies are also causing concerns, and there is a risk that policymakers are unprepared for sudden shifts in how AI technologies are used. This phenomenon is not new. Legislation gaps often occur with societal transitions to new technologies. It is, however, compounded by the fact that much of the technology is being developed by the private sector, including companies like

2 Kharpal, A. (2017, November 06). Stephen Hawking says A.I. could be 'worst event in the history of our civilization'. Retrieved from https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html.

3 AI is but one of a long list of threats to human civilization, including nuclear weapons, biological and radiological weapons, severe cataclysms and genetic experimentation.

4 Kissinger, H. A. (2018, May 16). How the Enlightenment Ends. Retrieved from https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

5 See for example: Open Letter on Autonomous Weapons. (n.d.). Retrieved from https://futureoflife.org/open-letter-autonomous-weapons.

6 Busby, M. (2018, April 09). Killer robots: Pressure builds for ban as governments meet. Retrieved from https://www.theguardian.com/technology/2018/apr/09/killer-robots-pressure-builds-for-ban-as-governments-meet.

7 Auslin, M. (2018, October 23). Can the Pentagon Win the AI Arms Race? Retrieved from https://www.foreignaffairs.com/articles/united-states/2018-10-19/can-pentagon-win-ai-arms-race.

8 Stewart, P. (2018, June 05). Deep in the Pentagon, a secret AI program to find hidden nuclear... Retrieved from https://www.reuters.com/article/us-usa-pentagon-missiles-ai-insight/deep-in-the-pentagon-a-secret-ai-program-to-find-hidden-nuclear-missiles-idUSKCN1J114J.

IBM, Google and Apple in the US, and Baidu, Alibaba and Tencent in China, leaving legislators struggling to regulate, control and mitigate some of AI's associated risks and to explore inherent opportunities. International organizations are beginning to respond to these challenges and governments are starting to develop their own national AI strategies and investment plans. In 2018, the EU, for example, released a civilian and economy-focused AI strategy,[9] and in the last several years a host of countries, including Canada, China, Denmark, Finland, France, India, Italy, Japan, Mexico, Singapore, South Korea, Sweden, Taiwan, the UAE, and the UK have released strategies to promote the use and development of AI.[10] In 2019, the US published its Department of Defense AI Strategy, which aims to accelerate the integration of AI across the US armed forces.[11]

Despite this growing attention, there are many areas of AI research in both the technical and political realms that are underdeveloped and have received surprisingly scant attention. This is especially true in the security and strategic studies disciplines in which the technical and practical aspects of AI development meet the political and doctrinal ones. How AI will affect military operations and how it can be harnessed to increase and enhance international security are questions that are only beginning to be addressed by security scholars.[12] Two schools of thought appear to be emerging in this nascent literature. The first argues that AI deployment in security and defense will have a revolutionary effect on operations (e.g. human-machine teaming), capabilities (e.g. swarms) and military structures (e.g. human-machine interfaces), and on how militaries interact with the civilian and political realms. Much of the literature in this school draws on the technical specifications of AI applications in the military field to derive conclusions about its likely revolutionary impact (which is arguable and speculative at this point in time). The second school of thought argues that AI will have a more evolutionary impact on international security, that its focus will be on increasing the efficiency of 'dull-dirty-and-dangerous' military tasks and on the speed of decision-making (through accurate situational awareness and actionable intelligence), and that it will not fundamentally change the nature of warfare.

9 Artificial intelligence: Commission outlines a European approach to boost investment and set ethical guidelines. (n.d.). Retrieved from http://europa.eu/rapid/press-release_IP-18-3362_en.htm.

10 Dutton, T. (2018, June 28). An Overview of National AI Strategies. Retrieved from https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd.

11 US Department of Defense (2019: February 28). Summary of the Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity. Retrieved from https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

12 Payne, K. (2018). Artificial Intelligence: A Revolution in Strategic Affairs? *Survival*, 60(5), 7-32. doi:10.1080/00396338.2018.1518374; Cummings, M. L., Roff, H. M., Cukier, K., & Parakilas, J. (2018, June 14). Artificial Intelligence and International Affairs: Disruption Anticipated. Retrieved from https://www.chathamhouse.org/publication/artificial-intelligence-and-international-affairs; Hoadley, D. S. and Lucas, N. J. (2019, January 30). Artificial Intelligence and National Security. *Congressional Research Service*. Retrieved from https://fas.org/sgp/crs/natsec/R45178.pdf; Sheppard, L. R. (2018, November 5). Artificial Intelligence and National Security: The Importance of the AI Ecosystem. Retrieved from https://www.csis.org/analysis/artificial-intelligence-and-national-security-importance-ai-ecosystem; Scharre, P., & Horowitz, M. C. (2018, June 22). Artificial Intelligence: What Every Policymaker Needs to Know. Retrieved from https://www.cnas.org/publications/reports/artificial-intelligence-what-every-policymaker-needs-to-know.

