I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

"What is an 'Artificial Intelligence Arms Race' Anyway?"

DR. PETER ASARO*

CONTENTS

I. INTRODUCTION

There has been a great deal of recent media coverage and discussion in international diplomatic circles of a coming (or rapidly accelerating) global artificial intelligence (AI) arms race. The phrase reached prominence in the U.S. and European media flurry following an *Associated Press* report on a speech by Vladimir Putin on September 1, 2017.[1] In that speech, an "Open Lesson" broadcast to

*Associate Professor, School of Media Studies, The New School; Visiting Professor, Munich Center for Technology in Society, TU Munich; Affiliate Scholar, Center for Internet and

over one million Russian schoolchildren on their first day of school, he said that "The one who becomes the leader in this sphere will be the ruler of the world."[2] While these comments could be argued to have been taken out of context—Putin continues by saying that Russia would share AI knowledge rather than monopolize it as they share nuclear technology—it is clear that the media was ready to quickly extrapolate these comments as acknowledging a previously unspoken arms race for AI. *Wired* ran their story on the September speech with the headline "For Superpowers, Artificial Intelligence Fuels New Global Arms Race,"[3] while just two months later *CNN* ran a story with the headline "US Risks Losing Artificial Intelligence Arms Race to China and Russia,"[4] and the *World Economic Forum* ran an article entitled "Artificial Intelligence is now an Arms Race. What if the bad guys win?"[5]

There is surely a complex story as to why the media was so readily primed for such sensationalism, one which involves Silicon Valley companies making massive investments into AI; the rising global

---

Society, Stanford Law School. This paper was originally presented at the National Security, Emerging Technologies and the Law Conference, American Bar Association SCOLANS, Moritz College of Law, The Ohio State University, Columbus, OH, March 23, 2018.

[1] *Putin: Leader in Artificial Intelligence Will Rule World*, AP NEWS (Sep. 1, 2017), https://www.apnews.com/bb5628f2a7424a10b3e38b07f4eb90d4 [https://perma.cc/CP73-5Z7H]; James Vincent, *Putin Says the Nation That Leads in AI 'Will Be the Ruler of the World'*, The Verge (Sep. 4, 2017), https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world [https://perma.cc/72SD-SAQQ]; President Vladimir Putin, '*Whoever leads in AI will rule the world': Putin to Russian children on Knowledge Day*, RT (Sep. 1, 2017), (transcript available at https://www.rt.com/news/401731-ai-rule-world-:putin/ [https://perma.cc/GF6S-XADG].

[2] *Id.*; President Vladimir Putin, *National Open Lesson Russia Focused on the Future*, (Sep. 1, 2017), (transcript available at http://en.kremlin.ru/events/president/transcripts/55493 [https://perma.cc/N3BX-9Q74]).

[3] Tom Simonite, *For Superpowers, Artificial Intelligence Fuels New Global Arms Race,* WIRED (Sep. 8, 2017), https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race/ [https://perma.cc/T4KL-LANN].

[4] Zachary Cohen, *US Risks Losing Artificial Intelligence Arms Race to China and Russia*, CNN (Nov. 29, 2017), https://www.cnn.com/2017/11/29/politics/us-military-artificial-intelligence-russia-china/index.html [https://perma.cc/D37Q-33LB].

[5] Mark Hughes, *Artificial Intelligence is Now an Arms Race. What if the Bad Guys Win?*, WORLD ECONOMIC FORUM (Nov. 10, 2017), https://www.weforum.org/agenda/2017/11/cybersecurity-artificial-intelligence-arms-race/ [https://perma.cc/H9AK-DT96].

competition for AI talent; a similar resurgent interest in AI at the Pentagon; the rise of international cyberattacks and the serious risks posed by cyberwarfare; the steady stream of revelations regarding Russia's use of social media to interfere in the U.S. elections as well as those of European states; and the rise of China as a serious economic competitor in the application of AI to Big Data.

But what do we really mean when we talk about an "AI Arms Race?" And what is at stake in "losing" such an arms race? This paper will examine several ways of interpreting this phrase, drawn from various definitions of both "arms races" and "AI"– both of which lack a precise technical definition applicable to the kind of broad global phenomena referred to by media coverage. A better understanding of what this means, or might mean, can help both separate the hype from the real issues, as well as help us to focus on what aspects of AI pose serious risks to national security in the U.S. and elsewhere.

Towards that end, the paper will start by reviewing some of the most common meanings of the terms "arms race" and "AI," and then review several of the most likely meanings behind the phrase "AI arms race" in its current usage. It is the conclusion of this paper that the phrase combines multiple meanings, some of which may be contradictory, and we should really focus on the separate underlying concerns, rather than adopt an umbrella term that confuses rather than clarifies the issues.

