2022

AI & Arms Control: The Cold War Continued

Robert Sobon

AI & Arms Control: The Cold War Continued

By: Robert Sobon

Abstract

As artificial intelligence ("AI") begins to play an ever-increasing role in military affairs, the traditional role of humans in warfare is likely to be diminished, if not eliminated, with time. The future of military intervention could depend entirely on the decisions and capabilities of AI-powered systems. This poses real-world, present-day questions, long before "intelligent" combat becomes a reality. One of those questions, includes the viability of arms control in artificially intelligent warfare. This paper addresses theories of arms control, introduces various historical and present-day examples of arms control, the concept of artificial intelligence, and examines whether conventional approaches could apply to artificially intelligent instruments of war. In doing so, the conclusion reached is that conventional obstacles plaguing arms control treaties are likely to be exacerbated by the inherent complexities involved in AI development and deployment. To effectuate success, nations must cooperate now to establish universally agreed-upon terms such as keeping the human warfighter behind the war machine and reducing automation when a system is responsible for taking a life. Further delay in establishing such agreements will continue to thrust society into what one can only describe as the AI Cold War – a conflict which may otherwise carry on indefinitely.

Introduction

In 1945, the world bore witness to the collapse of the Nazi regime in Germany and the end of World War II.[1] The conflict catapulted the world into a race towards military supremacy and technological might. World War II paved the way for the development of microwaves, radar, more

[1] A&E Television Networks, *World War II*, HISTORY, (Oct. 29, 2009), https://www.history.com/topics/world-war-ii/world-war-ii-history.

advanced computers, enhanced aeronautic capabilities, and nuclear weapons.[2] Intense competition between the United States and the Soviet Union to develop military technology defined the next several decades, culminating into what became known as the Cold War.[3] This conflict was a geopolitical chess match which lacked the traditional bloodshed and destruction of wars past.[4] Instead, the United States and the Soviet Union flexed their ideological and industrial muscles to prove which was the superior nation.[5] The exact period in which the Cold War transpired is often contested, and some would argue it never truly ended.[6] Instead, it lay in a state of dormancy until today. Today, we find ourselves in the midst of what could be considered, the AI Cold War.

Mankind has continued to make strategic advancements in science and technology at breakneck speeds.[7] We have progressed as a people from putting a man on the moon, to putting the moon in our living room with the help of virtual reality.[8] As these remarkable innovations continue to develop, and countries continue financing various technological endeavors, their defense-related concentrations begin to shift as well. Developed world powers have begun the transition from warheads to algorithms.[9] Specifically, the growth of Artificial Intelligence, otherwise known as AI.[10]

---

[2] WWII National Museum, *The Scientific and Technological Advances of World War II*, The National WWII Museum – New Orleans (Apr. 9, 2021), https://www.nationalww2museum.org/war/articles/scientific-and-technological-advances-world-war-ii.
[3] A&E Television Networks, *Cold War History*, HISTORY, (Oct. 27, 2009), https://www.history.com/topics/cold-war/cold-war-history.
[4] *Id*.
[5] *Id*.
[6] John Feffer, *The Cold War Never Ended*, Institute for Policy Studies, (Sep. 10, 2014), https://ips-dc.org/cold-war-never-ended/.
[7] Max Roser and Hannah Ritchie, *Technological Progress*, Our World in Data, (Apr. 26, 2021), https://ourworldindata.org/technological-progress.
[8] Alex Landon, *Visit the Moon with This Stunning New Virtual Reality Experience*, Secret London, (Mar. 8, 2019), https://secretldn.com/moon-virtual-reality-experience/.
[9] Ayran Olckers, *AI in War: "Algorithms Will Fight Each Other In 20 Years,"* Medium, (Mar. 7, 2020), https://medium.com/swlh/ai-in-war-algorithms-will-fight-each-other-in-20-years-4df66b346826.
[10] *Id*.

AI is slated to become the most important device, if not the only one, through which warfare will be conducted in the future.[11] China, Russia, the United States, and other sophisticated nations all view AI as a determinative factor of their world power in the future.[12] As Russian President Putin so eloquently stated in 2017, whoever wins the AI race, will rule the world.[13]

Mr. Putin is correct. Militaries around the world are investing money in the development of autonomous weapon systems as fast as their treasuries can print it.[14] The United States Department of Defense has invested over $15 million in the Joint Artificial Intelligence Center ("JAIC") so far, which is responsible for the assessment and application of AI defense projects across all branches of the military.[15] As these ventures continue to take place and expand, major implications begin to stem from AI's impact on conventional wartime concepts like rules of engagement, laws of war, arms control treaties, and whether the use of AI-based military systems constitutes warfare at all.[16]

This paper will address the tip of the spear and examine whether orthodox arms control agreements will ever be capable of governing the rules of engagement and arms in an artificially intelligent future. Or, on the other hand, whether there is any practicality in implementing such measures to begin with.

I. Artificial Intelligence

[11] *Id*.
[12] *Id*.
[13] *Id*.
[14] Laura Wood, *Global Artificial Intelligence in Military Market (2020 to 2025)*, BusinessWire, (Mar. 23, 2021), https://www.businesswire.com/news/home/20210323005739/en/Global-Artificial-Intelligence-in-Military-Market-2020-to-2025---Incorporation-of-Quantum-Computing-in-AI-Presents-Opportunities---ResearchAndMarkets.com.
[15] *See* Olckers, *supra* note 9.
[16] The Economist, *Artificial Intelligence and War*, (Sep. 7, 2019), https://www.economist.com/leaders/2019/09/05/artificial-intelligence-and-war.