In this context, we argue that empirical evidence and existing governmental AI strategies seem to suggest a middle path: that the role of AI will differ across military tasks. While AI may revolutionize tasks such as logistics and maintenance, it will be evolutionary for others, including decision-making (i.e. humans will continue to make political and military life-and-death decisions). In building this argument the aim of the paper is to shed further light on some of the crucial dynamics that will affect how AI is integrated into strategic planning and affect decision-making in relation to modern war and conflict. In particular, we focus on the process and implications of the weaponization of AI – meaning (a) how AI is and might be incorporated into weapons systems and platforms, and (b) how AI technologies themselves may be used with ill-intent to cause harm in the international arena. The paper seeks to understand the strategic implications of the process of weaponization and the results of that process, and in doing so to raise awareness and help contribute to emerging debates in the military and strategic studies communities about how AI affects military strategy.

The paper proceeds in four main sections. In the following section we outline the types of AI that are being developed that have usages in the military sector. This section works towards a typology of military AI that forms a foundation for the rest of the paper. The next section examines the uses of AI in cyberspace and the relationships between "cyber weapons" and weapons systems that are based on AI tools and capabilities. The following section examines how the embeddedness of AI across the land, air, naval and space domains may affect combined arms operations. The final section distils the main strategic implications of weaponized AI, which include changes in the speed of decision-making and action as well as implications for cross domain situational awareness.

2. TOWARDS A TYPOLOGY OF MILITARY ARTIFICIAL INTELLIGENCE

Much of the debate around the emergence of AI as a factor in military planning has suffered from a confusion about what exactly AI is and its various forms and utilities. This lack of clarity is not surprising given the complexity of the technology and the challenge of advancing scientific understanding in non-scientific communities. Across the international security and strategic studies disciplines, scholars are grappling with the implications of technologies that are opaque, highly technical, and developed by scientific disciplines with which they have had little interaction. The profusion of various forms of AI and their already widespread usage in the commercial sector has also complicated efforts to categorize and define the emerging AI marketplace. Voice recognition and commands are now built into everyday objects and platforms, and algorithms that predict and analyze information in real time are used extensively

across a range of societal activity, including in the financial sector, market decision-making, and in software and computer hardware development. Yet often, the blanket term "AI" is used to describe a range of technologies, methods and processes which are different and distinguishable from one another.

At the most basic level, AI is a form of technology that exhibits human characteristics – most notably that of intelligence. Intelligence is the ability to reason and perform complex tasks, to understand and adapt to one's environment, and to learn from previous interactions and situations.[13] Intelligent machines will be able to perform complex tasks, be able to learn and improve operationally over time, and do so without human input. Moving beyond this basic definition, the first type of AI classification is a disciplinary one: *practical AI* refers to technological development and computing requirements associated with technical progress; and *fundamental AI* refers to the social, economic, psychological, philosophical and political implications of AI use.[14] Practical AI has seen its ups and downs since the 1950s. In the last decade there has been an exit from the "AI winter" of the previous several decades, a period where technological advancement stagnated, and there have been some rapid technological advancements. Fundamental AI, however, has struggled to keep up with the technological progress in practical AI. The growing gap between the two was well framed by Henry Kissinger, who has said we are in the presence of "a potentially dominating technology in search of a guiding philosophy".[15]

A further distinction in the contemporary literature on AI technology relates to the number of tasks it can perform at a time. The first category is *narrow AI*, which is the most common type of AI already in civilian and military use: this refers to technology that can perform a single task at a time – the task it has been specifically built to perform. It does not have the ability to migrate the knowledge or behaviors it is taught or has learned in one context to other situations. Scholars refer to this limitation as "catastrophic forgetting", meaning *narrow* AI cannot be repurposed for other tasks.[16] The systems involved are either *reactive*, in that they are not capable of forming memories or using past interaction to shape decisions, or have *limited memories*, in that they might process simple pieces of past information but are not capable of using that information systematically to influence or make decisions.

The second category is *general AI*, which is not yet deployed either in the civilian or the military realms. Through analogy with human intelligence, general AI is supposed to be able to perform several tasks at a time. It has the ability to understand context, to successfully apply information and behaviors learnt in one context to other situations

13 Intelligence. (n.d.). Retrieved from https://www.merriam-webster.com/dictionary/intelligence.

14 The authors would like to thank the anonymous reviewer for raising this point of difference.

15 Kissinger, H. A. (2018, May 16). How the Enlightenment Ends. Retrieved from https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

16 Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521-3526. doi:10.1073/pnas.1611835114.

it encounters, and in circumstances other than the task it was designed to perform. In this category, intelligent machines will be able to adjust behavior depending on interaction with people and other technologies and understand the context, motivations and complex intentions of these actors. This type of AI has been referred to as "theory of mind" AI.[17] *General* AI in its most sophisticated form may become self-aware – this is a field of AI often referred to as artificial consciousness, machine consciousness, synthetic consciousness or singularity. The debates around the plausibility of the emergence of self-aware forms of AI are ongoing. In "Artificial Consciousness: Utopia or Real Possibility", Giorgio Buttazzo refutes the possibility that machines can exhibit consciousness,[18] but some scholars argue that AI may develop a level of sophistication *commensurate* with the human mind.