II. WHAT IS AN "ARMS RACE"?

The traditional notion of an "arms race" has no widely agreed upon definition, but is used in international relations to refer to a competition between two or more states in the development and production of military arms.[6] This is generally understood to take the form of one-upmanship in which each side seeks an incremental gain over the others in terms of either technical superiority (quality) and/or number (quantity) of weapons over their competitors. The sought-after advantage in military superiority can be an overall strategic advantage, or a tactical advantage in an area with strategic significance. The consequences of an arms race are generally viewed as negative insofar as it is at best expensive and self-defeating, and at

[6] DAVID ATKINSON, ARMS RACES, Oxford Bibliographies (database updated Mar. 2, 2011), http://www.oxfordbibliographies.com/view/document/obo-9780199743292/obo-9780199743292-0002.xml [https://perma.cc/HGT2-3CCF].

worst destabilizing, counter-productive, or even catastrophic.[7] But there remains some debate as to whether arms races are the consequence of geo-political instability or a cause of it. It could be argued that both views are correct insofar as both are involved in a positive feedback loop, or vicious cycle, of increasing instability and increasing arms investments, which serve to drive each other.

A more technical definition of an "arms race" comes from game theoretic analysis. In applying game theory to international relations, one can frame arms races as a form of the prisoner's dilemma. That is, it is actually in the (economic) interest of each state to not spend additional resources on arms, yet failing to do so in the event that their competitor does invest in arms has a large downside. In the absence of cooperation, both sides acting in their own narrow self-interest are worse off than they would be if they could cooperate (such as forming a treaty or otherwise trusting each other not to build up arms).

Evolutionary biology has also adopted and applied the concept of an "arms race." Starting with Van Valen's *Red Queen Hypothesis*,[8] biologists have sought to explain the evolution of species' traits through their competitions with other species. In the original hypothesis, the incremental advantages gained by a species over time through natural selection tend to be matched by the gains made by other competing species in their ecosystem. Thus, even though rabbits evolve to be faster, as the slower ones are less likely to reproduce, the foxes who prey on them also evolve to become faster. Thus, the traits of both species can evolve in relation to each other, even while the population sizes remain relatively stable over time. The original idea of the *Red Queen Hypothesis* is taken from *Alice Through the Looking* Glass, where in her chess game with the Red Queen, Alice had to keep running just to stay in place.[9] Sometimes these evolved traits follow a single path, leading to extreme biological forms. And so it can be with species getting ever fitter to their environment, while their ecological competitors similarly co-evolve resulting in extreme traits and little net gain. It is interesting to reapply the Red Queen hypothesis to military arms races, as it suggests that even as competing states invest

---

[7] Samuel P. Huntington, *Arms Races: Prerequisites and Results*, PUBLIC POLICY, vol. 8 (1958) pp. 41-86.

[8] Leigh Van Valen, *A New Evolutionary Law*, 1 EVOLUTIONARY THEORY 1, 21 (1973).

[9] LEWIS CARROLL, ALICE THROUGH THE LOOKING GLASS 11-12 (17th ed. 1994).

heavily in arms for incremental advantages, over time the relative military advantages disappear.

III. WHAT IS "ARTIFICIAL INTELLIGENCE"?

The definition of artificial intelligence ("AI") has been controversial for decades. Initially, when John McCarthy coined the term in 1956, it was meant to describe the work of a varied group of researchers. The common thread of their work was developing computer programs to perform tasks that were generally assumed to require human intelligence. Over time, this proved to be something of a moving target; as computers regularly achieved new performances, the scope of what requires human intelligence has shifted. As the algorithms developed by AI researchers became widely used, they were simply treated like other software techniques rather than being thought of as a special category of AI.[10]

Still, as a scientific and engineering discipline, AI has an identifiable history, along with an identifiable set of approaches, goals, methods, techniques, and algorithms. AI researchers have continued to explore various mathematical models for representing problems in computers, and algorithms for finding useful solutions in these models. While their approaches take many forms, they are all subsumed under the idea that intelligent action to solve a problem in a domain can be modeled mathematically, and that computers can effectively solve problems within those mathematical models. The various approaches are typically distinguished by their models. Even though the types of mathematics used to model problems is incredibly diverse, it mostly breaks down into two main types (and many hybrids) of logical and statistical approaches. Logical approaches attempt to encode the structure of a problem space in logical expressions and draw solutions from logical inference (deduction). Statistical approaches attempt to structure problems as parameterized datasets and draw their solutions from statistical inference (induction/prediction). And while AI has found success in solving

[10] Consider the algorithms that find the shortest path to your destination in a navigation application like Google Maps. This problem can be represented as a weighted graph theory problem and tackled by a variety of AI and other techniques, including A*, though we rarely think of it as "AI."

problems from checkers and chess, to speech and image pattern recognition, it has also gone through periods of hype and decline.

But it seems a bit silly to think that there might be a global arms race amongst the great powers to devise the best logical and statistical modeling techniques. Just substituting the word "math" for "AI" in the recent headlines above helps to highlight why there may be hype around "AI arms races," but not "math arms races." The public clearly associates something more than the underlying mathematical techniques to AI.