Artificial Intelligence is defined as the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.[17] Artificially intelligent systems are created with human-like capabilities such as the capacity to reason, to learn from experience, and to discover meaning.[18] Although they have not yet reached full human intellect or the capacity to be sentient, their near-human skills are stunning.[19]

A. A Brief History of Artificial Intelligence

The concept of artificial intelligence dates back to the days of the great Greek philosophers and their beliefs that human intelligence could be mechanized.[20] They imagined "an instrument that could accomplish its own work" and save humans from performing tedious labor.[21] In the 1700s, Jonathan Swift's novel *Gulliver's Travels* depicts what he calls the "engine" – which was essentially a modern-day computer.[22] These intelligent concepts continued to appear in popular literature and film throughout the following centuries. Movies like *Metropolis* and *The Wizard of Oz* depicted humanoid robots as far back as 1927.[23]

Fictional examples aside, in the 1940s, the Atanasoff Berry Computer was created and labeled the first electronic digital computer.[24] Although not quite artificially intelligent, the almost mystical appeal of these machines captured the popular imagination and influenced the work of individuals around the world.[25] One such individual, a man by the name of Alan Turing, proposed

[17] B.J. Copeland, *Artificial Intelligence*, Britannica (Apr. 8, 2021), https://www.britannica.com/technology/artificial-intelligence.
[18] *Id.*
[19] Carissa Veliz, *The Challenge of Determining Whether an A.I. is Sentient*, Slate, (Apr. 14, 2016), https://slate.com/technology/2016/04/the-challenge-of-determining-whether-an-a-i-is-sentient.html.
[20] Rebecca Reynoso, *A Complete History of Artificial Intelligence*, Learning Hub, (Mar. 1, 2019), https://learn.g2.com/history-of-artificial-intelligence.
[21] *Id.*
[22] *Id.*
[23] *Id.*
[24] *Id.*
[25] *Id.*

the idea that machines could be used to reason and solve problems just like humans can.[26] In 1950, Turing published his famous paper called *Computing Machinery and Intelligence*, in which he developed a framework for intelligent machines and proposed to his readers the famous question, "[c]an machines think?"[27]

Several years later, John McCarthy, a computer scientist, coined the term for what we now know as, artificial intelligence.[28] These events soon gave rise the first AI computer programs, machines, and catapulted research pondering the theory that the essence of human life – the ability to reason and think, could be harnessed by non-human entities like machines.[29]

The first human-like rendition of AI is arguably the project known as ELIZA.[30] The interactive computer program was capable of engaging in a somewhat interesting and human-like conversations between an end-user and its AI counterpart.[31] ELIZA could ask an individual to tell it what has been bothering them, and a user could say the weather is awful – to which she would respond, "I'm not quite sure I understand you fully."[32] Of course, after a few attempts the mock psychotherapist would respond properly with, "You say you hate the current weather?," and work with you to pass your plight.[33] Countless other real-world and entertainment examples came to fruition throughout the late 1900s building upon this foundation including C-3PO in Star Wars, WABOT-2 which could read and play music, and even the horrifying childhood toy we all grew to hate, Furby.[34]

---

[26] *Id.*
[27] *Id.*
[28] *Id.*
[29] *Id.*
[30] *Id.*
[31] *Id.*
[32] *Id.*
[33] *Id.*
[34] *Id.*

Modern examples include Honda's ASIMO robot, Roombas, NASA's rovers, Amazon Alexa, Google Deepmind's AlphaGo, Tesla's self-driving vehicles, and other well-known projects; all of which contain the innerworkings of a century's long journey with artificial intelligence.[35]

The future of AI lies not only in our convenient home gadgets and self-driving cars, but notably in the commencement and execution of warfare.[36] AI-powered autonomous weapons are capable or surveying land, identifying enemy combatants, and even engaging targets without traditional human intervention.[37] Lethal Autonomous Weapons Systems ("LAWS") like killer robots do not yet exist in the capacity we would imagine them to, however, researchers believe it is possible to develop such conscious technologies within the next few decades.[38]

B. How AI Works

Understanding the interplay between AI and arms control requires at least a rudimentary understanding of how AI operates. The first rendition of AI technology is often referred to as symbolic AI.[39] Symbolic AI was named so because it made use of symbolic reasoning which may also be described as binary "if-then" statements (*e.g.*, if X=Y and Y=Z then X=Z).[40] This approach to AI formation requires encoding the information one wants the machine to know into a set of "if-then" statements that the machine executes to solve problems.[41] The final product is known as an algorithm.[42] Humans are kept in the loop in this type of AI, because all of the machine's

[35] *Id.*
[36] *See* Olckers, *supra* note 9.
[37] Michael T. Klare, *Autonomous Weapons Systems and the Laws of War*, Arms Control Association, (Mar. 2019), https://www.armscontrol.org/act/2019-03/features/autonomous-weapons-systems-laws-war.
[38] *Id.*
[39] Philip Boucher, European Parliamentary Research Service, *Artificial intelligence: How does it work, why does it matter, and what can we do about it?*, EPRS, (2020), https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641547/EPRS_STU(2020)641547_EN.pdf.
[40] *Id.*
[41] *Id.*
[42] *Id.*

capabilities and limitations are based on developer input.[43] The machine is only capable of making the decisions and solving the problems it has been programmed to unravel, inhibiting true autonomy.[44]

The second leap in AI's growth came through what is called machine learning.[45] Machine learning automates the learning process of the algorithms described earlier.[46] Therefore, rather than requiring human intervention to make corrections or adjust the information embedded in a code, the algorithms are designed to identify patterns and apply what they have "learned" to create more "if-then" statements.[47] These statements continue to grow and make a system more "intelligent" through user-input or general data acquisition.[48] Although machine learning is as old as symbolic AI, it did not become prominent until recent times – sparking the development of modern AI applications.[49]

AI's third leap was the advent of what is called deep learning.[50] This process allows AI systems to identify objects and perform complex tasks, while continuing to learn from its actions, absent the human involvement mentioned earlier.[51] Deep learning is a subset of machine learning and works very much like a human brain.[52] Large data sets are used to teach the AI what it needs to know and although machine learning and deep learning are relatively similar, the latter does not need structured input to understand the data it is fed.[53] The AI is programmed to perform rapid calculations based on the ingestion of large quantities of data and progressively improves the