Thirdly, AI can be classified both as *software* and as *hardware*. Technically speaking, AI is an individual algorithm or system of algorithms (i.e. software). However, AI software is most generally deployed *together with and/or integrated on* physical platforms, be it robots, drones or systems of sensors. AI, either software or hardware, is dependent on being developed and deployed in a data ecosystem that it can monitor, exploit or adapt to achieve its tasks. In this sense, AI is fundamentally creating new capabilities and capacities for military institutions across the world, much like 'systems of systems' did in the late 1980s and early 1990s.

Another means of classification for AI refers to the *types of tasks or roles* it can perform. In the field of international relations and security, AI roles are generally considered to be analytical, predictive or operational.[19] Depending on the category, some roles are more important and likely to be more transformative than others: analytical roles provide decision-makers with actionable intelligence and improve situational awareness; predictive roles may have a significant transformative role at the tactical, operational and strategic level of military operations; whereas at the operational end of the spectrum, AI, robotics and automation are expected to take over a number of dull, dirty and dangerous tasks. Depending on the roles it is deployed to perform, AI software procurement is unlikely to result in easily quantifiable capabilities; AI in the form of lethal autonomous weapon systems, however, such as swarms, autonomous drones or autonomous underwater vehicles, will lead to the development of countable military capabilities. Swarm strategy and the intelligent collective behavior of these swarms is surely one of the most promising fields of AI R&D. Moreover, human-machine interaction, collectively and individually, and its

17 Minsky, M. L. (2007). *The society of mind*. New York: Simon & Schuster Paperbacks; see also Azarian, B. (2018, November 8). Intelligent Social Robots Must Have a "Theory of Mind". Retrieved from https://www.psychologytoday.com/us/blog/mind-in-the-machine/201811/intelligent-social-robots-must-have-theory-mind.

18 Buttazzo, G. (2001). Artificial consciousness: Utopia or real possibility? *Computer*, 34(7), 24-30. doi:10.1109/2.933500.

19 Cummings, M. L., Roff, H. M., Cukier, K., & Parakilas, J. (2018, June 14). Artificial Intelligence and International Affairs: Disruption Anticipated. Retrieved from https://www.chathamhouse.org/publication/artificial-intelligence-and-international-affairs.

technical and legal interfaces, will also create new capabilities. Therefore, AI is likely to significantly impact the qualitative and quantitative international balance of power.

AI is a dual-use technology, and as with all dual-use technology its specifications determine the degree to which it is likely to spread in the military or civilian realms. At the present time, the forms of AI in usage in the military sector are predominantly narrow AI, including reactive and limited memory AI. These forms of technology have been incorporated into a wide range of military platforms, systems and processes. At the softer end of the security spectrum, AI is in use in logistics and training; augmented reality systems, for example, are already in use in the Royal New Zealand Navy for training engineers to work on naval platforms.[20] In its perhaps most widespread and currently consequential role, AI is being used for Intelligence, Surveillance and Reconnaissance (ISR). One controversial example is the National Security Agency's (NSA) 'Prism' program, which applied AI systems to big data for counter-terrorism purposes.[21] At the harder end of the military spectrum, AI is being incorporated into missile defense systems, drones and other unmanned vehicles capable of deploying military force, and in targeting for weapons systems. The Israeli Harpy drone – a loitering munition also known as a 'fire and forget' system – is, judging by its technical specifications alone, a fully-autonomous weapon system.[22] The Japanese military is also considering acquiring ballistic missile defense drones that are capable of autonomously tracking incoming missiles.[23]

Conceptualizing AI weaponization

While there is a wide range of usages of AI in the military sector, the more consequential series of concerns exist at the harder end of the security spectrum. Significant concerns have arisen over the weaponization of AI. In this article we use this term to refer to two connected processes. The first is the use and integration of AI technology in weapons systems and platforms across the four domains of warfare (land, air, sea, space) for strategic advantage. In this first category, AI is used to enhance and multiply the effects of military operations, to enable rapid dispersion and concentration of force, to increase the lethality, precision and destructiveness of the application of military power, to give offensive operations an advantage and to erode an adversary's ability to defend itself. The second way we conceive of weaponization is through the use of AI as a stand-alone capability to undermine, disrupt and destroy enemy systems through computer network-enabled operations. Weaponization thus refers to both its use to enhance the power of conventional military assets, and the weaponization of the software and data through and within cyberspace (the 5th domain). The latter is dealt with in a following section.

20 Author visit to Devonport Naval Base, Auckland, NZ.

21 Kalakota, R. (2013, June 17). NSA PRISM – The Mother of all Big Data Projects - DZone Big Data. Retrieved from https://dzone.com/articles/nsa-prism-—-mother-all-big.

22 Harpy NG. (n.d.). Retrieved from http://www.iai.co.il/2013/36694-16153-en/Business_Areas_Land.aspx.

23 Sakhuja, V. (2018, June 27). Asian Militaries and Artificial Intelligence. Retrieved from http://www.indiandefencereview.com/asian-militaries-and-artificial-intelligence/.