Of course, public opinion and popular media is often informed as much by science fiction as it is by science. Consequently, there are many ideas about what AI is that are drawn from science fiction literature, film and television. AI in science fiction is something more like an artificial conscious being, which might embody a computer,[11] robot,[12] building or spaceship,[13] live in digital networks,[14] or even transcend embodiment altogether.[15] While science fiction can be a culturally rich way of engaging with contemporary and possible future forms of society and modes of technology, most researchers and governments are not seriously concerned about this kind of AI, at least in the near future.

Closely related to the science fiction view of AI is the recent interest in Super Intelligence, or Artificial General Intelligence ("AGI"). The idea of AGI is that advances in AI research could result in a system that has intelligence that exceeds human capability not just in one specialized area (like chess) but in a broad array of domains, or in general. That is, right now you can buy a computer program that will beat you in chess, even if you are a grand champion player. But no one worries about that program doing anything other than playing

---

[11] COLOSSUS: THE FORBIN PROJECT (Universal Pictures 1970); WARGAMES (United Artists 1983).

[12] BLADE RUNNER (The Ladd Company, Shaw Brothers, & Blade Runner Partnership 1982); SHORT CIRCUIT (TriStar Pictures 1986); STAR TREK: THE NEXT GENERATION (Paramount Domestic Television 1987-1994); AUTOMATA (Nu Boyana & Green Moon Espana 2014); EX MACHINA (Film4 & DNA Films 2014).

[13] 2001: A SPACE ODYSSEY (Stanley Kubrick Productions 1968); DEMON SEED (Metro Goldwyn Mayer & Herb Jaffe Productions 1977); I, ROBOT (20th Century Fox 2004).

[14] WILLIAM GIBSON, NEUROMANCER, (Ace, 1984); THE TERMINATOR (Orion Pictures 1984); THE LAWNMOWER MAN (New Line Cinema 1992).

[15] GHOST IN THE SHELL (Kodansha 1995); TRANSCENDENCE (Warner Bros. Pictures 2014).

chess, much less that it might become conscious or seek world domination. An AGI or Super Intelligence, on the other hand, might devise its own goals, or develop its own understanding of the world in such a rapid and vast way that humans might no longer be able to predict or control it. While devoting some effort to mitigating such an existential risk might be prudent, the predictions of the likelihood of such a development vary widely, and its consequences are largely speculative. For our purposes, this does not seem to be the type of AI that concerns diplomats and the media—even if it provides fuel for the media's sensationalism around AI arms races.

If we follow the long history of AI hype and decline cycles up to the most recent peaks of hype, we find a technique called Deep Learning. Deep Learning applies machine learning to large datasets using decades-old techniques of neural networks and convolution. The real breakthrough that has occurred in recent years is not so much a novel technique or algorithm, but a combination of factors, which suddenly made existing techniques practically effective for a much greater number of significant problems. Namely, the availability of large training datasets, access to powerful and inexpensive computing resources to researchers, including cloud computing and the use of graphical processing units ("GPUs") as massive parallel computers within desktop personal computers ("PCs"), has resulted in the ability to make very large neural networks (hence "Deep") that can be trained with large datasets to solve practical problems. Another way to look at the current situation is that it has become much cheaper and easier to use these techniques.

With the rise of internet usage over recent decades, especially more recently through smart phones and social media, a few companies have become massive platforms for collecting data on their users about an increasing variety of human activities and commercial services. This has also meant a massive surge in "big data" that includes many real-world examples of all sorts of things, from spam email to photos to traffic patterns to dating to shopping to voter behavior, that can now be easily used for training neural networks. The potential opportunities to exploit this data to increase economic efficiencies are enormous, and many parties are interested in finding new ways to capitalize on it–economically, politically, or otherwise.[16]

[16] Shoshana Zuboff in her book THE AGE OF SURVEILLANCE CAPITALISM (Profile Books Limited, 2019) argues that this is a powerful new form of capitalism that has turned the personal data of the public commons into private resources for generating private wealth. This tends to support the view that "data is the new oil" insofar as natural resources like oil were not seen as valuable until technologies like the internal combustion engine and

AI and Deep Learning seem to offer the means of doing this, and to do so in an automated and scalable way, as opposed to requiring the labor of skilled statisticians and analysts to study datasets looking for useful patterns. Deep Learning is especially attractive insofar as it does not seem to require human engineers to really figure out the detailed structure of a given problem space–so long as engineers have a deep enough neural network and big enough dataset, they will get some good results. As such, much of the current hype around AI is really about applying Deep Learning to anything and everything to discover hidden patterns, optimize efficiencies, and thereby maximize profits.