[43] *Id.*
[44] *Id.*
[45] *Id.*
[46] *Id.*
[47] *Id.*
[48] *Id.*
[49] *Id.*
[50] *Id.*
[51] *Id.*
[52] *Id.*
[53] *Id.*

accuracy of the calculations it generates over time.[54] This is the science behind contemporary gadgets like digital assistants, voice-enabled devices, and self-driving vehicles.[55]

To simplify the concept even more, consider digital assistants like Apple's Siri, Amazon's Alexa, or Google's Assistant. Years ago, one had to train their device to recognize their voice, perhaps by repeating their name or a certain phrase several times; that is machine learning.[56] Now one can open their new product, power it on, and immediately command it to play music; that is deep learning.[57]

As one can infer from how AI functions, its success is highly dependent on the availability and acquisition of data.[58] DOMO, a cloud software company, releases annual reports measuring the amount of data collected each minute for a given year.[59] In 2019, that report held that Americans used over 4,416,720 gigabytes of Internet data per minute.[60] To put that into perspective, users consume about 7GB streaming high-definition video on Netflix each hour.[61] As the amount of data we consume and produce grows, the capabilities of AI systems increase proportionally.[62]

C. Concerns with Autonomous Weapons

The mounting fears surrounding autonomous weapons and the like have sparked outcries from prominent figures in society.[63] In 2015, Stephen Hawking, Elon Musk, Steve Wozniak and

---

[54] *Id*.
[55] *Id*.
[56] IBM Cloud Education, *Deep Learning*, IBM Cloud Learn Hub, (May 1, 2020), https://www.ibm.com/cloud/learn/deep-learning.
[57] *Id*.
[58] *See* Boucher, *supra* note 39.
[59] DOMO, *Data Never Sleeps 7.0*, DOMO, (Apr. 10, 2021), https://www.domo.com/learn/data-never-sleeps-7.
[60] *Id*.
[61] Brady Gavin, *How Big Are Gigabytes, Terabytes, and Petabytes?*, How-To Geek, (May 25, 2018), https://www.howtogeek.com/353116/how-big-are-gigabytes-terabytes-and-petabytes/.
[62] *See* Boucher, *supra* note 39.
[63] Eric Levitz, *Elon Musk and Stephen Hawking call for a ban on autonomous weapons*, MSNBC, (Jul. 28, 2016), https://www.msnbc.com/msnbc/elon-musk-and-stephen-hawk-call-ban-autonomous-weapons-msna649206.

many other leaders in the AI and robotics space published an open letter condemning the development of autonomous weapons systems.[64] The authors collectively stated that,

> The key question for humanity today is whether to start a global AI arms race or to prevent it from starting. If any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow.[65]

Issues such as these are at the forefront of AI development and present a plethora of ethical, moral, and regulatory concerns.

In response, various coalitions and AI think-tanks have been established to bring some semblance of uniformity in approach when developing AI-based products, irrespective of their application.[66] For example, the International Technology Law Association ("ITLA") published a global policy framework entitled, *Responsible AI*, which calls for actions including grounding AI development in human-centric principles like accountability, that makes organizations developing, deploying, or using AI systems responsible for harm caused by AI.[67] These principles are not all encompassing, and this organization is not the only entity promoting AI-centric tenets.[68] Companies like Microsoft and other major competitors in the AI space have come together to promote the idea of responsible AI development.[69]

Notions of responsibility, legitimacy, and accountability have long been significant components of ethical warfare.[70] Humans have attempted to define and regulate agents of war

---

[64] *Id*.
[65] *Id*.
[66] Falon Fatemi, *How Companies Should Answer the Call for Responsible AI*, Forbes, (Feb. 28, 2020), https://www.forbes.com/sites/falonfatemi/2020/02/28/how-companies-should-answer-the-call-for-responsible-ai/?sh=78425e0913f5.
[67] International Technology Law Association, *Responsible AI A Global Policy Framework*, (1st ed, 2019).
[68] *See* Fatemi, *supra* note 66.
[69] Microsoft, *Responsible AI*, (Apr. 25, 2021), https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6.
[70] Peter Rowe, *Law of War*, Britannica, (Apr. 9, 2021), https://www.britannica.com/topic/law-of-war.

since the dawn of humanity.[71] The *Mahabharata* introduces concepts like the rule of proportionality stating, "[o]ne should not attack chariots with cavalry; chariot warriors should attack chariots. One should not assail someone in distress, neither to scare him nor to defeat him."[72] Proportionality implies a balancing act when conducting offensive military operations in which the methods of warfare cannot be disproportionate to the advantage sought.[73] Other examples are evident in religious texts like the Torah which states:

> When you besiege a city for a long time, making war against it in order to take it, you shall not destroy its trees by wielding an axe against them. You may eat from them, but you shall not cut them down. Are the trees in the field human, that they should be besieged by you? Only the trees that you know are not trees for food you may destroy and cut down, that you may build siegeworks against the city that makes war with you until it falls.[74]

Treaties, customs, and ordinary human values have influenced what are now known as the laws of war and rules of engagement.[75] These ideas seek to mitigate the suffering induced by war on an international scale and have produced countless declarations, conventions, treaties, and judgments ranging from the 1864 Geneva Convention to the 2008 Convention on Cluster Munitions.[76]

As the AI Cold War continues, the application of such principles continues to worry developers, academics, and civilians alike.[77] Arms researchers pose questions like, should a machine be making life and death decisions on the battlefield?[78] Or, whether in these warfare settings a military advantage is outweighed by the harm imposed on noncombatants.[79] The

---

[71] Shubhendu Shekar, *War Under International Law*, Legal Bites, (Jun. 1, 2018), https://www.legalbites.in/war-international-law/.
[72] *Id.*
[73] International Committee of the Red Cross, *How Does Law Protect In War?*, (Apr. 8, 2021), https://casebook.icrc.org/glossary/proportionality.
[74] *See* Shekar, *supra* note 71.
[75] *See* Rowe, *supra* note 70.
[76] *See* Klare, *supra* note 37.
[77] International Committee of the Red Cross, *Autonomous Weapon Systems: Is It Morally Acceptable For A Machine To Make Life And Death Decisions?*, (Apr. 13, 2015), https://www.icrc.org/en/document/lethal-autonomous-weapons-systems-LAWS.
[78] *Id.*
[79] *Id.*