The process of weaponization has been studied in various security-related fields, the most prominent being the weaponization of nuclear materials and programs.[24] Similar concerns have been documented concerning the weaponization of toxins and biological and chemical agents, and the manipulation of weather and climate has even been examined in the concept of weaponization.[25] There is also a substantial literature on weaponizing outer space, most often referring to placing military assets and capabilities in earth's orbit. More recently, the notion that information is being weaponized has received significant attention, especially in the context of Russian information operations, active measures and the use of cognitive behavioral algorithms to achieve 'mass manipulation' effects.[26] Common to existing analyses of weaponization processes is the use of civilian or dual-use technologies for military purposes. This basic dynamic applies to nuclear, outer space, biological agents and much of the other weaponization literature. AI has widespread uses across societal functions and, unlike the internet, which was originally a military network, has not been developed with military purposes at the forefront of planning and funding. However, the military has clearly been interested in the functionality of AI technologies for some time, including for the purposes of achieving strategic surprise, achieving a military advantage over one's opponent or otherwise creating politically-driven military effects.

The process of weaponization – be it in the nuclear or information area – entails considerable risks. These are associated with the instability that the proliferation of technologies within the international arena creates, the prospect of arms races and security dilemmas, the risk that non-state actors will acquire weaponized agents, the risk that states will not be able to effectively control the weaponized technology, and that AI technologies will be uncontainable and result in unintended consequences when used. The risks associated with the weaponization of AI have not been outlined systematically[27] but include the development of bias within AI systems. This dynamic was demonstrated recently when a Microsoft chatbot called 'Tay' was given its own Twitter account and allowed to interact with the public and, as a result of being fed malicious data, began to exhibit racism, sexism, and extremist political viewpoints. If bias develops within AI that is integrated into military systems, either as a result of manipulation or by the nature of the algorithm or data it processes, it will not serve to enhance military effectiveness. Another significant risk with AI systems is that they can be manipulated, and their integrity altered by malicious actors and even

24 Thakur, R. (2014). The inconsequential gains and lasting insecurities of India's nuclear weaponization. *International Affairs*, 90(5), 1101-1124. doi:10.4324/9781315749488-8.

25 Pincus, R. (2017). 'To Prostitute the Elements': Weather Control and Weaponisation by US Department of Defense. *War & Society*, 36(1), 64-80. doi:10.1080/07292473.2017.1295539.

26 Waltzman, R. (2017, April 27). The Weaponization of Information: The Need for Cognitive Security. Retrieved from https://www.rand.org/pubs/testimonies/CT473.html.

27 Cummings, M. L., Roff, H. M., Cukier, K., & Parakilas, J. (2018, June 14). Artificial Intelligence and International Affairs: Disruption Anticipated. Retrieved from https://www.chathamhouse.org/publication/artificial-intelligence-and-international-affairs.

programmed to perform unintended functions.[28] AI has also created concerns over social manipulation. Sophisticated data algorithms were used to affect social media in the run-up to the 2016 US general election and to exacerbate societal tensions, thus exhibiting the utility of weaponization of information by authoritarian states to undermine democratic ones. There have also been several concerns highlighting the misalignment of goals between humans and machines, where an AI is programmed and intended to accomplish a specific task but may not proceed according to the expectations of the programmer.[29] The lack of transparency of most AI algorithms in performing designated tasks is a significant problem and creates obstacles to their deployment in active security and defence roles.

3. THE WEAPONIZATION OF AI IN CYBERSPACE

Enhancing cyber security is becoming increasingly challenging due to the growing number of internet-connected devices and the exponentially increasing volume of data produced that needs securing. These basic dynamics affect the deployment of AI in cyberspace directly. The volume of data produced is such that humans will never be able to monitor data networks without assistance from machines. Cyber networks are vast and carry vast amounts of data. Monitoring the security of these networks is an exponentially increasing challenge in the 21st century. The potential for AI to have a positive impact in this area is obvious, particularly in enhancing the ability of human operators to monitor and respond to adversarial and abnormal events. As Vinod Vasudevan argues:

> Today's systems generate so much security data that human experts are rapidly surpassed. People cannot find the attack elements fast enough or reliably enough. By comparison, computers excel at these operations. AI then helps them to make sense of what they find. It can even help by offering suggestions to security teams of processes to handle them.[30]

AI may thus help mitigate offensive actions. It may also help to more effectively attribute cyber-attacks to specific actors by enhancing information and digital evidence collection and by providing probabilistic models to assess contradictory and uncertain data.[31]

28 Hoadley, D. S. and Lucas, N. J. (2019, January 30). Artificial Intelligence and National Security. Congressional Research Service. Retrieved from https://fas.org/sgp/crs/natsec/R45178.pdf.

29 Worley, G. G., III. (2018, February 19). Formally Stating the AI Alignment Problem. Retrieved from https://mapandterritory.org/formally-stating-the-ai-alignment-problem-fe7a6e3e5991.