Other new technologies that are capturing the public imagination, such as voice-recognizing personal assistants, drones, self-driving cars and even private space flight, also draw heavily upon recent advances in computation, miniaturization, economies of scale in electronic components, and, of course, machine learning techniques. While these have been quite successful, they have not yet revolutionized social life in the way of the train, telephone, light bulb, airplane, television, or the Internet itself have. Yet, these AI-enabled technologies remain a cultural symbol for the promised power of future AI applications. Given the current state of the information economy, Deep Learning would seem to promise deep profits, and nation-states are right to pay attention to the profits to be made across a broad swath of industries and sectors, and the potential impacts on their economic competitiveness in global markets.

With these various definitions of arms races and AI in mind, we now turn to several of the best candidates for what an "AI arms race" might be. Again, the view of the author is that each of these views, rightly or wrongly, at least partially informs the public notion of what an AI arms race means. After laying them out, we will consider their relationships to one another and what might be productively done about them.

A. *The AI Arms Race is an Economic Competition*

The idea that there is a race to develop the most capable AI and to translate this into economic dominance by capturing markets, users, data, and customers is probably the most salient interpretation of the

---

organic chemistry devised ways to turn those resources into wealth, setting off international competition in the age of industrial capitalism to acquire and control those resources.

phrase. Indeed, one can look to the existing economic competition between Silicon Valley companies like Google, Apple, and Facebook, along with Amazon and Microsoft to develop AI technologies and acquire AI talent as just such an "arms race." Of course, this is a figurative or metaphorical sense of "arms race" insofar as the competitors are not really building weapons or engaged in armed conflict.[17] The competition is for markets and profits, and the competitors are mainly companies based in the same country. These companies perceive the ability to develop and deploy a specific technology or set of technologies as being critical for business success, and thus, there is a technological race underway.

One could also look at this as a competition between the technology companies of different countries. To the extent that language barriers and the Great Firewall have created separate Internet ecosystems in China and the U.S., these companies mostly compete with each other for those customers within their own countries. However, U.S. and Chinese companies are directly competing to gain new Internet users throughout Asia, Africa and South America. As more of the world comes online each year, there is competition to bring new users from all over the globe into one sphere or the other of digital influence. And this is a competition not only for sales or users, but primarily a *competition for their data*. Insofar as the current hype around AI is really about the collection and utilization of big data, then the real global competition underway is to suck up the world's data. And mostly, that means user data. If data is the new oil, then this global competition for users and their data is also a strategic competition between states for control of the key resource in the future global economy.

It is hardly a stretch to view U.S. Internet companies and their Chinese competitors as being in a competition for global influence and

---

[17] Google recently announced it would not renew a contract with the Pentagon to use its AI to analyze drone footage, following pressure from employees and academics. *See* Kate Conger, *Google Plans Not to Renew Its Contract for Project Maven, a Controversial Pentagon Drone AI Imaging Program*, GIZMODO (June 1, 2018), https://gizmodo.com/google-plans-not-to-renew-its-contract-for-project-mave-1826488620 [https://perma.cc/TV63-4BV4]; Amazon, Microsoft and Google are currently competing for the multi-billion dollar JEDI contract to provide cloud computing services to the Pentagon, which is deeply tied to various AI services. *See* Naomi Nix, Ben Brody & Kathleen Miller, *Pentagon's Winner-Take-All Move on Cloud Contract Expected to Favor Amazon*, BLOOMBERG NEWS (July 26, 2018), https://www.bloomberg.com/news/articles/2018-07-26/pentagon-goes-with-winner-take-all-10-billion-cloud-contract [https://perma.cc/65MV-ZE6N].

market domination. Chinese companies like Baidu, Alibaba, JD.com and Tencent are very much engaged in the very same technological competition to acquire massive amounts of users and their data, and to utilize AI to leverage that data for economic advantage as are Amazon, Microsoft, and Google. In Russia, the public uses primarily Russian social media platforms like VK.com, OK.ru, and Rutube.ru, alongside U.S. platforms like Facebook, Twitter, and Instagram. If we are indeed in a new age of surveillance capitalism, then the great powers are ultimately competing for access to personal data by becoming essential platforms and capturing market share of users globally. We could argue whether the "great powers" here are the technology companies or the nation-states with which they are aligned, as one could for the Dutch East India Company and the Dutch military in the 17th century. And if AI is the technology that amplifies the value and impact of all the data being collected, then it is certainly an integral part of this great competition, and a metaphor for one's standing in that competition.

