YRIS

The Yale Review of International Studies

ESSAYS / WINTER ISSUE

Narrow Artificial Intelligence Weapons Systems and their Impact on the Balance of Power

Posted on March 2020 by Alexandra Tsitsiringos

By:Alexandra Tsitsiringos, Tufts University

Narrow Artificial Intelligence Weapons Systems and their Impact on the Balance of Power

Research Question: How might the diffusion of military driven narrow AI weapons platforms shift the balance of power between states that invest in such platforms and states that have not invested in them?

Part 1 – Executive Summary

Artificial intelligence is one of the fastest-growing emerging technologies with important effects on current leading powers and the relationship between states. In June 2018, the Pentagon created the Joint Artificial Intelligence Center (JAIC), responsible for around 600 AI projects across the Department of Defense (DoD). One such project, named "Project Maven", aims to increase the precision of already existing weapons, such as drones, by including AI algorithms.[1] The project has already been used to identify insurgent targets in Iraq and Syria. The Joint Enterprise Defense Infrastructure (JEDI) will aid AI weaponization as it will allow the military to concentrate its data into a modern cloud platform and use machine learning to analyze information.[2] In September 2018, the Defense Advanced Research Projects Agency (DARPA) announced it would be investing up to $2 billion over the next five years into AI weapons research. These are only some steps the US has taken to achieve its AI ambitions. At the same time, major US competitors such as China and Russia are making significant strides in military AI growth. In 2017, China released a plan that included its intention to become a leader in AI development by 2030.[3] On the other hand, Russia's AI focus lies in "robotizing" the Russian Armed Forces and increasing funding for military robotics.[4] How might the diffusion of military driven narrow AI weapons platforms shift the balance of power between states that invest in such platforms and states that have not invested in them?

The nature of this research question is speculative as it attempts to peer into the impact of an emerging technology on the balance of power. Due to the recency of AI development, the traditional disputes found in international relations cease to exist and security literature is less focused on the realist versus liberal approach but more on ideas of disruption, innovation, and how impactful those can be. Thus, this paper draws from existing security studies literature and follows the trend of examining existing technologies and determining impacts based on the level and effectiveness of AI adoption. There is no certain path to understanding how consequential the diffusion of narrow AI in weapons platforms might be for state affairs, as many of these weapons platforms are in early R&D stages or do not currently exist. However, assessing the possible impact of these technologies will allow us to understand how shifts in the world order occur, to determine which states might be the key players during an AI revolution in weapons systems, and to understand the factors that might allow states to have a competitive edge. As such, the paper explores the model of public-private cooperation as a necessary factor in the development of AI weapons platforms and its potential to produce first mover advantages. It also examines how the broader application of AI will cause disruption that may lead to two outcomes: the emergence of a new leader or an AI arms race.

The remainder of this paper has three parts. The first examines the theoretical framework used to create the hypotheses that are part of this analysis and then presents those hypotheses with an overview of the evidence that points to their validity. The second part is a presentation and analysis of the empirical evidence, which is qualitative in nature. Finally, the paper includes a policy prescription portion and a conclusion on these findings.

Part 2 – Theoretical Framework:

Artificial intelligence can be divided into two subcategories: modular (or narrow) artificial intelligence and general artificial intelligence. Modular AI focuses on applications of this technology on devices that perform a specific task. An example of this type of AI is software that perfects playing scrabble; it can improve on its strategies and approaches in achieving its purpose. General AI on the other hand can be multi-purpose; it can break out of its "domain expertise and acquire generalisable intelligence".[5] Both of these types of artificial intelligence can "learn" from repeating tasks and their environment in order to improve their performance. However, they are not able to interpret meaning from data as they are not sentient, cognisant, or self-aware. This research paper will focus on modular AI and specifically its application to weapons systems, for example AI that improves precision targeting for drones or missiles.

The technological nature of AI weapons systems as well as their current stage of preliminary development create two important theoretical considerations in security studies: i. the traditional schools of thought do not dominate disagreements ii. the body of literature used is both limited and specific. While many global conflicts and changes can be examined through the traditional schools of thought in international relations, such as realism, disagreements are focused elsewhere in the discussion of AI weapons systems. One of the disagreements lies in what types of comparisons are more valuable. Some argue that it is more valuable to compare the potential impact of AI on the military to the impact of electricity or the combustion engine.[6] Others compare AI to nuclear, biotech, aerospace, and cyber applications to understand the probable effect of this technology on international security.[7] Another area of discussion is whether or not AI technologies will be disruptive; whether their contribution will be more consequential than enhancing the performance of current weapons.[8] In international relations, and particularly in security studies, AI literature is sparse and thus I draw insights from both scholarly articles as well as official government papers and news reports. I also use work of authors in historical studies to examine past patterns, as well as some literature in business management and innovation. This paper contributes to the subfield of security studies in international relations, and particularly to the debates of what influences the balance of power, whether the private sector can effectively cooperate with the government, whether innovations have disruptive or sustaining effects in military applications, and how these emerging technologies impact international order.

The paper focuses on two hypotheses in an attempt to determine the possible impact of AI weapons systems on the balance of power. First, if states follow the model of close public-private cooperation the United States DoD adopted in creating Project Maven, also known as the Algorithmic Warfare Cross-Function Team, it is probable that they will reap first-mover advantages in the development of narrow AI weapons systems. Second, if AI has a disruptive effect on the defense industry then two possibilities emerge: a transition of power or a period of intense competition between states. These hypotheses are not directly competing in the sense that they are not mutually exclusive; both of the arguments made in this paper might come true. However, one hypothesis is stronger than the other due to the availability of current evidence that points to its validity. While this paper is speculative, the more predictive nature of the second hypothesis makes it inherently more vulnerable to weakness.