30 Vasudevan, V. (2018, July 24). How AI Is Transforming Cyber Defense. Retrieved from https://www.forbes.com/sites/forbestechcouncil/2018/07/24/how-ai-is-transforming-cyber-defense/#8b13293bb20a.

31 Nunes, E., Shakarian, P., Simari, G. I., & Ruef, A. (2018). *Artificial intelligence tools for cyber attribution*. Cham, Switzerland: Springer.

But how does the deployment of AI in cyberspace relate to the weaponization debates introduced in this article? First, there has been increasing concern in scholarly and policy circles about the vulnerability of AI to malicious interference affecting the integrity and operability of those systems. As we have stated, AI is software that exists on hardware. It is present on computers and computer networks that are just as vulnerable to intrusion and exploitation as any other computer network. AI is also based on sophisticated algorithms which can be manipulated or corrupted in the same way that other data can. Hackers are already developing tools to manipulate AI and turn it against the controller/user. This is beginning to be interpreted as an emerging security crisis.[32] There are several crucial concerns here. The first is that AI may be fooled into seeing things that are not there, misclassifying objects and processes, and/or failing to identify patterns or processes within data that has become corrupted or corruptible.[33] Researchers at University of California, Berkeley, for example, recently invented a stop sign that could fool driverless cars. The implications of this in the military realm are significant. If military vehicles are manipulated into taking or not taking actions that are based on adversarial mal-intent, then serious consequences could ensue. Military satellites could be fooled into misclassifying military assets, which could have negative implications for situational awareness and decision-making. Manipulation of AI-based image identifiers could also be used to deliberately misidentify terrorist suspects, for example.

Advances in AI may also make malware itself more damaging, more sophisticated and better able to precision-target its intended recipient. One recent example is the Deeplocker malware, developed by IBM Research, which is highly evasive and able to conceal its malicious intent before it reaches its target. The malware identifies targets through social media indicators, including facial recognition, geolocation and voice recognition, and avoids detection until delivering its 'payload'. It has the potential to operate across millions of devices and was demonstrated recently as a mechanism to distribute the Wannacry virus covertly through video conferencing apps.[34] This is just one example in an expanding range of offensive capabilities enhanced or facilitated by AI. Others include spear-phishing campaigns that harness big data for more targeted social engineering attacks; 'hivenets' – artificial intelligence enabled botnets that harvest data to compromise additional devices; extensive-tailored attacks – which are large numbers of targeted attacks conducted simultaneously through the application of AI; and advanced obfuscation techniques – including efforts to misdirect defenders by learning from data from past campaigns.[35]

32 Kobie, N. (2018, September 12). To cripple AI, hackers are turning data against itself. Retrieved from https://www.wired.co.uk/article/artificial-intelligence-hacking-machine-learning-adversarial.

33 Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J. Erhan, D., Goodfellow, I, and Fer, R. (2014, February 19). Intriguing properties of neural networks. Retrieved from https://arxiv.org/abs/1312.6199.

34 Smith, Ms. (2018, August 08). Weaponized AI and facial recognition enter the hacking world. Retrieved from https://www.csoonline.com/article/3296098/security/weaponized-ai-and-facial-recognition-enter-the-hacking-world.html.

35 Artificial intelligence technologies boost capabilities of cyber threat actors. (2018, February 28). Retrieved from http://thetimesofafrica.com/artificial-intelligence-technologies-boost-capabilities-cyber-threat-actors/.

A related concern is that AI could be used to enhance information operations and target populations with the intent of causing instability or division. In that way, AI might be a multiplier or amplifier of information warfare. More generally, the use of AI in cyber operations poses many risks similar to those that have been identified with 'cyber weapons' (loosely defined as malware designed and intended to cause damage). These have been amply documented elsewhere, but include the ability of states and non-state actors to reverse engineer malware, collateral damage (Wannacry and Stuxnet spread to hundreds of thousands of computer systems in over a hundred countries), the dangers that investment in cyber weapons can create security dilemmas and arms races within the international system,[36] that cyber weapons can be stolen and reused,[37] and the fear that proliferation of AI to less restrained and less deterrable non-state actors may create heightened levels of danger and instability.[38] In this sense, concerns over the weaponization of AI within cyberspace are closely related to (although not necessarily the same) as the weaponization of malware for strategic objectives.

4. BATTLEFIELD AI? USE OF AI IN COMBINED ARMS OPERATIONS

While AI can be weaponized within and through cyberspace and has the potential to cause considerable harm when used with malicious intent within computer networks, the ability to integrate AI into existing weapons systems or deploy it on next generation military platforms is equally apparent. In this section we explore how AI might be used on the battlefield in combined arms operations to achieve strategic objectives.

At this juncture, there are two possible paths through which AI could be utilized in joint operations to generate military advantage: either it will be integrated within existing doctrines and battle concepts (*evolutionary* perspective), including being deployed to enhance existing capabilities, or to improve the speed of action and effectiveness of the human environment. Alternatively, the application of AI in the military field, either independently or in conjunction with other emerging technologies such as quantum computing, big data analytics, advanced robotics, human enhancement technologies, and automation, will lead to the development of new doctrines that defy the existing physical and legal boundaries of today's battlefield (*revolutionary* perspective). The application of AI into combined armed operations will likely depend more on the national models of inclusion of AI into the military field and the usefulness of this

36 Buchanan, B. (2017). *The Cybersecurity Dilemma: Hacking, trust and fear between nations*. Oxford: Oxford University Press.

