# HOW COULD ML/AI TRANSFORM THE DIMENSION OF WEAPONS OF MASS DESTRUCTION IN THE SECURITY ENVIRONMENT?

*CHIARA ARTICO*

*ABSTRACT*

*The introduction of artificial intelligence (AI)-based systems in warfare has raised several concerns over the last decades, such as the many issues related to the potential deployment of increasingly autonomous weapons systems. The debate becomes even more intense when analysing what the potential uses of AI could be on weapons of mass destruction, although it is mainly considered only as nuclear arms in the literature. This article will first introduce some of the potential assets and evaluate the challenges of AI applications in warfare. Then, it will proceed to an assessment of its uses in relation to chemical, biological, radiological and nuclear (CBRN) weapons. In particular, analysed applications will include the use of artificial intelligence for nuclear early-warning and command-and-control systems, and the contribution of machine learning algorithms to the detection and classification of dual-use goods in non-proliferation. Furthermore, it will assess the potential for the employment of artificial intelligence in augmenting CBRN delivery systems capabilities and the potential for complementing CBRN weapons. It is important to note that AI is a very broad field that includes several different developments of this technology, such as machine learning (ML) and deep neural networks (DNNs), which will all be put under the same umbrella of AI in this article, when not specified otherwise.*

*KEYWORDS: Artificial intelligence; Machine learning; Warfare; Autonomous weapons; CBRN weapons*

## INTRODUCTION

Optimistic views of AI-based military systems in warfare often rely on the argument that the increasing use of this kind of technology can improve the preciseness of military operations and circumvent human error to avoid collateral damage. Furthermore, artificial intelligence systems have already demonstrated their efficacy in reducing costs related to logistics as well as enhancing communication and transparency during military operations (Sisson, 2019). In particular, the successful implementation of artificial intelligence elements in warfare is thought to positively influence combat in two ways: the changes in military organisations and combat philosophy that it could imply, and the changes in the speed of operations (Spindel, 2019).

## CHANGES IN MILITARY ORGANISATION AND COMBAT PHILOSOPHY

On the one hand, it is argued that when applied to intelligence, surveillance and reconnaissance (ISR) operations, AI systems are able to go through countless more images and pieces of information than humans in a much faster way, systemically isolating the most meaningful ones and directing them to the human agent that will evaluate them. This process saves a significant amount of time because it makes it possible to avoid the analysis of thousands of identical and irrelevant images. This changes the distribution of resources that are needed in warfare by splitting human and machine elements and entrusting them with different, though consequential, tasks. Therefore, this allows human actors to focus directly on relevant issues in war and war-related operations on the strategic level, rather than committing so much time to tactics (Spindel, 2019).

Moreover, artificial intelligence is believed to allow for a more reliable prediction of conflict outcomes in two ways. First, "with the support of advanced algorithms and supercomputing capabilities, the calculation and prediction results of artificial intelligence systems are more accurate than human brains" (Chen, 2018). Second, by using "wargame systems to more effectively test and optimize combat plans" (ibid) to further enhance the effectiveness and preciseness of those predictions. Consequently, combat philosophy could drastically change when considering the chances for a state to enter war since it would only do it when its AI-predicted possibilities of winning are sufficiently high to offset potential losses. All these elements contribute to the modification of military organizations and combat philosophy by reducing and sharpening steps that would otherwise be much more time-consuming and inaccurate.

## CHANGES IN THE PACE OF WARFARE

On the other hand, ML/AI applications in warfare are expected to drastically change the pace of war. In fact, it is evident that automated systems can make decisions on how to react to external stimuli more quickly than human soldiers. However, this speed could have ambivalent consequences on the window for human decision-making. Depending on the offensive context and the technology employed, in fact, a situation in which AI-based systems are employed and quickly react on the battlefield can take two forms: the quicker the reaction of the automated system, the shorter the time-span available to human decision-makers before their adversaries' counterattack, paving the way for potential miscalculations and mistakes.As opposed to this, however, the same situation could also create a framework in which human decision-makers have more time thanks to the speed and accuracy of the machines (Spindel, 2019). It is argued that "whether speed is a net positive or a net detriment to wartime decision-making will depend on how AI systems are integrated into military systems and what range of actions are pre-delegated" (*ibid*, p. 7).

## RELEVANCE TO NUCLEAR APPLICATIONS

These two sets of elements, and in particular the capacity of reliably predicting conflict outcomes and the need for augmented speed in action and reaction, are relevant to nuclear applications of AI when assessing whether an actor would opt for increased automation in its systems or would not employ such sophisticated technology. It is important to note that because the stakes are so high, this field is already one of the most automated ones. Automation contributes to maintaining nuclear actors' deterrence capabilities and their role as a potential threat credible and reliable (Conn, 2018). However, it is argued by Scharre and Horowitz that countries that are confident in their second-strike capabilities tend to not prioritise (nearly-) full automation of their nuclear systems. In line with this, Horowitz declares that United States (US) military leaders seem to have explicitly denied their interest in acquiring autonomous systems armed with nuclear warheads (ibid).