Pentagon has even adopted its own principles calling for individuals to "exercise appropriate levels of judgments and care" when deploying and using AI systems, and that such automated weaponry should be "traceable" and "governable" – in other words created with some type of kill switch mechanism.[80]

These concerns and historical approaches to warfare have caused numerous AI-based weapons programs to be delayed or terminated.[81] In 2018 over 3,000 Google employees signed a petition in protest of Google's joint venture with the Department of Defense's known as Project Maven.[82] The project aimed to develop and integrate "computer-vision algorithms needed to help military and civilian analysts encumbered by the sheer volume of full-motion video data that DoD collects every day in support of counterinsurgency and counterterrorism operations."[83] In other words, the venture uses AI to rapidly process vast imagery utilized for target acquisition.[84] Unlike conventional warfare where Soldiers are responsible for identifying and subsequently engaging an enemy, AI weapons like Project Maven can replace that causal link.[85]

Despite mounting concerns, nations around the globe are investing heavily in the dream of forming AI-powered militaries.[86] In 2016, the United States Navy and Defense Advanced Research Projects Agency ("DARPA") launched what it named Sea Hunter.[87] Sea Hunter is an unmanned anti-submarine vehicle capable of operating completely autonomously for long periods

---

[80] Matt O'Brien, *Pentagon Adopts New Ethical Principles for Using AI In War*, AP News, (Feb. 24, 2020), https://apnews.com/article/73df704904522f5a66a92bc5c4df8846.
[81] Scott Shane and Daisuke Wakabayashi, *'The Business of War': Google Employees Protect Work for the Pentagon*, The New York Times, (Apr. 4, 2018), https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.
[82] *Id*.
[83] Adam Frisk, *What is Project Maven? The Pentagon AI project Google employees want out of*, Global News, (Apr. 5, 2018), https://globalnews.ca/news/4125382/google-pentagon-ai-project-maven/.
[84] *Id*.
[85] *Id*.
[86] *See* Wood, *supra* note 11.
[87] *See* Klare, *supra* note 37.

of time.[88] The U.S. Navy aims to introduce fleets of autonomous war vessels into the high seas and begin moving away from large numbers of manned ships.[89] The U.S. Army utilizes variants of what are known as Autonomous Remote Engagement Systems ("ARES").[90] These mounted machine guns are managed by AI which identifies targets and allows the end-user to fire the weapon remotely.[91] South Korea military hardware manufacturer DoDAMM introduced Super aEgis 2 in 2010, described as "an automated gun turret that can detect and lock onto human targets from kilometers away, day or night and in any weather conditions, and deliver some heavy firepower."[92] These few examples and applications of AI in war machines are only the tip of the spear.

The fundamental complexities of war are difficult to grasp by Soldiers on the battlefield, let alone autonomous weapons.[93] These systems, even when well developed, can commit fatal errors; errors for which we have not yet conceptualized effective regulatory or recovery procedures.[94] Some military leaders have addressed the fact that, autonomous weapons lack an ability to feel empathy and be merciful in war – restraints which limit the potential for brutality in conventional human combat.[95] The intricacies inherent in AI development as well as the execution of war seem almost insurmountable as the two concepts begin to merge.

---

[88] *Id.*
[89] *Id.*
[90] Ari Shapiro, *Autonomous Wepaons Would Take Warfare To a new Domain, Without Humans*, NPR, (Apr. 23, 2018), *https*://www.npr.org/sections/alltechconsidered/2018/04/23/604438311/autonomous-weapons-would-take-warfare-to-a-new-domain-without-humans.
[91] *Id.*
[92] Loz Blain, *South Korea's Autonomous Robot Gun Turrets: Deadly from Kilometers Away*, (Dec. 7, 2010), https://newatlas.com/korea-dodamm-super-aegis-autonomos-robot-gun-turret/17198/.
[93] *See* Shapiro, *supra* note 90.
[94] *Id.*
[95] *Id.*

II. Arms Control Defined

Arms control is defined as a "limitation of the use, exchange, or manufacture of military weapons by nations often as a policy established through diplomatic negotiation."[96] Arms control implies collaboration between states, generally in areas of military policies, to reduce the likelihood and destructiveness of war.[97]

A. A Brief History of Arms Control

In the shadow of hostility, humans have long attempted to make the world more secure and peaceful to curb the consequences of war.[98] Ancient civilizations like the Greeks made pacts during times of conflict in which sources of water could not be restricted or destroyed.[99] In Europe during the Middle Ages, agreements were made to protect the clergy as well as individuals identified as non-combatants.[100] In 1675, France and Germany prohibited the use of bullets laced with poisonous agents.[101] In the early 1900s, international peace conferences such as the League of Nations and the Geneva Conference were held to ban the use of noxious gas and to encourage the reduction of arms.[102] These small but mighty agreements led to more prominent ones such as the Geneva Protocol and the Geneva Convention; and paved the way for international bodies like the United Nations.[103] Each of these treaties, conventions, and international bodies shared a common goal – to limit the destructiveness of war that has plagued humanity since the beginning of time.

---

[96] *Arms Control*, Merriam-Webster Online Dictionary, (Apr. 8, 2021), https://www.merriam-webster.com/dictionary/arms%20control.
[97] *Id*.
[98] Richard Burns, *Encyclopedia of Arms Control and Disarmament*, (1993), *see also*, Bonn International Center for Conversion, A Short History of Disarmament and Arms Control, Warpp, (Apr. 8, 2021), http://www.warpp.info/en/m7/articles/m7-01.
[99] *Id.*
[100] *Id.*
[101] *Id*.
[102] *Id*.
[103] *Id*.