In hypothesis one, the independent variable is the adoption of a model that focuses on private-public cooperation for AI development and the dependent one is the emergence of first-mover advantages in the field of AI through the early development of modular military AI. The causal mechanism that links the two is the subjective level of this model's adaptation by each state: whether states will choose to adopt this model or not at all, at what rate, and against what odds. In the realm of business and marketing, first mover advantage is defined as the advantage that a company gains by being the first to introduce a new technology, thus facing virtually no competition from other companies.[9] In the realm of security studies, first mover advantages are reaped by states that develop new technologies earlier than others, thus providing them a strategic edge.[10]

The examination of my first and preferred hypothesis is three pronged. First, I compare public-private cooperation to other factors, such as the availability of computing resources, to determine that it is the most important and necessary condition for first mover advantages in AI. Second, I examine the origins of the public-private model in the US and its success as a six-decade old system and present its current application in AI development. Third, I explore public-private cooperation and first mover advantages with regard to two major competitors: China and Russia. Overall, the purpose of this part of the paper is to show that the earliest impact of AI on the balance of power is the emergence of technological leaders when they apply a public-private model through different means.

The second hypothesis claims that if AI has a disruptive effect on the defense industry then two possibilities emerge: a transition of power or a period of intense competition between states. The independent variable is whether AI will have a disruptive or sustaining effect on the defense industry. The dependent variable is that the balance of power might be affected either by the emergence of a new leading power or by the existence of intense competition between states: an AI arms race. The causal mechanism that links the two is the extent to which AI will be adopted by each state and the mechanisms through which this adoption might occur, for example the emergence of a replication system for AI technologies. To prove this hypothesis, I begin with a presentation of innovation theory, developed by business management scholars, and a presentation of its application to defense by security scholars. Elevating this theory to nation states, I assume two possible outcomes to the balance of power: that current leaders will either maintain their power, or that their relative power will decline as new leaders emerge. To determine which will truly occur, I examine past technologies such as nuclear, aircraft, cyber, and biotechnological weapons, to establish whether or not the US maintained technological leadership. I also explore the possibility of an arms race and how regulations and the AI community can assist in "managing" such an arms race.

Part 3 – Empirical Analysis

Proving Hypothesis One

Why the Public-Private Model is Necessary

The existence of a model of public-private cooperation is necessary for first-mover advantages in the development of AI weapons systems. The very first applications of AI in technology occurred through the private sector, and specifically through industry leaders such as Google and Apple. Siri was created by SRI International and is considered one of the first applications of AI on consumer products since it was applied to Apple's iPhones in 2009.[11] There are several other examples of such integration of AI in the private sector, from simpler ones such as autocorrect to more complex ones like Google's algorithms and their ability to use specific filters to determine the most relevant results. These early AI applications demonstrate the important role of the private sector in AI research and development. Michael Horowitz et. al identify other factors that may lead to first mover advantages in AI weapons systems, which I will evaluate to show that public-private cooperation emerges either as more important than these factors or or as a necessity for their existence.

Owning large quantities of the right type of data is crucial as the most powerful machine learning techniques need large data sets to be efficient.[12] Information has perhaps become the most important commodity, and therefore governments that use larger data sets in developing AI will have a competitive edge. An important part of this calculus is that the government often relies on private companies for information. This type or reliance can range from simpler to more complex. The Snowden leaks revealed that the US government demanded disclosure of large quantities of data directly from companies or was seizing it as it moved over communications links between data centers.[13] US Law enforcement agencies continually request data from private companies and each complies to a varying level. Between January and June of 2015, Microsoft complied to 66%, Google to 78%, Facebook to 80%, and Apple to 81% of requests.[14] The US government serves as an example of how governments rely on private companies for user data, and for information in general. Gathering large sums of data for AI applications is not possible without public-private cooperation.

Training, sustaining, and enabling an AI-capable talent pool is another necessary factor according to Horowitz and his colleagues.[15] The authors state that the human capital skills required for AI development are rare as engineers that are able to create and implement current AI technologies are scarce, resulting in high private sector salaries for such talent.[16] States that are able to educate and train engineers, and develop immigration policies to allow top talent to work in their countries, will gain a competitive edge on others. However, identifying, recruiting, and training engineers are tasks that cannot be achieved without the private sector and its ability to find talent, incentivize through high payment, and prepare engineers for specific AI development. These engineers need to then be further incentivized by companies to join projects for military applications of AI, which can be hard due to moral concerns. It is clear that in the matter of human capital, public-private cooperation is inextricable from the equation as companies provide the basis for identifying and keeping top talent.

Computing resources are another necessary tool for the creation of AI weapons systems. Machine learning demands access to high level technology that is expensive. Actors with fewer resources may utilize previously trained systems or buy "off the shelf AI". States with greater resources, however, will be able to build original AI systems.[17] Computing resources are any physical or virtual components of limited availability within a computer system, including files, network connections, and memory areas. Computers, their components, as well as devices connected to them are all developed in the commercial sector. As a result, the government would benefit from creating projects under which companies such as Amazon use their own computing resources for the government's AI purposes. Without this public-private cooperation, it is likely that the government would have to purchase computing resources at a high price and under budget constraints, while the lack of cooperation with the private sector could lead to a lack of the latest or more secure technology needed for AI weapons systems. Public-private cooperation has been established as a bedrock for another one of the factors discussed by Horowitz.