37 Baram, G. (2018, June). The Theft and Reuse of Advanced Offensive Cyber Weapons Pose A Growing Threat. Retrieved from https://www.cfr.org/blog/theft-and-reuse-advanced-offensive-cyber-weapons-pose-growing-threat.

38 Maurer, T. (2018). *Cyber Mercenaries: The state, hackers, and power*. Cambridge: Cambridge University Press.

emerging technology, rather than a general set of technical specifications. Application of AI in combined operations, however, will likely, at a minimum:

(a) Facilitate real-time analysis and improve situational awareness of the battlefield;
(b) Provide troops on the ground with actionable intelligence and enhanced decision-making;
(c) Facilitate dispersion or rapid concentration and application of lethal power, thereby enhancing mission precision and improving military effects;
(d) Act as a logistical aide by providing predictive maintenance and supply for military equipment, increasing the safety of operating equipment, reducing operational costs and thereby improving the readiness and deployability of troops;
(e) Enable robotics systems to serve a variety of military functions, including the use of lethal force;
(f) Fulfill jobs in the military that are dull, dangerous or dirty, including enhancing force protection and reducing casualties.

At a broader level, the effect of the application of AI in the military field will affect the balance of power at least through doctrinal changes and adaptations or through the creation of new capabilities; a new computer powerful enough to perform real-time big data analysis in ISR and discern actionable intelligence, for example. It will also affect the interplay between different levels of action, creating opportunities for tactical maneuvers (especially because of superior speed of decision and action) to have operational or even strategic effects, particularly through *fait accompli*, increasing strategic surprise and creating perceptions of first mover advantage (i.e. intensifying the security dilemma).

AI will likely create the conditions for the return of warfare operations 'in mass' again. Mass will become increasingly important, whether in data and intelligence or in actual capabilities deployed on the ground. In this context, as well as in the context of AI-cyber jointly, it is interesting to consider the idea of attrition: are these new capabilities likely to be used for attrition purposes or for disruption purposes, or both? This leads us to the question: is AI, together with cyber and a number of emerging technologies, likely to lead to the emergence of a new era of weapons of mass attrition or weapons of mass disruption? For example, active measures doctrine in Russia is a type of attrition in that it seeks to deplete the opponents' sources of power (be it the integrity of their democratic institutions, the integrity of their information systems, and public support) but it may also act as a type of disruption, including disrupting the functioning of a national power apparatus and incapacitating the opponent from acting at the speed of relevance. Russia has not released a formal strategy for AI and

is encumbered in some areas of technology by a lack of industrial and technological innovation, but its operational doctrine appears to suggest that the main current function of its AI capability is attrition – i.e. it is aimed at undermining the political cohesiveness and solidarity of the 'West' over time. That is not to say, however, that the Russian government will not use the technology for mass disruption, especially at a time of armed conflict and or international crisis.[39]

Interoperability will be increasingly affected by AI. Developing and deploying AI that is compatible across different branches of the armed forces will be challenging. The ability of two or more different AI-enabled systems to cooperate seamlessly in pursuit of combined mission objectives will be critical to achieve military advantages and mission effects. There are a number of states developing AI-enabled capabilities that have expressed an interest in maintaining interoperability with allies and partners,[40] but there are equally powerful protectionist forces in the defence industry which may present obstacles to seamless multinational interoperability.

Critical decision-making at the political level and on the battlefield will remain human in the age of AI. However, human-machine teaming and other blending solutions will enhance the application of power. Ultimately, it is unlikely that humans will be able to exert *full* control and authority over AI systems *at all times*. The notion that has been often stated on the military side of the LAWS debate, that there will always be an element of human control, appears to be fanciful in the current context. Trust will be an integral factor – military decision-makers will have to either trust from ignorance or from verification. In this context, testing and exercises involving AI and the generation of data pertaining to reliability and integrity will be paramount. This also raises questions about process, and how military decisions are made, including the centralization of command functions relating to AI. This is an old issue in many ways – centralized command structures have always had to adapt to the deployment of new battlefield technologies. In the field of AI, however, we believe it will be important to assess and resolve the balance between AI-based decision-making being distributed to commanders in the field, based on actionable AI generated intelligence, and the slower (but perhaps safer) centralization of AI command and decision-making.

Politics in this respect will be integral to outcomes in the deployment of military AI across domains. Strategy has always been the use of force to achieve political objectives, but we assert that politics will shape how AI is used as much as being the goal of the deployment of AI. What AI will not be able to do for combined arms operations, or any other type of operation for that matter, any time in the near future,

39 Polyakova, A. (2018, November 16). Weapons of the weak: Russia and AI-driven asymmetric warfare. Retrieved from https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/.