Some of the most well-known arms control agreements include the Antiballistic Missile Treaty ("ABM"), which limits the use of anti-ballistic missiles created to counter nuclear weapons in an effort to limit production of more offensive arms.[104] The Chemical Weapons Convention ("CWC") prohibits the development, production, stockpiling and use of chemical weapons.[105] The Nuclear Nonproliferation Treaty ("NPT") was established to stop nuclear propagation and contains agreements from major nuclear powers including the United States, Russia, the United Kingdom, France, and China.[106] The Strategic and Tactical Arms Reduction Treaty ("START") was devised to foster working agreements between the United States and former Soviet Union to slow the nuclear arms race.[107] These are just a few examples of the countless arms control deals aimed at ceasing, slowing, or eliminating armaments entirely.

B. Arms Control Purpose, Enactment, and Enforcement

Arms control is typically accomplished through a collective settlement between adversarial or competing nations.[108] The governments of participating states agree to certain terms regarding arms, conduct, or any other matter, and confirm the agreement by signing a treaty, holding a periodic convention, or any other similar act.[109] To ensure compliance, independent international bodies sometimes conduct inspections in person, verifications via satellite or airplane flyovers, or participating states will inspect one another.[110] These inspections are largely agreed upon within the terms of the arms control treaty as well.[111]

---

[104] Jack Mendelsohn and David Grahame, *Arms Control Chronology*, Center for Defense Information, (2002), https://carnegieendowment.org/pdf/npp/acc.pdf.
[105] *Id*.
[106] *Id*.
[107] *Id*.
[108] Kenneth Thompson, *Arms Control*, Britannica, (Jul. 20, 1998), https://www.britannica.com/topic/arms-control.
[109] *Id*.
[110] *Id*.
[111] *Id*.

Arms control is also considered by many thinkers to be directly responsible for tackling what is known as the security dilemma.[112] The security dilemma describes a situation in which states increase their security posture; typically, by expanding the size of their militaries or their number of warheads, which prompts other nations to do the same.[113] Arms control introduces conditions which states agree upon to prevent this type of risky competition.[114] Additionally, arms control provides a means for cost reduction as well as limits on potential damage resulting from war.[115] For example, states may agree to prohibit the use of certain types of munitions to lessen the cost of recovery following a conflict.[116]

C. Issues with Arms Control

Despite their promises, arms control measures have been generally ineffective throughout history.[117] Early renderings of arms control treaties were afflicted by imprecise language.[118] By identifying loopholes in a treaty's text, signatories have circumvented their terms by cleverly altering designs or modifying the fabrication of the weapons or items of focus.[119] For example, the Washington Naval Treaty of 1922 was established by the allied powers of World War I and sought to prevent an arms race by limiting the production of certain naval assets by imposing restrictions on weights, sizes, and variants of vessels.[120] Countries like the United States found loopholes in the treaty by designing better ships whilst adhering to the weight restrictions.[121] Other countries

[112] Anders Wivel, *Security Dilemma*, Britannica, (Apr. 8, 2021), https://www.britannica.com/topic/security-dilemma.
[113] *Id*.
[114] *Id*.
[115] *Id*.
[116] *Id*.
[117] Andrew J. Coe and Jane Vaynman, *Why Arms Control Is So Rare*, Cambridge University Press, (Dec. 18, 2019), https://www.cambridge.org/core/journals/american-political-science-review/article/abs/why-arms-control-is-so-rare/BAC79354627F72CDDDB102FE82889B8A.
[118] *Washington Naval Treaty*, Wikipedia, (Apr. 8, 2021), https://en.wikipedia.org/wiki/Washington_Naval_Treaty.
[119] *Id*.
[120] *Id*.
[121] *Id*.

like Italy misrepresented their weights entirely as the treaty lacked any means of effective review.[122] Due to the lack of compliance, the treaty was scrapped shortly before the end of 1936.[123]

The ability to withdraw or to remain a non-signatory from a treaty's inception weighs heavily against arms control success.[124] For example, the Treaty on the Non-Proliferation of Nuclear Weapons ("NPT") has 4 non-signatories including Israel, South Sudan, Pakistan, and India, none of whom have signed the treaty.[125] The objective of the NPT was to promote nuclear weapon disarmament while permitting peaceful applications of nuclear power like, generating energy.[126] Each country cites various explanations for their lack of cooperation. For example, India argues that the NPT creates restrictions on possession of nuclear weapons upon newly developing nations, because states which tested nuclear arms prior to 1967 are permitted to retain them.[127] Other countries like Israel cite foreign policy decisions in which the country refuses to divulge information regarding nuclear arms.[128] Setting aside a given country's motivations, lacking important participants makes valuable treaties like the NPT virtually worthless.[129]

Foreign policy decisions like deliberate ambiguity, frustrate arms control agreements as well.[130] Deliberate ambiguity refers to a country's attempt to protect certain information to avert risk.[131] It gives rise to its own set issues, as these actions may often be misinterpreted by other

---

[122] *Id*.
[123] *Id*.
[124] *Id*.
[125] United Nations Office for Disarmament Affairs, *Treaty on the Non-Proliferation of Nuclear Weapons (NPT)*, United Nations, (Apr. 8, 2021), https://www.un.org/disarmament/wmd/nuclear/npt/.
[126] *Id*.
[127] Aritry Das, *Modi Wants Nuclear Path for India to Go Green*, Asia Times, (Sep. 30, 2019), https://asiatimes.com/2019/09/modi-wants-nuclear-path-for-india-to-go-green/.
[128] Marvin Miller and Lawrence Scheinman, *Israel, India, and Pakistan: Engaging the Non-NPT States in the Nonproliferation Regime*, Arms Control Association, (Apr. 10, 2021), https://www.armscontrol.org/act/2003-12/features/israel-india-pakistan-engaging-non-npt-states-nonproliferation-regime.
[129] *Id*.
[130] World Heritage Encyclopedia, *Policy of Deliberate Ambiguity*, (Apr. 8, 2021), http://self.gutenberg.org/articles/eng/Policy_of_deliberate_ambiguity.
[131] *Id*.