Horowitz et al. include a state's willingness to act, meaning a state's willingness to adopt AI, as the final factor that may lead to having a competitive edge in AI weapons systems development.[18] The authors argue that states often prioritize other values or needs over efficiency and cite the example of health data restrictions on the grounds of privacy. While this is a valid example, it ignores the violations that constantly occur in both the government and the private sector for the sake of efficiency. Through the Snowden leaks, it was revealed that the US engaged in mass surveillance and is currently struggling with setting a higher privacy standard and reforming the NSA.[19] While the General Data Protection Regulation (GDPR) and its adoption in the EU is a positive stride in legislation to protect consumers, enforcement of these measures is still in question.[20] Evidence showing deep invasions of personal privacy, such as Uighur surveillance in China, the Snowden leaks in the US, and the shutdown of anonymous chat applications in Russia, suggests that states prioritize defense over privacy, and thus will prioritize becoming technological leaders over values. Erik Gartzke and Jo Dong-Joon discuss the proliferation of nuclear weapons and identify technological capabilities and security concerns as the key factors in nuclear weapons creation. They also mention the role of domestic politics and economic capabilities.[21] Their approach shows that there was a lack of consideration of values in nuclear weapons procurement; decisions were focused on capabilities and security concerns. Similarly, the adoption of AI weapons will likely depend on a state's capabilities and security reasons for doing so instead of its willingness to act.

Through an evaluation of the factors presented by Horowitz et. al, public-private cooperation emerges as the most necessary and important factor in development of AI weapons systems. Through cooperation, the government and the public sector can co-exist in programs such as Project Maven where the commercial sector offers its resources, data, human capital, and current knowledge in AI development. A nation's willingness to act cannot be considered a factor in itself as it depends on each actor's capabilities and security concerns.

The Contract State: Origins and AI Application

Aaron L. Friedberg offers an explanation of the origins of public-private cooperation in the United States; its historical use of contracting with industry and universities for scientific and technological developments that have military applications.[22] Harold D. Lasswell argues that the Cold War era demanded a level of military advancement that would create an elite group of political and military "specialists in violence".[23] Friedberg explains that instead of a garrison state, the Cold War led to a contract state that did not increase taxation and conscription, focused on arms and research instead of broader economic development, and relied on public-private contracts for state purposes. During the interwar period, the United States produced arms through public manufacturing in government laboratories. The private sector was solely engaged when the possibility of actual conflict arose in World War II.[24] This model could have continued if it had not been for the growth of military forces and the fact that their demands exceeded the ability of the government to supply after WWII. Therefore a new model of public-private cooperation emerged, which demanded "a steadier, more continuous relationship with private industry".[25] Since the 1950s, research and weapons spending has been stable as research often occurs in the same companies that build the end products.[26] This new relationship between companies and the government was complemented by a shift in strategy that began in the early 1960s: deterrence through the preservation of technologically advanced forces.[27] Judging by US supremacy in WWII and the Cold War, public-private cooperation is a successful, six-decades old national security system that is in place to this day. While Friedberg provides an understanding of the basis of this model and its early existence, it is important to examine how this model of public-private cooperation is applied to AI development.

Public-private cooperation for AI development in the United States is strong but not without challenges. In November 2018, the Project Maven team hosted technology companies in Maryland, where the government viewed private demonstrations. Large tech companies such as Intel, IBM, GE, Oracle, as well as defense company Raytheon, were among the 42 businesses that expressed interest in "showing off" their AI for the military.[28] It is unclear whether Microsoft or Amazon are currently participating in Project Maven. Google, a single company out of many technology innovators in the US, may serve as a helpful study case for the potential problems that will arise in the current cooperation model. After Google's initial participation at the Pentagon's Drone AI Imaging Program, about 4,000 Google employees signed a petition demanding "a clear policy stating that neither Google nor its contractors will ever build warfare technology".[29] The company complied with its employees' wishes and did not renew its Pentagon contract. This incident shows that the extent to which companies listen to their employees' concerns and act on them is an important factor in whether a company's participation in government contracts is possible. Another consideration is that the petition allows for the separation of Google employees into two distinct groups: those that do not want to participate in any sort of military project and those that have specific concerns over the lethality and morality of facilitating the creation of AI weapons systems. At Google, there seems to be an overall lack of desire to cooperate with the government as it opted out of taking part in the Pentagon's JEDI Cloud Contract, while the CEOs of Microsoft and Amazon stood by the $10 billion contract.[30] Challenges to the participation of industry in military development are not new; Friedberg explains that during the interwar period the public's anger towards private weapons manufacturers was one of the factors that led weapons R&D to occur mostly through federal efforts.[31] In spite of such challenges, public-private cooperation for weapons procurement exists to this day. This analysis of news sources regarding companies and their levels of participation shows that while the platform for public-private cooperation remains, its strength might be in question due to mixed positions in employee willingness to participate in defense and AI projects.

Public – Private Cooperation and Major Competitors: Russia and China

The first part of this analysis focused on singling out public-private cooperation as a necessity for first mover advantages while the second part showed how this model can be successful through an examination of its origins along with its current application to AI development in the US. This gives rise to the question: is the US model the only way? Do other nations need to become "contract states" and refine such a system over decades? I would argue that China and Russia provide an example of how other states use their own public-private model for success in AI development. More specifically, the authoritarian and corporatist nature of the Chinese and Russian governments has allowed them to rise to a level of significant competition with the United States.