40 For example, the 2019 US Department of Defense AI Strategy, the EU's 2019-2020 Work Programme for the European Defence Industrial Development Programme and the 2019 Work Programme for the Preparatory Action on Defence Research reference interoperability in AI-enabled capabilities.

is lift the fog of war: the veil of uncertainty around the interests driving opponents' actions. It will not alleviate the security dilemma and may complicate arms control and disarmament efforts as barriers to entry are lowered due to the acceleration of technological progress in the civilian sector.

5. STRATEGIC IMPLICATIONS

The purpose of this article is not to provide definitive conclusions as to how AI will affect strategy. As Clausewitz often stressed, the unseen complexities involved in military affairs do not allow for clear answers.[41] The purpose of this paper is rather to enhance understanding of different aspects of what policymakers and military officials will face as AI technologies are integrated into war and conflict. In that spirit, we see several considerations as paramount to current and future strategy and policy.

The first is the requirement for and the simultaneous challenge of greater military-civilian fusion. We recognize this as a tautology that has always been true. However, it seems clear that militaries will need to develop much closer cooperation with the private sector in the development and use of AI technology through 'spin in' effects. China has already recognized this, as detailed by Elsa Kania in a recent report, and is working to fuse military and state-owned enterprise efforts to enhance China's AI capabilities and technologies.[42] In this respect, the extent to which China has an inherent advantage over the US because of state control of private enterprise is likely to influence the emerging power struggle over AI. China certainly has some advantages, including a productive and innovative economic and industrial base, and the clear articulation of national strategies around AI, but the notion that direct control over industry confers an advantage should be questioned. Much of historical innovation in technology has been derived from research conducted in private enterprises and research labs, sometimes with government funding. China's technological progress has also been driven, at least in part, by illegal appropriation of technologies and copyright theft, largely through cyber espionage. This has been amply documented.[43] The latest research suggests that China faces significant challenges in developing technologies due to the exponential increase in the complexity of military technology and the difficulties involved in replication and imitation.[44] In the US and Europe, conversely, the challenge will be to develop effective cooperation between the

41 Otte, T. (2002). Educating Bellona: Carl von Clausewitz and Military Education. In G. Kennedy & K. Neilson (eds.), *Military education: Past, present, and future*. Westport, CT: Praeger.

42 Kania, E. B. (2017, November). Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power. Retrieved from https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

43 Laskai, L. L., & Segal, A. (2018, December 6). A New Old Threat: Countering the Return of Chinese Industrial Cyber Espionage. Retrieved from https://www.cfr.org/report/threat-chinese-espionage.

44 Gilli, A., & Gilli, M. (2019). Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage. *International Security*, 43(3), 141-189. doi:10.1162/isec_a_00337.

military and private sector in the development of AI, while managing concerns around ethics and privacy. Recent reports suggest that the US military is now more trusted to develop AI systems than some of the big tech companies such as Google and Facebook, reflecting recent controversies around social media being used as a platform for AI-enabled information warfare and data privacy breaches.[45] However, it is our contention that technology must be jointly and collaboratively developed, and that military control of AI innovation will ultimately be counterproductive, largely because of the need to apply the technology across a wide range of societal activity.

Second, we expect that there will be an ongoing evolution (not revolution) from information warfare to intelligent warfare and that this process will define technology's use in conflict.[46] The outcomes of military conflict will not just be decided by who controls the information environment, but the application of AI to that information, to monitor it, to manipulate it, to degrade it and to harness it with the aim of achieving political ends. We recognize that there is no AI without information processing and that AI is already a social and collective technology that relies on information being fed into it. But the acceleration of this process as a result of big data trends is clearly significant. Battlefield commanders will need to gain an accurate view of the operational environment and achieve an understanding of how information flows through it, the extent to which AI systems can better inform military decisions, enhance insight, better predict what enemy forces might be planning, and minimize error. Access to information and large volumes of data will be paramount, and there will be increased competition, particularly in the early stages of military conflicts, over gaining access to and denying adversaries information.

Third, there will be a scale of human involvement depending on the military function. To express this simply, there will always be human control over AI pertaining to the deployment of nuclear weapons; authority is unlikely to be delegated to computers and algorithms at the high end and in the most destructive areas of military power. However, military decision-making and autonomous decision-making are likely to occur in other military functions such as logistics and situational awareness, for example. In this respect there is a spectrum of decision-making in AI and not a binary with humans involved or not. The novelty of AI should be noted here. We already have AI platforms – such as in the area of missile defense, the Israeli Harpy drone, and automated Russian tanks – that are fully capable of being autonomous, but they have not yet been fully deployed or relied upon. This is because of: (a) the fallibility of human control or decision-making; (b) the competition between states restricting the extent of deployment; (c) the lack of determination of the acceptable uses of AI; and

45 Kahn, J. (2019, January 10). U.S. Military Trusted More Than Google, Facebook to Develop AI. Retrieved from https://www.bloomberg.com/news/articles/2019-01-10/u-s-military-trusted-more-than-google-facebook-to-develop-ai.