nations.[132] For example, in the early 2000s, Saddam Hussein took part in a form of deliberate ambiguity regarding Iraq's possession of weapons of mass destruction ("WMD").[133] Saddam used this ambiguity to his advantage by becoming a fearmonger for Iraq's neighboring states and to control the country's citizens.[134] Under the auspice of Saddam Hussein owning WMDs, the United States invaded Iraq in 2003.[135] Other instances of strategic ambiguity have caused lesser, albeit noteworthy consequences even between allied nations. For example, the United States conceals whether certain naval surface ships contain nuclear arms.[136] This policy of deliberate ambiguity has led to New Zealand banning United States Navy ships from its ports.[137]

The lack of implications resulting from a signatory's failure to abide by or join an arms control agreement is likewise disadvantageous.[138] Russia has repeatedly violated the Intermediate-range Nuclear Forces Treaty ("INF") by developing, testing, and deploying intermediate-range missiles which the treaty prohibited.[139] The result of this noncompliance led to the United States withdrawal from the treaty in 2019 and it is no longer in force.[140] In 2013, Syria was found to be using chemical agents in attacks in Ghouta, and consequently decided to acknowledge their existence and partake in the Chemical Weapons Convention ("CWC).[141] Since then, Syria has experienced two known chemical attacks. The Khan Shaykhun chemical attack in 2017 killed at

---

[132] *Id*.
[133] *Id*.
[134] *Id*.
[135] *Id*.
[136] *Id*.
[137] *Id*.
[138] Chelsea Bailey and Erik Ortiz, *Syria Airstrike: Trump Declares 'Mission Accomplished' After Chemical Weapons Targets Hit*, NBC News, (Apr. 14, 2018), https://www.nbcnews.com/politics/donald-trump/syria-airstrikes-trump-declares-mission-accomplished-after-hitting-weapons-targets-n866001.
[139] Bill Gertz, *Large-Scale Russian Treaty Violations Revealed on Eve of New Talks*, The Washington Times, (Jun. 21, 2020), https://www.washingtontimes.com/news/2020/jun/21/russian-arms-control-treaty-violations-revealed-ah/.
[140] *Id*.
[141] Julian Borger and Patrick Wintour, *Russia Calls on Syria To Hand Over Chemical Weapons*, The Guardian, (Sep. 9, 2013), https://www.theguardian.com/world/2013/sep/09/russia-syria-hand-over-chemical-weapons.

least 89 individuals and injured over 541.[142] The Douma chemical attack in 2018 claimed the lives of nearly 50 people and injured over 100.[143] The penalty for these violations included repeated air strikes against alleged chemical weapons facilities and nothing more.[144]

Despite the multitude of existing and past arms control treaties, their efficacy continues to be limited or nonexistent.[145] Countries involved in these agreements have the difficult task of trading transparency and security.[146] The methods of verification are often intrusive, which some states fear will inform their adversaries about their capabilities or current developments.[147] Nations may simply elect not to participate, withdraw, or violate agreements facing little reaction from other signatories.[148] The difficulties inherent in enacting and enforcing traditional methods of arms control, will surely inhibit the practicability of such measures when applied to complex AI-powered instruments of war.

III. AI and Arms Control: Combined Obstacles

Considering the various shortcomings of arms control treaties past and present, as well as the complexities associated with AI development, the following aims to combine the matters into distinct subjects that stem from the hybridization of these issues.

A. Responsibility

Responsibility is arguably the largest hurdle in AI development and deployment.[149] Conventional weapons require user input, meaning the combatant is responsible for pulling the

[142] BBC News, *Syria Chemical 'Attack': What We Know*, BBC News, (Apr. 26, 2017), https://www.bbc.com/news/world-middle-east-39500947.
[143] BBC News, *Syria war: At Least 70 Killed in Suspected Chemical Attack In Douma*, BBC News (Apr. 8, 2018), https://www.bbc.com/news/world-middle-east-43686157.
[144] *See* Bailey and Ortiz, *supra* note 138.
[145] *See* Coe and Vaynman, *supra* note 117.
[146] *Id.*
[147] *Id.*
[148] *See* Gertz, *supra* note 139.
[149] *See* International Technology Law Association , *supra* note 67.

trigger or otherwise engaging a device.[150] When it comes to AI-based weapons however, many of these technologies may be completely automated or require such little human input, that the firing of a weapon, for example, may involve little consideration. Take for instance, the SGR-1, a 5.5mm machine gun and 40mm grenade launcher developed by Samsung.[151] This sentry robot is equipped with heat and motion detectors that can operate at distances of more than 2 miles.[152] As of 2014, the devices were deployed at the demilitarized zone in South Korea and are fully capable of identifying and engaging targets with little to no human intervention.[153]

The introduction of war machines like the SGR-1 raises both practical and philosophical questions. If the system misidentifies a target, engages said target, or advises its user to engage a target to which the user submits, who is to be held responsible? Is this machine error or user error? Is the developer who did not create a failsafe responsible for an AI system's resulting actions? Does one nation hold another responsible for implementing autonomous systems without complex testing? There are certainly more questions than answers for responsibility-bearing in arms control, specifically in cases where we are not dealing with traditional arms – pushing against the efficacy of an arms control agreement today and certainly in the future with AI-powered weapons.