Government involvement in the Chinese private sector has provided China with first mover advantages comparable to those of the United States. Both the United States and China have moved ahead of the rest of the world due to their strong technology industries. On an international scale, there are roughly 4,500 companies involved in AI development, with half of them active in the US and one third active in China.[32] By 2030, it is expected that these two countries will capture 70% of the 15.7 trillion AI is likely to generate in global markets.[33] China is expected to rise as an international AI power, with Beijing as an innovation center at the level of Silicon Valley[34]. Government documents echo these expected outcomes since The "New Generation AI Development Plan" outlines China's aims to catch up on AI technology and applications by 2020, achieve major breakthroughs by 2025, and become a global leader in AI by 2030.[35] Ryan Hass and Zach Balin attribute part of both countries' success to the highly competitive innovation systems within their private sectors.[36] Overall, China benefits from a regulatory environment that fosters AI development through "unparalleled government support" in the commercial sector. However, it is important to consider that there are structural disadvantages that occur from the government's regime and accelerated involvement.[37] Chinese firms are often pushed to develop products that aim to support the Communist Party and its efforts. The country has also insisted on a principle of self-sufficiency. Due to this policy, China suffers from a lack of high-level technologies from abroad and a lack of cooperation with important global players in key data sectors. In the future, it is likely that the government will demand that technological components originate solely from China instead of other countries. Similar to the moral hesitations of employees in the US, ethical concerns could arise if China applies AI to intrusive surveillance or targeted repression. Both the US and China have become AI frontrunners in an intense bilateral rivalry that many have compared to the Cold War, which has invoked criticism by those who suggest that instead of creating a rift between the two countries, AI development should invite a healthy level of competition.[38] While China's structural environment has allowed it to become a frontrunner in AI development, the weaknesses that result from the Communist Party's involvement should not be ignored.

Russia's AI development is also more government focused than in the US, with the Russian Department of Defense taking the lead and industry having an assisting role. The lesser role of the private sector in this regard contributes to the less powerful innovation ecosystem in Russia, which has led to its rank below the US and China in AI development. While private investment in AI is expected to increase to $500 million by 2020, the current $12.5 million commercial and federal spending is well below Chinese and US efforts[39]. These efforts have invited ethical concerns, and in a statement to the UN Group of Government Experts on Lethal Autonomous Weapons Systems, Russia echoed those concerns stating that there are "serious ethical, legal, operational, technical challenges raised by these weapons."[40] Russia's Foundation for Advanced Studies is working on AI weapons systems such as image recognition software and systems that imitate human thought process.[41] Russia is also planning to use AI in information warfare, on par with its recent efforts to spread fake data, and to build AI weapons systems such as a combat weapon equipped with a machine gun that uses neural network technologies for target identification and decision-making. The country has been focusing on AI applications on robotics, with claims that there is already a super tank with an autonomous turret, slowly leading to the creation of fully autonomous tanks. The commander in chief of the Russian Air Force has confirmed that AI-guided missiles are in early stages of development.[42] Russia also aims to create a nuclear delivery vehicle in the form of an autonomous underwater vehicle: Status 6.[43] The pace of AI R&D in the country seems to be on par with Vladimir Putin's statement that "Artificial intelligence is the future, not only for Russia but for all humankind…Whoever becomes the leader in this sphere will become the ruler of the world."[44]

The case studies of China and Russia demonstrate that while the US "contract state" model has contributed to the country's frontrunner status, other public-private models can have comparable levels of success. Government involvement in the private sector is inherent to the regime structure of these two competitors that rank right below the US in AI development. However, an examination of other states that have made strides in AI weapons procurement could alter these findings as the consideration of three cases is by no means exhaustive.