The military metaphor only goes so far here, however. Countries and companies are in perpetual economic competition, sometimes deploying military force to protect their economic interests. The war metaphor further breaks down when we consider the nature of AI research. Most of the advances in basic AI research still occur primarily within academia and are shared in conferences and academic journals. Of course, most of the major advances in AI applications occur within companies and some government research labs– mostly because they hold the data. Even the major corporate players have established entities like the Partnership on AI and OpenAI to promote the widespread distribution of AI techniques, algorithms, standard data sets, and best practices.[18] As such, it seems unlikely that any single country or company will retain a significant advantage due to a technological breakthrough for very long. From a business perspective, such innovations can result in capturing market shares, brand loyalty, and network effects that have longer term advantages. But these will not preclude others from competing with similar products and technologies. Significant advantages will come through network effects– those countries that foster more and bigger AI businesses, educate and train more AI researchers, invest more in

[18] *See* THE PARTNERSHIP ON AI, https://www.partnershiponai.org/ (last visited March 5, 2019) [https://perma.cc/52LM-PW76]; OPEN AI, https://openai.com/ (last visited March 5, 2019) [https://perma.cc/ZER8-7GZG].

basic science and public works applications will inevitably be the ones with the best AI, and the countries that reap the greatest rewards from it.

B. *The AI Arms Race is a Proxy for Technical Dominance*

Another way to view the AI arms race is as the space race of our generation. The Cold War between the U.S. and U.S.S.R. played out not only in real proxy wars, but also in symbolic proxy wars. One need not look too far to find symbolic proxy wars during the Cold War, from chess championships to Olympic hockey. The most spectacular of these was the Space Race in which the two countries sought to put satellites and people into space, to reach the Moon, and probe the planets and deep space. Of course, the technologies developed in the space race had many direct military applications, from intercontinental ballistic missiles, to supersonic aircraft, to spy satellites and telecommunications satellites, and a host of advanced sensor technologies. But these probably could have been developed without the cost and spectacle of the Space Race. That spectacle was about capturing the public imagination and demonstrating technical superiority over one's competitors. As such, it was largely a cultural battle fought through technological innovation.

Insofar as the AI arms race is a cultural battle to convince the world which country has the greatest technical prowess, and which country holds the keys to the technological (and economic) future, then this is an apt analogy. Of course, like the Space Race, the culturally symbolic aspect of the AI arms race does not preclude the application of AI to more traditional forms of economic and military competition. Indeed, such an AI arms race would likely entail many of the same dual-use capabilities and applications that motivated government investments in the Space Race.

The U.S. has also worried about technological competitiveness with other nations along various dimensions since World War II. In the 1980s, there was great concern that Japanese industrial productivity and quality would so far outstrip the U.S. as to result in massive economic consequences. Coupled to this were fears of losing out in specific technologies, from automobiles to consumer electronics, and from microchips to 5th Generation computing. Similar fears continue today regarding U.S. competitiveness in 5th Generation

cellular networking,[19] and to a lesser extent in solar and green energy technologies. Such rhetoric seems an inevitable part of mobilizing private capital and companies, and spurring interest and investment in strategic areas.

However, the global strategic costs of "losing" such competitions are mostly economic and political, not military. Ultimately, what is at stake in such a symbolic cultural race is prestige, and perceived power and influence, which are somewhat intangible. The general perception of a nation as having technological prowess, respect, and influence has real political impact, but is difficult to measure or observe. But the economic AI competition has very little to do with the military, arms, or armed conflict, except as it becomes directly applied to military applications, or provides economic support for traditional military build-ups.

C. *The AI Arms Race is About Cyberwarfare and Cybersecurity*

Another view is that since AI is essentially software, "AI weapons" will be cyberweapons. Accordingly, the main strategic advantage to be sought for in AI developments will be in the cyber domain—the capture and destruction of data and control of the information infrastructure. As the tools, techniques, and software used in cyberattacks becomes increasingly intelligent by utilizing AI, it should become increasingly capable of overcoming defenses and having greater effects. Similarly, the best cybersecurity defenses against these cyberattacks will also depend more and more on AI. Taken together states will race each other for cyber-superiority by developing AI for both offensive and defensive capabilities. What are we to make of such a view?

On the one hand, it seems obvious that cyberattacks and cyberdefenses will become increasingly sophisticated. And it does not seem improbable that various AI techniques might add to their sophistication and accelerate their improvement. If we take AI to mean "better IT/software" then this is simply a claim about the increasing strategic importance of tools for cyber operations. If we take AI to mean a specific set of computational techniques, then it remains to be seen how valuable those techniques are in cyber

[19] Stu Woo & Drew FitzGerald, *How Cellphone Chips Became a National-Security Concern*, WALL STREET JOURNAL (Mar. 7, 2018), https://www.wsj.com/articles/how-cellphone-chips-became-a-national-security-concern-1520450817 [https://perma.cc/JRK4-GVBR].

operations. Are deep networks better at finding zero-day exploits in systems than humans, and are they more practical and cost effective to implement?