B. Transparency

Transparency is already a point of contention for contemporary arms control treaties.[154] Should arms control enter the space of AI, transparency is likely to be impossible to establish. In this context, transparency means, "a duty for businesses and governments to inform people that they are interacting with an AI system and to provide information about its specifications,

[150] *Id*.
[151] Kim Tae-gyu, *Machine Gun-Armed Robots to Guard DMZ,* The Korea Times, (Jun. 24, 2010), http://www.koreatimes.co.kr/www/news/biz/2010/06/123_68227.html.
[152] *Id*.
[153] *Id*.
[154] *See* World Heritage Encyclopedia, *supra* note 130.

including the origin and nature of the data used to train the algorithm and to describe what it is that the AI system does, and how it does it."[155] This is a difficult goal for civilian product implementation, let alone arms control.[156] AI is often considered a proprietary technology, and defense companies and developers alike are prone to maintaining some level of secrecy regarding their technology.[157]

If arms control agreements are adopted by a group of nations, the likelihood that states will share the algorithmic makeup of their technology seems slim.[158] Policies of deliberate ambiguity are likely to exist in complex AI arms control agreements as they have in conventional agreements like the NPT. Countries will likely not want to share developments in AI-based defense technology because it gives them a significant advantage should they enter conflict.[159]

Even if international agreements are established to promote AI transparency, it may be impossible to achieve from a scientific perspective.[160] The way AI functions is unlike any other conventional weapon.[161] It operates by reaching conclusions via the data provided in an increasingly unsupervised manner.[162] AI does not explicitly share how or why it reaches certain conclusions.[163] This mysterious process is often referred to as the AI "black box."[164] The "black box" means that we understand what information was fed to an AI system as well as what that information leads the system to conclude, however, we do not know how the system turned said input into that output.[165] This may not appear troublesome with devices like digital assistants and

[155] *See* International Technology Law Association , *supra* note 67.
[156] *See* World Heritage Encyclopedia, *supra* note 130.
[157] *See* International Technology Law Association , *supra* note 67.
[158] *Id*.
[159] *See* World Heritage Encyclopedia, *supra* note 130.
[160] *See* Boucher, *supra* note 39.
[161] *Id*.
[162] *Id*.
[163] *Id*.
[164] *Id*.
[165] *Id*.

vehicles perhaps, but unexpected or fatal outputs produced by an autonomous weapon introduces a flood of new concerns.

Transparency is paramount to the success of any arms control agreement.[166] It is the characteristic with which past treaties have succeeded or failed.[167] In the case of treaties like the NPT or others regarding physical warheads, it may be easier to enforce and monitor the efficacy of these treaties and understand the innerworkings of their development. In the area of AI war machines, it appears far easier to hide code, provide false information, or even claim absence of the technology to begin with – and so transparency for AI-powered weapons seems unrealistic at first glance.

C. Enforcement

Unjust punishments or failure to enforce penalties will likely diminish any potential for AI-based agreements. The penalties imposed by arms control agreements have been historically deficient and larger states with more geopolitical power tend to walk away unscathed, while less powerful states are punished harshly for failing to comply.[168] As the events taking place in Syria illustrated, some states may face devastating repercussions in the event they disobey a treaty.[169] Others, like Russia, may have had the ability to exit or defy an agreement with little to no backlash.[170] If the provisions of a treaty are imbalanced and stronger states can consume a penalty while failing to comply, they will do so whilst continuing the offensive act.[171]

Additionally, a state can conceal the fact that it is developing AI arms by masking the code designed to automate weapons more easily than it can hide a physical armament which signatories

[166] *See* World Heritage Encyclopedia, *supra* note 130.
[167] *See* Coe and Vaynman, *supra* note 117.
[168] *See* Gertz, *supra* note 139.
[169] *See* Bailey and Ortiz, *supra* note 138.
[170] *See* Gertz, *supra* note 139.
[171] *Id*.

can monitor through satellite or directed inspections.[172] Even in situations where defiance is found, signatories' support for checks can wane, poor reporting procedures may exist, or a target state may heavily mislead an inspecting organization – prolonging any implementation of penalties.[173] Enforcement remains a difficult objective for modern treaties and will likely remain so with any AI-based agreements.

D. Inclusivity

An arms control agreement is only as effective as the parties involved. Classic arms control measures have been riddled with feeble membership and virtually no retention.[174] As the examples discussed illustrate, nations participating in arms control treaties tend to exit when the terms of an agreement no longer align with their foreign policies or objectives.[175] All signatories of an arms control agreement must have a real interest in a treaty's success otherwise they will likely see the associated disadvantages as overburdensome and remain uninvolved.[176]

Moreover, many countries simply never become signatories despite a seemingly international desire for them to do so.[177] If society is to create and employ ethical and legal forms of AI onto the battlefield, adversaries must agree that each will conduct themselves appropriately by developing their arms in accordance with the terms of an agreement. Currently, the world is hard-pressed to see any type of unity in approach. Most world powers including the United States, Russia, and Australia are vehemently against regulating or banning lethal autonomous weapons,

[172] U.S. Department of State, *New START Treaty Inspection Activities*, (Apr. 25, 2021), https://www.state.gov/new-start-treaty-inspection-activities/.
[173] Arms Control Association, *Iraq: A Chronology of UN Inspections*, (Apr. 25, 2021), https://www.armscontrol.org/act/2002-10/features/iraq-chronology-un-inspections.
[174] *See* United Nations Office for Disarmament Affairs, *supra* note 125.
[175] *See* Gertz, *supra* note 139.
[176] *See* Das, *supra* note 127.
[177] *See* United Nations Office for Disarmament Affairs, *supra* note 125.

whereas China supports legally binding bans or treaties against LAWS.[178] A lack of international cooperation or signatories would render any such ban or arms control measure fruitless.

E. Language

Arms control documents or convention policies must be drafted with precise and living language. Documents cannot contain language that is vague or ambiguous and must be void of any potential loopholes that opposing nations could take advantage of.[179] If a global conference agrees to require human intervention in autonomous weapons when faced with the question of taking a life, the drafted agreement must ensure that there are no gaps in the language, while providing means for amendments as technologies progress and invariably alter the landscape of treaties. Examples like the Washington Naval Treaty have demonstrated the consequences of hastily written treaties.[180] The language of an agreement, above all else, is the backbone of its success and must be drafted in a manner which fosters compliance and encourages involvement.