46 Kania, E. B. (2017, November). Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power. Retrieved from https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

(d) the shadow of the future – i.e. fear of the normative and political consequences of AI's use in the battlefield.

Fourth, we expect that situational awareness both within computer networks and on the battlefield in tactical and operational environments will be considerably enhanced. There are already trials of battlefield AI that can significantly enhance the awareness that soldiers have of the environment, allowing them to be notified of enemy troop presence and movements, and these will lead to a more proactive approach to threat identification and mitigation. Mission control has always been based on sensing, perception, comprehension and prediction (battlefield situational awareness) and has always been meant to provide effective real-time decision support.[47] AI will accentuate the importance of these functions. Trials of these types of battlefield AI have already taken place, such as those developed by the Defence Science and Technology Laboratory (Dstl) and UK industry partners (SAPIENT).[48] Because of this, we expect that the role of humans in the battlefield will be reduced: drones, for example, have enabled us to place distance between ourselves and violence, and this trend will likely accelerate with advances in AI. Automated systems will be increasingly capable of doing the dirty work that soldiers used to do, and AI will enable commanders to keep forces out of harm's way more effectively.

Relatedly, while AI has been presented in certain debates (and certainly in *The Terminator* films) as posing a great threat to humankind, the prospect that 'killer robots' might take the place of human combatants is not without its benefits. Military commanders will likely be focused on harnessing AI to minimize danger, for force protection, and for deterrence as much as for offensive actions. In this respect, while the weaponization of AI is likely to be an ongoing driver of AI adoption in the military, the technology can clearly be harnessed to enhance security as well as destroy.

6. CONCLUSIONS

This article has sought to highlight some of the key strategic implications resulting from the weaponization of AI, but it is but one of a handful of early scholarly ventures into the strategic use of AI technologies. We are sure it will not be the last. The state of AI research in the strategic studies and security studies areas is still in its infancy. In the next decade, the literature is likely to expand, just as the cyber security literature has done in the previous decade. This will bring much-needed answers to questions over how AI will affect war, conflict, and strategy.

47 Endsley, M. R. (2002). *Designing for situation awareness: An approach to human-centered design*. London: Taylor & Francis.

48 Evans, V. W. (2018, September 24). Artificial intelligence weaponry successfully trialled on mock urban battlefield. Retrieved from https://www.telegraph.co.uk/news/2018/09/24/artificial-intelligence-weaponry-successfully-trialled-mock/.

Overall, we believe AI will continue to shape the battlefield and provide a driving force for the evolution of strategy itself as we move further into the 21st century. It will do so because AI systems will continue to be integrated into weapons systems and used to enhance the precision, lethality and destructiveness of the use of military force. Furthermore, AI will have varied and influential impacts on cyber defense and offense and is likely to continue to be weaponized – to be used with the intent to cause harm and damage – within and between computer networks. We see several other key impacts related to the emergence of AI. These include the magnification of the cognitive ability of military commanders, and, provided AI can be secured from intrusion and manipulation, that decision-making will become more intelligent and less prone to error. Again, this will be a revolutionary or evolutionary process depending on the task AI is set to perform and the domain it is activated in. Clearly the structure of militaries will also need to adapt to AI – especially as swarm technologies and multi-agent systems are developed – and new decision-making processes will need to be adopted. We are at the early stage of that process. Relatedly, constant attention will need to be given to the legal, ethical and strategic debates around human enhancement – including the physical and cognitive development and evolution of military forces, and how psychical and cognitive processes might change and evolve as weaponized AI is increasingly integrated into war fighting.

This leaves us with some big questions. Is weaponization desirable? Should the international community be seeking to control and stop these processes, and what effect might that have on non-military uses of AI? In this respect we believe that the sometimes hyperbolic debate about 'killer robots' somewhat misses the point. AI is already being weaponized and the debate about banning fully autonomous weapons systems ignores much of the other weaponization processes pertaining to AI that are already in full swing. A final point for further theoretical and scholarly reflection is what role AI will play in multilateral fora such as NATO, and how the use of AI within multilateral security missions will be shared and harnessed among contributing nations. Developing common operational standards, requirements and ethical guidelines for AI-enabled capabilities through NATO's Defence Planning Process (NDPP) and Science and Technology Organization (STO), or through the EU's European Defence Fund (EDF), Coordinated Annual Review on Defense (CARD) and Permanent Structured Cooperation (PESCO), will be both necessary and challenging.[49] NATO has taken a big step forward in announcing the use of offensive cyber operations by its members to support its missions,[50] but this leads to the question of how AI will be integrated into operations such as those in Afghanistan involving dozens of allies and partners deployed to highly complex, fractured intra-state conflicts.

49 European Commission – the Independent High-Level Expert Group on Artificial Intelligence. (2019, April 8). Ethical Guidelines for Trustworthy AI. Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

50 Ricks, T. E. (2017, December 07). NATO's Little Noticed but Important New Aggressive Stance on Cyber Weapons. Retrieved from https://foreignpolicy.com/2017/12/07/natos-little-noticed-but-important-new-aggressive-stance-on-cyber-weapons/.