For instance, a system could automatically scan for vulnerabilities in networks. Once access is gained to systems, software could automatically scan the system, and seek out higher levels of access, gather intelligence, and/or sabotage data as it goes. Of course, this is already what is done, but perhaps AI could improve on existing scripts and software. What we know from the recent history of cyberattacks is that the human elements are usually the weakest point in system security. It is often a human who is lured into clicking on malware in phishing attacks, or is convinced by a charming voice over the phone to reset passwords, that compromises system security. To the extent that AI could be developed to manipulate humans more effectively, or detect when humans may be being manipulated, then it could also be effective in this area. This is perhaps one of the scarier applications of AI– to optimize human psychological manipulation. It is already being used for extensively advertising, and may soon be applied to increasingly sophisticated spear-phishing types of operations.

But it is hard to see how the incremental gains in cyber from applying AI technology could result in a dramatic strategic shift. Ultimately, AI is about automating human labor– in this case programmer and hacker labor. Perhaps in all-out cyber warfare, having an army of hacker-bots could be a force multiplier for a small but skilled set of human hackers. They might serve a similar function in low-intensity cyber harassment operations, but it is more difficult to see any strategic advantage to be gained by massive investments here, unless the expectation is to escalate into high-intensity cyber conflict, or that cyber warfare could supplant conventional warfare as primary in armed conflict. States likely already have the ability to wreak significant havoc, or even shut down, each other's information infrastructures if they wanted to, without massive investments in AI.

D. *The AI Arms Race is About Weaponizing AI for Social Manipulation*

Related to the idea of applying AI to cyberwarfare– attacking information networks, infrastructure and data– is to apply AI to information warfare and propaganda– essentially conducting psychological operations by shaping the information environment of mass media, social media, and the internet. Just as AI could be applied to the human engineering side of cyber operations, it could

also be used to shape public understanding and political action more generally.

This is essentially the implication of recent revelations about for-profit companies like Cambridge Analytica and their actions to influence U.S. voters through targeted advertising. Their "added value" to more traditional public relations and marketing strategies, is the claim that they developed psychological models derived from Facebook user profile data– essentially a basic AI technique. These models are claimed to reveal how best to identify persuadable voters, their innermost fears and desires, and thus which messages are most likely to influence them and how to pitch those messages for maximum effect. These claims should be treated cautiously, like most PR and advertising, in the absence of much empirical verification of the efficacy of the end result. Still, quite a lot can be done with very basic demographic information (age, gender, race, income level, home address) in terms of predicting political views. And much can be gained from the real-time measurement of user responses on webpages and responsive adjustments to those, such as mass-scale real-time A-B testing to optimize advertisements and messaging. These techniques will become even more powerful when integrated with eye-tracking technology that helps to identify and isolate the focus of attention of users on images and messages. Moreover, the filter bubbles of like-minded and socially-tied users that exist on social media platforms like Facebook are ideal for influence and manipulation by ideologically divisive tactics.

The takeaway here is that the real or potential efficacy of automated mass deception relies primarily on mass data collection under weak or non-existent privacy controls, coupled with rudimentary multi-variable correlation. The actual analysis of that data beyond basic correlations might yield subtler or more powerful forms of manipulative messaging, but that is an empirical question. Again, what we have learned about the production of "fake news" is that it was largely conducted by human labor given the perverse incentives of social media platforms. It was less about content than clicks, and sensationalism and extremism are good at getting clicks. Automating that labor, of course, will inevitably make the production of such misinformation more efficient and effective.

Russia has a known history of interfering in elections, and this new set of tools and platforms, coupled with cyberattacks, appears to

have had a significant impact on the U.S. election of 2016.[20] From a global and national security perspective we should be just as concerned that such messages might come from foreign governments and their agents, illegal campaign operations, profit seeking opportunists, malevolent individuals, extremist groups, or misguided teenagers. Such undermining of the public sphere threatens democratic states regardless of the agents and aims, though the various agents and aims may require different types of policy responses.

E. *The AI Arms Race is About Weaponizing AI for Conventional Warfare*

One of the more literal interpretations of the "AI arms race" idea is that this involves building "AI weapons" or weaponizing AI for conventional warfare. This can also be tied into two other major initiatives– the United Nations discussions at the Convention on Certain Conventional Weapons (CCW) on the possible regulation of Lethal Autonomous Weapons Systems (LAWS or just AWS), and the U.S. Pentagon's efforts to advance the "Third Offset Strategy."[21] While these issues may overlap in terms of hype and buzzwords, the underlying issues are only partially related to AI and its advance.

In terms of the UN discussions of AWS, the issue there is really about the control of weapons systems, and more narrowly, the control of the targeting and engagement of weapons in an attack. The aim of the Campaign to Stop Killer Robots is to ensure that there is always meaningful human control over the use of weapons, and that humans do not delegate the authority to kill to machines, or otherwise abrogate their legal accountability and moral responsibility.[22] The Campaign is often mischaracterized as trying to ban a broad category

---

[20] Jane Mayer, *How Russia Helped Swing the Election for Trump*, THE NEW YORKER (October 1, 2018) https://www.newyorker.com/magazine/2018/10/01/how-russia-helped-to-swing-the-election-for-trump [https://perma.cc/LBG4-VCNQ].