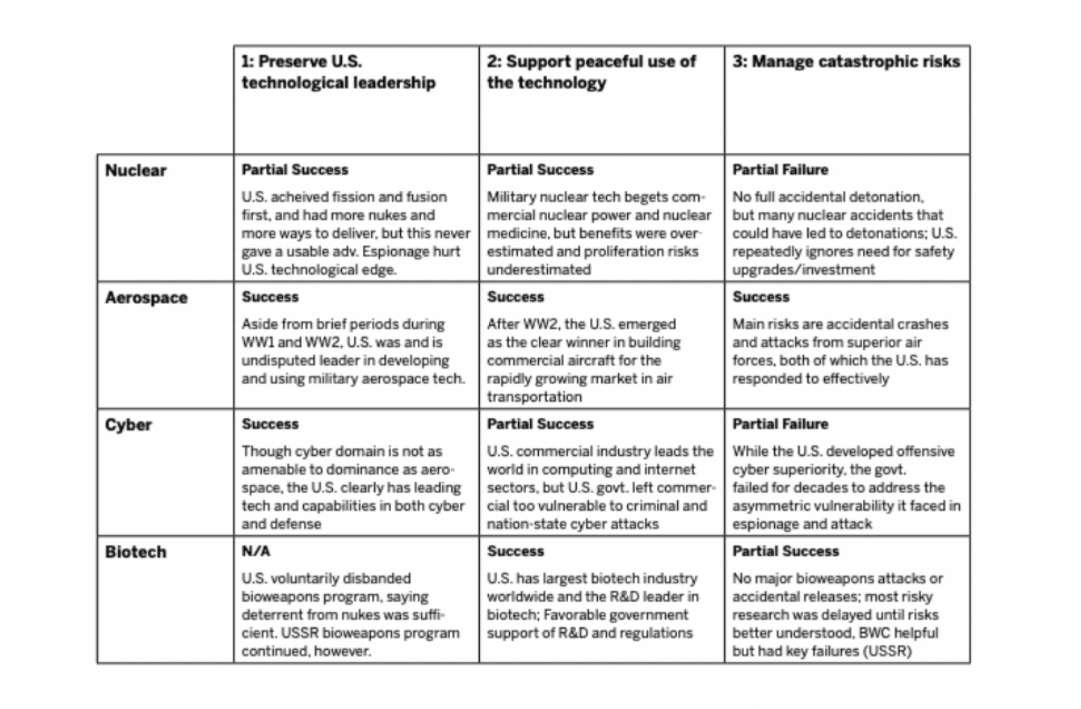
Proving Hypothesis Two

Innovation Theory, The Defense Industry, and Nation States

This part of the paper utilizes two theoretical frameworks that arise from Christensen's Innovation Theory and applies them to nation states to determine whether AI might result in a new global power or enhance the status of current leaders. Clayton M. Christensen developed a hypothesis regarding businesses and potentially disruptive innovations, according to which incumbent firms produce sustaining innovations but rarely make disruptive innovations – thus allowing new entrants to dethrone established market leaders.[45] If applied to the defense industry, this theory would imply that established firms would stop contracting with the government due to their focus on more obsolete technologies or older products, thus losing ground to newer companies and startups. Peter Dombrowski and Eugene Gholz challenged this hypothesis, claiming that "new technology for military communications mostly requires sustaining innovation."[46] The weakness in Christensen's theory is that transformational technologies for the military are not low cost or low quality products but high-end innovations. The two authors examine the cases of Littoral Combat Ships (LCS) and networks in military communications to prove that established prime contractors remain. In some cases, new entrants provide innovative processes and technologies to these established defense companies – a win-win scenario[47]. In short, they show that military transformation does not require a new group of suppliers. I propose an extension of these different assumptions to the nation-state level. If Christensen's theory is valid, surveying past disruptive technologies would uncover that the balance of power changed to favor newer powers. If the Dombrowski-Gholz theory is correct, leading nation-states adopted innovation and remained in power. Apart from presenting the outcomes created by past innovative technologies, I will also include a predictive component about the impact of AI.

Past Technologies and US Success Levels – AI and its Potential

Allen Greg and Taniel Chan provide insights into the varying levels of US success with regards to nuclear, aerospace, cyber and biotech weapons.[48] United States success with these transformative technologies are evaluated with regards to a. the preservation of US technological leadership b. the support of peaceful use the technology c. the ability to manage catastrophic risks.



	1: Preserve U.S. technological leadership	**2: Support peaceful use of the technology**	**3: Manage catastrophic risks**
Nuclear	**Partial Success** U.S. acheived fission and fusion first, and had more nukes and more ways to deliver, but this never gave a usable adv. Espionage hurt U.S. technological edge.	**Partial Success** Military nuclear tech begets commercial nuclear power and nuclear medicine, but benefits were overestimated and proliferation risks underestimated	**Partial Failure** No full accidental detonation, but many nuclear accidents that could have led to detonations; U.S. repeatedly ignores need for safety upgrades/investment
Aerospace	**Success** Aside from brief periods during WW1 and WW2, U.S. was and is undisputed leader in developing and using military aerospace tech.	**Success** After WW2, the U.S. emerged as the clear winner in building commercial aircraft for the rapidly growing market in air transportation	**Success** Main risks are accidental crashes and attacks from superior air forces, both of which the U.S. has responded to effectively
Cyber	**Success** Though cyber domain is not as amenable to dominance as aerospace, the U.S. clearly has leading tech and capabilities in both cyber and defense	**Partial Success** U.S. commercial industry leads the world in computing and internet sectors, but U.S. govt. left commercial too vulnerable to criminal and nation-state cyber attacks	**Partial Failure** While the U.S. developed offensive cyber superiority, the govt. failed for decades to address the asymmetric vulnerability it faced in espionage and attack
Biotech	**N/A** U.S. voluntarily disbanded bioweapons program, saying deterrent from nukes was sufficient. USSR bioweapons program continued, however.	**Success** U.S. has largest biotech industry worldwide and the R&D leader in biotech; Favorable government support of R&D and regulations	**Partial Success** No major bioweapons attacks or accidental releases; most risky research was delayed until risks better understood, BWC helpful but had key failures (USSR)

For the purposes of this paper, the most important factor is the first since it pinpoints whether a leader remains in power technologically. The U.S. partially succeeded in its preservation of leadership with nuclear weapons, since it achieved fission and fusion first, had more weapons than other states, and more ways to deliver them. However, this never provided a usable advantage and espionage hurt U.S. advancement, making *this a partial success*. In terms of aerospace, the US *succeeded* since aside from small periods during WWI and WWII, the US was and is the undisputed leader in developing and using military aerospace technology. The US also *succeeded* in cyberspace as it has leading technology and capabilities in both cyber and defense. It is important to note that success levels are not as strong as in the aerospace domain, especially because the US has not faced its vulnerability to attack and espionage at an appropriate level. The level of success in biotechnology *cannot be assessed* because the US voluntarily disbanded its bioweapons programs, claiming that nuclear weapons were a sufficient deterrent against the USSR – even though the latter continued bioweapons development. An examination of these four disruptive technologies suggests that the US succeeded as a leading power to maintain technological and defensive leadership, but not without challenges. Taking the US case study as the only source of evidence, it seems that when elevated to the nation state level the Dombrowski-Gholz hypothesis holds true. I recognize that surveying the status of other leading powers such as Russia and China with regards to such weapons would be useful and constitutes a weakness in this analysis as this theory could be proven wrong. AI has not yet been adopted to its full potential and most AI weapons are currently in development, but a comparison with these past technologies allows us to understand which innovation hypothesis holds true.