IV. AI and Arms Control: Potential Pathways Toward Success

A. Exposure

Despite the many drawbacks and unanswered questions regarding the efficacy of arms control agreements for autonomous weapons, there are ways in which employment of treaties may bring about success to varying degrees. For one, defiance may in fact be equally as helpful in protecting states from rogue AI applications. Nations which fail to obey or enter agreements can be identified and closely monitored by cosignatories as evidenced by commonly uncooperative states like Russia.[181] Even in situations where arms control is not directly impacted, a notorious

[178] Future of Life Institute, *World AWS Policy Map*, Future of Life, (Apr. 11, 2021), https://futureoflife.org/lethal-autonomous-weapons-systems/.
[179] *See* Washington Naval Treaty, *supra* note 118.
[180] *Id*.
[181] *See* Gertz, *supra* note 139.

nation's actions tend to remain under a vigilant public eye.[182] Take for instance, China, and their alleged development of biologically engineered super soldiers.[183] This development is now international news and countries are keen to follow any further developments.[184] The same could apply to arms control involving AI. Should a nation fail to comply, join, or otherwise be involved in questionable advancements, they will surely be monitored by adversaries and allies alike, in which case signatories could also amend treaties to better suit their needs in the face of rogue state actors.

Additionally, time may be of the essence not only in enacting arms control measures but pressuring them into existence as well. As the world grows more aware of artificial intelligence and its potentially devastating uses in combat environments, growing concerns could pressure lawmakers and world leaders to establish some type of controlling agreements. Organizations have already begun promoting the idea that all AI-based weapons systems should require strict human control.[185] The United Nations Secretary General believes that "machines with the power and discretion to take lives without human involvement are politically unacceptable, morally repugnant and should be prohibited by international law."[186] International bodies such as the United Nations and entities including the "Campaign to Stop Killer Robots," the "Future of Life Institute," and others have been taking strides to keep the human warfighter behind the fight.[187] Maintaining meaningful human control is likely an aspect many nations would prefer over banning autonomous

[182] Natalie O'Neill, *China Doing Human Testing to Create 'Super Soldiers': US Official*, New York Post, (Dec. 4, 2020), https://nypost.com/2020/12/04/china-doing-human-testing-to-create-super-soldiers-u-s-official/.
[183] *Id*.
[184] *Id*.
[185] Lethal AWS, *Lethal Autonomous Weapons*, Autonomous Weapons, (Apr. 11, 2021), https://autonomousweapons.org/.
[186] United Nations Secretary-General, *Secretary-General's Message to Meeting of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, (Mar. 25, 2019), https://www.un.org/sg/en/content/sg/statement/2019-03-25/secretary-generals-message-meeting-of-the-group-of-governmental-experts-emerging-technologies-the-area-of-lethal-autonomous-weapons-systems.
[187] *See* Lethal AWS, *supra* note 185.

weapons entirely in order to remain competitive, and the continued growth of the AI field along with public awareness will likely lead to the existence of additional organizations.

B. Internationally Promoted Uses

Despite the negative light surrounding AI and arms control, there are numerous potential and pacific benefits to these systems. The promotion of these benign applications could be achieved through international arms control agreements in a non-conventional sense where development is encouraged in select fields while restricting harmful uses. Advocates of AI war machines look to AI as a replacement for human fighters.[188] AI can be integrated into health care systems to provide remote care or assistance under hazardous conditions.[189] It can even go as far as providing streamlined transportation solutions lowering costs and reducing if not eliminating human operations entirely.[190] Proponents argue that targeting systems may even enhance the accuracy of target recognition and limit the hazards of misidentifying targets in complex combat situations.[191] States may agree to promote these specific uses in warfare as a means of counterbalancing the negative results of combat – the very objective arms control is set out to achieve.

C. Sheer Necessity

The final and perhaps most likely method in which a pathway towards success exists for AI-based arms control is by way of necessity alone. AI's complexities including its "blackbox" design, the inability to concretely identify bad actors, and society's historical desire to employ ethics even in times of armed conflict may further the trend for autonomous weapons awareness

[188] Tejaswi Singh and Amit Gulhane, *8 Key Military Applications for Artificial Intelligence in 2018*, Market Research Blog, (Oct. 3, 2018), https://blog.marketresearch.com/8-key-military-applications-for-artificial-intelligence-in-2018.
[189] *Id.*
[190] *Id.*
[191] *Id.*

and the need for international regulation. The potential hazards and uncertainties associated with autonomous weapons could demand collaboration as did the Treaty on the Prohibition of Nuclear Weapons ("TPNW"), which was entered into force January 22, 2021, aimed at banning and eliminating nuclear weapons. [192] The risks of a nuclear holocaust and horrors of wars past may generate the same reaction regarding the theoretically devastating and inhumane consequences of an autonomous war – barring international efforts to not act preemptively.

V. Conclusion

As AI continues to play an ever-increasing role in our lives, it will undoubtedly do the same for military capabilities. Nations around the world continue to harness the power of AI in the hopes of remaining a contender within our technologically dependent civilization. The uncertainties and complexities inherent in AI creation and application will frustrate the already struggling notion of arms control and continue an arms race of global proportion far into the future.

To prevent this, nations must cooperate to establish universally agreed-upon terms such as keeping the human warfighter behind the war machine and promoting benign uses of artificial intelligence in combat environments. The same level of cooperation is paramount for the success of any arms control measure, and global leaders must come to terms with the likelihood that adversarial nations will employ strategies of deliberate ambiguity or total secrecy. AI's exponential growth is unlikely to cease, and inaction now will continue the development of an AI Cold War. As Alan Turing stated in *Computing Machinery and Intelligence*, "[w]e can only see a short distance ahead, but we can see plenty there that needs to be done."[193]

[192] United Nations, *Treaty on the Prohibition of Nuclear Weapons*, (Apr. 28, 2021), https://www.un.org/disarmament/wmd/nuclear/tpnw/.
[193] Turing Trust, *Who Was Alan Turing?*, (Apr. 11, 2021), https://turingtrust.co.uk/alan-turing/.