[21] *See* THE UNITED NATIONS OFFICE AT GENEVA, *2018 Group of Governmental Experts on Lethal Autonomous Weapons Systems* (LAWS), UNOG (2018), https://www.unog.ch/__80256ee600585943.nsf/(httpPages)/7c335e71dfcb29d1c1258243003e8724?OpenDocument&ExpandSection=6%2C3#_Section6] [https://perma.cc/6CM4-TAZ7].

[22] *See* CAMPAIGN TO STOP KILLER ROBOTS, https://www.stopkillerrobots.org/ (last visited March 5, 2019) [https://perma.cc/22FX-AASW].

of technology, or non-existent weapons. But really, it is focused on clarifying the law and establishing norms going forward on the limits of machine autonomy and authority in armed conflict.

AI, as a specific set of technologies, could be used to enable or enhance the autonomous functions of a weapon system, but this is not necessary. Very dumb and simple machines can operate autonomously, and there are many ways to automate targeting and firing that need not involve any AI techniques. Of course, AI might offer some elegant solutions and new capabilities for autonomy, and certainly the history of autonomous robotics is closely intertwined with AI. But the point is not really about the technologies used, but rather preserving human agency, authority, and accountability in the use of lethal force in armed conflict. How best to do this, exactly, is a matter of ongoing debate, and will hopefully emerge with some level of international consensus around a norm that could form the basis of a legally-binding instrument.

AI does challenge the efforts at consensus building insofar as some proponents, such as Ron Arkin, have argued that AI controlled systems may perform much better than humans in some respects or domains.[23] In this view, AI will be so good that we need not worry about autonomy in weapon systems. I have challenged this view on several fronts, including the fact that legal reasoning and judgment is not likely to be easily imitated by AI, and even if it were, it would still not constitute the right kind of moral or legal agent to actually take responsibility for its actions.[24]

Moreover, there are a host of risks to automating targeting that are not resolved even if we imagine highly reliable systems. Among those are the risks of hacking and arms races, which are relevant in the current discussion. By automating weapons systems, one exposes them to the risk of unauthorized agents taking over, including AI-enhanced cyber attacks. In this sense, meaningful human control

---

[23] RONALD C. ARKIN, PATRICK ULAM & BRITTANY DUNCAN, AN ETHICAL GOVERNOR FOR CONSTRAINING LETHAL ACTION IN AN AUTONOMOUS SYSTEM: TECHNICAL REPORT GIT-GVU-09-02, GA. TECH. UNIV. (2009), https://www.cc.gatech.edu/ai/robot-lab/online-publications/GIT-GVU-09-02.pdf [https://perma.cc/N7RA-5V2S].

[24] *See* Peter Asaro, *On Banning Autonomous Lethal Systems: Human Rights, Automation and the Dehumanizing of Lethal Decision-making*, INTERNATIONAL REVIEW OF THE RED CROSS, 94 (886), Summer 2012, 687-709, http://www.icrc.org/eng/resources/international-review/review-886-new-technologies-warfare/index.jsp [https://perma.cc/A2ZL-AS7H].

could serve as a "human firewall" in the event of such cyberattacks, preventing weapons release.

In terms of arms races, the retrofitting of manned weapons platforms into autonomous platforms, or the development of new autonomous weapons platforms could easily foment and feed new arms races. So too could the introduction of destabilizing new capabilities such as robotic swarms. And insofar as autonomous weapons could become a new form of weapon of mass destruction (WMD), due to the ability of an individual or small group to unleash mass destruction and death, they would be very politically destabilizing.

F. *The AI Arms Race is the Third Offset Strategy*

While Deputy Defense Secretary, Bob Work promoted the development of advanced digital technologies as part of what he calls the Third Offset Strategy.[25] The notion of the offset is that, while the U.S. has a history of trailing competing militaries, primarily the Chinese and Soviet militaries in terms of numbers of soldiers and equipment, it has historically offset that perceived deficiency through technological superiority—*i.e.* quality over quantity. The first offset was considered to be nuclear weapons and deterrence following the Second World War, while the second offset following the Vietnam War sought general technical superiority in terms of intelligence, surveillance, and reconnaissance, stealth, and precision guided munitions– each offsetting more traditional forms of numerical advantage.

The focus of the third offset is on remote and autonomous platforms, big data and information processing, and information dominance. Again, most of these amount to simply applying software and IT solutions to military planning and operations, along with cyber warfare, information warfare and autonomous platforms as discussed above. AI, as a set of software techniques will have a place in this, but it is difficult to say it would be more important than the networks or

---

[25] Cheryl Pellerin, *Deputy Secretary: Third Offset Strategy Bolsters America's Military Deterrence*, U.S. DEPARTMENT OF DEFENSE (October 31, 2016), https://www.defense.gov/News/Article/Article/991434/deputy-secretary-third-offset-strategy-bolsters-americas-military-deterrence/ [https://perma.cc/9S2H-LQ5T].

databases the Third Offset requires, which raises the question as to why this is an "AI arms race" and not an "IT arms race."