The most important caveat when it comes to AI technologies is their potential for replication, which can lead to unexpected world order outcomes. Replication may seem impossible initially, given that high-level technologies for machine learning and AI are costly with companies spending millions or billions on R&D. However, small groups can use open source code libraries and commercial off the shelf software, as well as rented hardware, to develop powerful AI technologies or weapons for less than one million dollars.[49] Greg and Chan suggest that leaked copies of AI software will be "virtually free". Similar implications may arise with regards to complexity. To conduct basic AI research and reach initial advancements, states need to recruit world-class talent from a highly limited pool. Once fundamental research exists, applying it to more specific, smaller problems can be more straightforward and solved without top talent.[50] Furthermore, conversion of commercial AI tech to military systems requires high levels of technical expertise but as our understanding of AI improves, a decline in these needs is probable. Through this analysis of capabilities and talent, the potential to replicate AI in the future is high, which creates the possibility for other nation states or even non-state actors that are not current front-runners to advance fast or even catch up. This type of "skipping ahead" in the R&D process mirrors current Chinese and Russian efforts to engage in economic cyber-espionage against the US. According to the Office of the Director of National Intelligence, their "efforts compromise intellectual property, trade secrets, and technological developments that are critical to national security.. [and] espionage against the private sector increases the danger to long-term U.S. prosperity".[51] The high destructive potential of AI and its vulnerability to espionage and monitoring are also risk factors to current AI leaders, namely the US, China, and Russia.[52] According to Greg and Chan, it seems that aerospace technology can be considered similar to AI applications, given that it became almost synonymous with military power.[53] Businesses do not have a choice in whether to adopt machine learning, simply because not doing so would result in competitive losses. In an analogous vein, militaries and intelligence agencies might expand their military AI for fear of other countries gaining an advantage.

Overall, the ability of countries to copy AI military weapons systems may have unpredictable results for the balance of power since weaker states or non-state actors could gain advantages. It is also possible that AI will fit into the pattern of past technologies mentioned in this paper, which means that current powers will remain leaders. The predictive nature of this paper does not allow for a conclusive argument, however, the competitiveness introduced by this emerging technology begs the question: will AI cause an arms race? The first arms race was the pre-WWI naval arms race followed by the nuclear arms race during the Cold War, and an AI arms race could follow suit.

An AI Arms Race?

Security literature on the history of AI weaponization and the current competition between suggests that an arms race is likely, with researchers focusing on how it can be managed instead of questioning whether it will occur. Edward Moore Geist summarized this pattern by claiming that our choice is between "a well-managed AI arms race that reinforces mutual security and a poorly managed one that could lead to disastrous outcomes".[54] The US first developed weapons to engage targets without human input through acoustic homing torpedoes that were used in WWII[55]. During the 1960s, DARPA started funding AI R&D in the United States. At the same time, the USSR commenced research into "voennaia kibernetika", military cybernetics.[56] Thus, Geist identifies the origins of an AI arms race at the Cold War, though today the competition between China, Russia, and the US is more intensified due to advancements in the field. Greg and Chan mirror Geist's logic by pointing out that arms races may be unavoidable, but they can also be managed.[57] An AI arms race is unavoidable primarily because of how useful AI is proving to be, along with the idea that there is vast unlocked potential for military applications. Other authors, such as Ben Tarnoff, suggest that apart from an arms race, AI will allow for "algorithmic forever wars".[58] The War on Terror that started after 9/11 is still ongoing, and is characterized by its unconventional enemies and lack of set boundaries or battlefields. This setup makes the question of who to target the most important one, the vagueness of the adversary being a factor in how prolonged this war is. AI has the potential to extend this war to an unending period of time, given that it will permit the US to see "enemies everywhere" depending on the hostile behavior pattern identified by machine learning.[59] This potential is strengthened by adding more players that would gain from an extended conflict, as Silicon Valley companies profit from their inclusion in military projects. Tarnoff points out that "the problem isn't the quality of the tools but the institution wielding them" and Geist that "human foolishness" rather than automation is the issue – indicating that danger lies in human behavior rather than weapons, and opening grounds for the discussion of AI regulation.[60]

Policy Prescription: Managing an AI Arms Race

Regulations on previous transformative technologies, and particularly aircraft and biotechnological weapons, are helpful in forming a blueprint for AI regulation. In 1899 during a peace conference in the Hague, diplomats decided on a five-year moratorium on all offensive military uses of aircraft. At the second conference in 1907, the same agreement that was supposedly going to be permanent was abandoned as states realized the potential of airspace battles. As a result, during WWI, multiple capitals were bombed from the air and thousands of civilians suffered. Greg and Chan consider AI applications similarly irresistible and suggest that regulations to completely ban AI will be fruitless. Instead, they recommend the adoption of a framework similar to the one that limited the risks of aerospace technology.[61] On the other hand, Geist explains that AI used for monitoring purposes can be as dangerous as weapons that target individuals. AI-controlled undersea drones may make the seas "transparent", rendering missile-carrying submarines effectively unusable, an advancement that might have unpredictable and significant geostrategic consequences.[62] The first step in AI regulation is for researchers to agree not to contribute to AI applications with undesirable social consequences, like biotechnology researchers did at the 1975 Asilomar Conference on Recombinant DNA. Given that the USSR continued to develop biological weapons despite singing the 1972 Biological Weapons Convention, it seems that norms are necessary but insufficient.[63] There is a three-pronged approach researchers can take to prevent catastrophic consequences of an AI arms race. There should be a focus on *verification* – the process through which states determine whether other states comply with arms-control agreements. The bioweapons convention of 1972, included no meaningful verification measures. In the nuclear realm, the SALT I treaty set limits on nuclear arsenals that both the US and the USSR could verify through reconnaissance satellites, an example AI researchers and policymakers should emulate. This process is hard because it is both technical and political, and AI complicates it further given that software can be developed domestically with appropriate data and hardware resources. The AI community can also create global *monitoring mechanisms* to push for stronger arms control measures. They can also use *Track II Diplomacy*, a type of unofficial channel for bargaining, by engaging with fellow AI researchers in enemy states.[64] Regulations on biotechnology, aircraft, and nuclear weapons can provide valuable insight on how to craft policy to restrict the risks of AI military applications.

The history of AI weaponization and the currently intense competition between states leads to the conclusion that an AI arms race is possible. At the same time, AI may allow for more "forever wars" like the US War on Terror through its speed in identification and precision. Military aircrafts show how emerging weapons can be irresistible to nation states, while biotechnology and nuclear weapons regulations point to norms and practices that can help in "managing" an AI Arms Race.

Part 4 – Conclusion

"Artificial intelligence is the future, not only for Russia but for all humankind… Whoever becomes the leader in this sphere will become the ruler of the world[65]."

- Vladimir Putin

This paper evaluated how states can reap first mover advantages in AI development, and how future advancements point to an unavoidable arms race. The first part of my analysis showed how public-private cooperation is the most necessary factor in AI development, pointed to a six-decade old model of such cooperation in the United States, and discussed the status of the US and its major adversaries, Russia and China, with regards to narrow AI military applications. These findings point to first mover advantages through public-private cooperation, regardless of how that cooperation occurs. The second part of the paper examined innovation theory and its application to defense and elevated it to the nation state level with two possible outcomes: new leaders emerging or current powers remaining. Through an examination of past technologies, the most likely outcome appears to be that current leaders will remain in power as AI advances, but not without challenges. The examination of past technologies also gives rise to an AI arms race, which is considered by scholars as probable but also manageable through appropriate regulations. My first hypothesis that public-private cooperation will lead to first mover advantages seems more concrete, given the fact that we already know that the US, China, and Russia are leaders in artificial intelligence development. The evidence presented in this paper confirms this hypothesis. The second hypothesis posits that the disruptive nature of AI might lead to a new leader or an arms race. Historic evidence on aerospace, nuclear and cyber weapons suggests that leaders remain after new technologies emerge, which partly disproves my hypothesis that a transition of power might occur. However, it has not been completely ruled out since AI weapons have not been developed fully and their ability to be replicated might have unpredictable results. Research suggests that intense competition between states is currently occurring and will continue, but the scale of conflict might not reach the levels that would warrant its categorization as an "arms race" like the pre-WWI naval and the Cold War nuclear arms races. The first part of this hypothesis is weakened but not completely eliminated through the evidence, while the arms race component is proven relevant but not confirmed. Overall, the first hypothesis has been proven stronger than the second. There are many faults with modern AI technology, which can already be observed in private sector applications such as self-driving cars. AI weapons may have even more catastrophic unintended consequences if left unchecked. Thus, drawing on past technologies and their impact on the balance of power as well as their successful regulation frameworks should be the highest priority of academics, researchers, and policymakers involved in AI weapons systems.

[1] Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven). Deputy Secretary of Defense. April 26, 2017. https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.

[2] Tarnoff, Ben. "Weaponized Artificial Intelligence is Coming. are Algorithmic Forever Wars our Future?" *Turning the Tide* 30, no. 5 (Nov, 2018): 7.

[3] China State Council, "A Next Generation Artificial Intelligence Development Plan," July 20, 2017, translated by New America, https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf,

[4] Samuel Bendett, "Red Robots Rising: Behind the Rapid Development of Russian Unmanned Military Systems," *The Strategy Bridge*, December 12, 2017, https://thestrategybridge.org/the-bridge/2017/12/12/red-robots-rising-behind-the-rapid-development-of-russian-unmanned-military-systems.

[5] Ayoub, Kareem & Kenneth Payne. "Strategy in the Age of Artificial Intelligence". Journal of Strategic Studies (November 2015) 39:5-6, 793-819, DOI: 10.1080/01402390.2015.1088838, pages: 795-796

[6] Horowitz, Michael, Elsa B. Kania, Gregory C. Allen and Paul Scharre. *Strategic Competition in an Era of Artificial Intelligence*. Center for a New American Security, 2018. Page: 3

[7] Allen, Greg and Taniel Chan. *Artificial Intelligence and National Security*. Belfer Center for Science and International Affairs, 2017

[8] Ayoub, Kareem & Kenneth Payne. "Strategy in the Age of Artificial Intelligence". Journal of Strategic Studies (November 2015) 39:5-6, 793-819, DOI: 10.1080/01402390.2015.1088838

[9] “First-Mover Advantage”. Cambridge Dictionary. https://dictionary.cambridge.org/us/dictionary/english/first-mover-advantage

[10] Silverstein, Andrew Bernard, “Revolutions In Military Affairs: A Theory On First-Mover Advantage” 01 April 2013. CUREJ: College Undergraduate Research Electronic Journal, University of Pennsylvania, http://repository.upenn.edu/curej/169.

[11] Naone, Erica. “TR10: Intelligent Software Assistant”. MIT Technology Review. March/April 2009. http://www.morgenthaler.com/press-releases/Siri%20Named%20Top%2010%20Emerging%20Tech%20of%202009.pdf

[12] Horowitz et al, 5

[13] Cate, Fred H and James X. Dempsey Bulk Collection: Systematic Government Access to Private-Sector Data. Oxford Scholarship Online. (2017) DOI: 10.1093/oso/9780190685515.001.0001. Page: 308

[14] Joon Ian Wong. “Here’s how often Apple, Google, and others handed over data when the US government asked for it”. Quartz. February 19, 2016.

https://qz.com/620423/heres-how-often-apple-google-and-others-handed-over-data-when-the-us-government-asked-for-it/

[15] Horowitz et al, page: 5

[16] Cade Metz, “Tech Giants Are Paying Huge Salaries for Scarce A.I. Talent,” *The New York Times*. October 22, 2017, https://www.nytimes.com/2017/10/22/technology/artificial-intelligence-experts-salaries.html

[17] Horowitz et al, page: 6

[18] Horowitz et al, page: 6

[19] Edgar, Timothy H. Beyond Snowden: Privacy, Mass Surveillance, and the Struggle to Reform the NSA”. Brookings Institution Press, August 29 2017. https://www.brookings.edu/book/beyond-snowden/

[20] “Chapter 1 – General Provisions.” *General Data Protection Regulation* (GDPR), gdpr-info.eu/chapter-1/.