Within defense policy discussions, there is considerable hype around both the Third Offset Strategy and the sense that the world is at a critical moment in the strategic use of data and control of information. Of course, this was realized during the Second World War by cyberneticians and the architects of the information age. The fact that so much of the world is computerized and datafied is a direct result of the competition for data and information processing since then. Ultimately, however, the Third Offset will fail as an offset if the sought-after capabilities are matched by competitors, which is likely as it would set the conditions for an AI arms race. Thus, if the U.S. pursues this third offset, and other states do the same, they will all pour increasing economic resources into technologies with diminishing military returns in a classic case of the Red Queen hypothesis.

G. *The AI Arms Race is About Building a Super Intelligence/AGI*

The concepts of Super Intelligence and Artificial General Intelligence are a bit fuzzy and speculative. The basic idea is that it might be possible to create a computational system that vastly outperforms humans in many, or any and all cognitive areas. Whereas existing AI systems can do this in limited domains, an AGI could learn new domains so rapidly that it could quickly outperform humans in many areas. The idea is fuzzy because we do not really know what it means or what technologies would be required to realize it. Without a clear understanding of how such a system might work or its contours, one must speculate as to what capabilities it might have and how it might behave.

There could conceivably be a "race" between states to create such a technology. In this view, such a technology would be capable of giving a strategic advantage in a broad array of domains– from stock trading to logistics to military operations, to scientific and technological innovation itself. These might be completely unforeseeable until the AGI technology exists, and starts creating exotic new technologies. But it also raises a host of questions about whether it would be benevolent or malevolent towards its creators, what side it might take in a conflict or towards conflict in general, whether it might be completely indifferent to human concerns, or whether any of the technologies it might develop could or would be controlled by humans. An AGI, to use another cinematic reference, is a bit like the Ark of the Covenant

in the Indiana Jones film.[26] It may or may not exist, and if it does, and if it is indeed powerful, it is not clear that whoever discovers it will manage to control it or will merely destroy themselves, and possibly everyone else, with it.

IV. CONCLUSION

While there is a great deal of concern and hype around the idea of an "AI arms race," a careful analysis of the ideas and notions caught up in this hype reveal a complex tapestry of concerns over technology and its role in economic, political, and military competition. It does not really make sense, from a policy perspective, to lump all of these concerns together because policies that promote success in one will inevitably conflict with and undermine others. Consider where economic and intellectual resources should be applied by states in this "AI arms race". If the real arms race is about economic dominance, then the best AI programmers should be going into industry, and creating consumer and business AI applications. This is pretty much the case today, and as such supports the economic competition interpretation of the AI race over others. But if the concern is with AI cyber operations, or the Third Offset, then talent and resources should be directed there, at the expense of consumer and business applications. In the mobilization of a nation for war, the state directing resources in this way might make sense, despite its negative economic impacts in the short and long terms. While it makes sense for states to recruit small numbers of cyber warriors, few industrialized nations would consider conscripting computer programmers from successful companies into military duties. Indeed, employees at companies like Google are protesting their company's involvement in military projects altogether.[27] But clearly an AI arms race framed in economic terms is very different from one framed in military terms. And of the ultimate goal of nations is economic and political control, it is probably counterproductive to overinvest in military applications. Further, as IT platforms come to control more data, capture vast economic wealth and political influence, and deploy AI to predict and exert more control over human behavior, the

---

[26] RAIDERS OF THE LOST ARK (Paramount Pictures 1981).

[27] Ben Tarnoff, *Tech Workers Versus the Pentagon*, JACOBIN (June 6, 2018), https://jacobinmag.com/2018/06/google-project-maven-military-tech-workers [https://perma.cc/GZB2-RH2C].

greatest threat to the political hegemony of nation-states could become the technology companies themselves, rather than other nation-states.

That said, there are certain policy approaches that would benefit both framings. Namely the promotion of computer science education and academic AI research. These areas already receive more funding than many other areas of education and research, but given the shortages of tech workers and academic researchers with these skills, and their increasing demand, this seems like the most obvious and effective investment states could make in this technological race.

The deeper concerns of citizens, as well as business and political leaders, as to the increasing surveillance and collection of personal data, and its use in manipulating people for economic and political purposes will not be addressed or resolved by any simple policy solution. It is not enough to eliminate "fake news," or enhance cyber security or privacy policies. The social changes being wrought by new technologies appear likely to accelerate rather than slow down, and a broad variety of policies will be needed to reign it in. Moreover, some of its forms, including "fake news" and information manipulation, point to a future where public discourse and public opinion are increasingly divided and irreconcilable in ways that make it difficult to discuss which policies will bring about a shared vision of society.