[21] Dong-Joon, Jo and Erik Gartzke. “Determinants of Nuclear Weapons Proliferation”. *The Journal of Conflict Resolution*

Vol. 51, No. 1 (Feb., 2007), pp. 167-194 (28 pages).

[22] Friedberg, Aaron L. “Why Didn’t the United States become a Garrison State?”

International Security. Vol. 16, No. 4 (Spring, 1992), page: 114

[23] Friedberg, page: 112

[24] Friedberg, page: 136

[25] Friedberg, page: 137

[26] Friedberg, page: 140

[27] Friedberg, page: 117

[28] Gershgorn, Dave and Justin Rohrlich. "These companies are pitching AI to the US military". Quartz, November 13, 2018. https://qz.com/1461910/these-companies-are-pitching-ai-to-the-us-military/

[29] Shane, Scott and Daisuke Wakabayashi. "Google Will Not Renew Pentagon Contract That Upset Employees". The New York Times, June 1st 2018. https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html

[30] Konkel, Frank. Microsoft, Amazon CEOs Stand By Defense Work After Google Bails on JEDI. NextGov, October 15, 2018. https://www.nextgov.com/it-modernization/2018/10/microsoft-amazon-ceos-standby-defense-work-after-google-bails-jedi/152047/

[31] Friedberg, page: 137

[32] Hass, Ryan, and Jack Balin. "US-China Relations in the Age of Artificial Intelligence." *Brookings*, January 10, 2019.

[33] "*AI to drive GDP gains of $15.7 trillion with productivity, personalisation improvements*". PriceWaterhouseCoopers June 27, 2017. https://press.pwc.com/News-releases/ai-to-drive-gdp-gains-of–15.7-trillion-with-productivity–personalisation-improvements/s/3cc702e4-9cac-4a17-85b9-71769fba82a6

[34] Lee, Kai – Fu and Paul Triolo. China's Artificial Intelligence Revolution: Understanding Beijing's Structural Advantages. Eurasia Group and Sinovation Ventures, December, 2017. https://www.eurasiagroup.net/files/upload/China_Embraces_AI.pdf

[35] China State Council, "A Next Generation Artificial Intelligence Development Plan," July 20, 2017, translated by New America, https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf,

[36] Hass, Ryan, and Jack Balin

[37] Hass, Ryan, and Jack Balin

[38] Hass, Ryan, and Jack Balin

[39] Samuel Bendett, "In AI, Russia Is Hustling to Catch Up," Defense One, April 4, 2018, https://www.defenseone.com/ideas/2018/04/russia-races-forward-ai-development/147178

[40] "Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects," November 10, 2017, https://admin.govexec.com/media/russia.pdf.

[41] Samuel Bendett, "Red Robots Rising: Behind the Rapid Development of Russian Unmanned Military Systems," The Strategy Bridge, December 12, 2017, https://thestrategybridge.org/the-

bridge/2017/12/12/red-robots-rising-behind-the-rapid- development-of-russian-unmanned-military-systems.

[42] Tom O'Connor, "Russia's Military Challenges U.S. and China By Building a Missile That Makes Its Own Decisions," Newsweek, July 20, 2017, http://www.newsweek.com/russia- military-challenge-us-china-missile-own-decisions-639926.

[43] Valerie Insinna, "Russia's nuclear underwater drone is real and in the Nuclear Posture Review," January 12, 2018, https://www.defensenews.com/space/2018/01/12/russias-nuclear- underwater-drone-is-real-and-in-the-nuclear-posture-review/

[44] "Putin: Leader in artificial intelligence will rule world," Associated Press, September 4, 2017, https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html.

[45] Dombrowski, Peter and Eugene Gholz."Identifying Disruptive Innovation Innovation Theory and the Defense Industry". *MIT Press Journals.* (Spring, 2009) page: 102

[46] Dombrowski, Peter and Eugene Gholz, page: 112

[47] Dombrowski, Peter and Eugene Gholz, page: 114

[48] Allen, Greg and Taniel Chan. *Artificial Intelligence and National Security.* Belfer Center for Science and International Affairs, 2017, page: 45

[49] Allen, Greg and Taniel Chan. page: 46

[50] Allen, Greg and Taniel Chan. page: 47

[51] "Economic Espionage". Office of the Director of National Intelligence. https://www.dni.gov/index.php/ncsc-what-we-do/ncsc-threat-assessments-mission/ncsc-economic-espionage

[52] Allen, Greg and Taniel Chan. page: 46 – 47

[53] Allen, Greg and Taniel Chan. page: 50

[54] Geist, Edward Moore. "It's already too late to stop the AI arms race—We must manage it instead". Bulletin of the Atomic Scientists (August 2016), 72:5, 318-321, DOI: 10.1080/00963402.2016.1216672, page: 318

[55] Ibid.

[56] Geist, Edward Moore, page: 3

[57] Allen, Greg and Taniel Chan. page: 49

[58] Tarnoff, Ben

[59] Ibid.

[60] Geist, Edward Moore. Page: 319

[61]Allen, Greg and Taniel Chan. page: 51

[62] Allen, Greg and Taniel Chan. page: 51

[63] Geist, Edward Moore. Page: 320

[64] Ibid

[65] "Putin: Leader in artificial intelligence will rule world," Associated Press, September 4, 2017, https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html.

featured, latest, recent

Alexandra Tsitsiringos

View all posts by Alexandra Tsitsiringos →