As opposed to this, it appears also true that "the more a country fears that its nuclear arsenal could be placed at risk by a first strike, the stronger its incentives to operate faster" (Conn, 2018) and to entrust automatic systems with the decision of launching an attack. Smaller nuclear actors might be prone to use enhanced automatic systems because of their fear of being disarmed by a first strike. Therefore, automation is perceived as enhancing their retaliatory capabilities and, thus, could be a good motivation to invest in it. Nevertheless, the evaluation of the role of AI in warfare needs to consider the numerous vulnerabilities of AI-based systems and how that will affect their overall performance on the battlefield.

Two key challenges are revealed in the early stages of the process of programming AI-based systems, namely in their training to operate in wartime frameworks and in the difficulties that arise from ethical issues (Scharre, 2019). For the purpose of this article, the ethical discussion will not be evaluated since it is believed to require an extensive analysis which falls beyond the scope of this work.

Nowadays, machine learning, which forms part of the broader field of AI, is ultimately defined as "the process by which a computer system, trained on a given set of examples (or data), develops the ability to perform a task flexibly and autonomously" (Stephens, 2019). However, ML/AI systems are currently still described as narrow, meaning that "they are only capable of performing the specific tasks for which they have been programmed or trained" (Scharre, 2019: 12). Consequently, they are not flexible when it is required to adapt autonomously to new tasks or environments in which they are inserted, unlike human brains. This also means that, while they seem ready to perform correctly and predictably, they are "susceptible to sudden and extreme failure when pushed outside the bounds of their intended use" (*ibid*). The war field is often chaotic and unpredictable, which also means that any artificial intelligence system cannot be programmed for all eventualities (Lin *et al.*, 2008), and that their performance will be only as good as the data with which they are trained (Scharre, 2019). In particular, as regards the nuclear application of ML/AI, this problem is fuelled by the fact that there is a significant lack of data with which to train artificial intelligence systems in order to teach them how to make appropriate nuclear actions-related decisions (Spindel, 2019). In the past 70 years, over 30 nuclear accidents and near-catastrophes have occurred, and several were related to some technological malfunctioning (Outrider Foundation, n.d.).

Moreover, warfare is also one of the contexts in which the higher number of adversarial examples can, intentionally or accidentally, mislead AI systems (Spindel, 2019).

Besides the possibility of being targeted with cyberattacks and suffering from bugs, automated systems can also be subject to other kinds of problems that might arise from human biases. In particular, projection bias and automation bias are likely to affect nuclear behaviours. The former is "a cognitive tendency where humans fail to accurately predict their own preferences in future situations" (Scharre, 2019: 14), which might lead to serious consequences in case decision-makers changed their minds only after the machine is programmed to react to a certain action in a specific way. The latter is the tendency to over-trust machines believing they would act better than humans, and thus entrusting them with a certain degree of authority that could be equally dangerous in case of malfunctioning, which is sometimes challenging to detect (Scharre, 2019). Both these biases are argued to be applicable to and especially thorny in the nuclear field, where the stakes are particularly high. Due to this, maintaining human actors in the loop does not seem a conclusive solution. For all these reasons, it seems evident that full or near-full automation of nuclear systems is neither achievable nor recommendable. Instead, what can be achieved is applying AI to command-and-control and early-warning systems. Because of a tendency of nuclear actors to maintain a certain degree of secrecy around this potential application of ML/AI, Borrie (2019) draws upon Cold War examples to outline how this can be done.

## EARLY-WARNING AND COMMAND-AND-CONTROL APPLICATIONS DURING THE COLD WAR

Both the US and the USSR were developing early-warning capabilities based on various types of sensors, which were usually checking the same inputs in different ways. However, both soon realised that such systems were fallible and avoided entrusting them with the authority to take vital decisions. This was mostly due to several near accidents on both sides, such as the famous Petrov case in 1983. Lieutenant Colonel Petrov concluded

that the five incoming missiles from the US detected by the Soviet early-warning system were a false alarm. He thought that the US would have never launched only five missiles perfectly knowing that the Soviet retaliation would have been massive. Another example is when a wrong file was mistakenly uploaded on a North American Aerospace Defense Command (NORAD) computer, which suggested an imminent attack from the Soviet Union that was soon proven wrong by another parallel sensors system. Stemming from these early attempts of applying AI to such systems, however, recent developments in this technology are allegedly increasing reliance on automation especially in ISR and automated target recognition (ATR), and allowed the US to adopt DNNs for detection, early-warning and target identification (Borrie, 2019; Geist and Lohn, 2018).

As regards command-and-control systems, following the US threat to deploy a missile defence system to defeat any incoming Soviet attack and the consequent fear of not being able to effectively retaliate, the USSR decided to develop and deploy the so-called 'Dead Hand' (Borrie, 2019).

This semi-automated command-and-control system, believed to still be in force in Putin's Russian Federation, was designed to guarantee nuclear retaliation in case of a verified attack by not necessarily requiring an order from the command-and-control system's top authority to launch a counterattack. In fact, in the case of the designated officials being offline, the system would automatically transfer launch authority to whoever was governing it at that moment, bypassing many steps of the chain of command. The idea behind this was to create a back-up communication system able to launch a retaliatory attack even if a first strike destroyed crucial elements of the chain . This option was supposed to convey stability to the Russians and guarantee a certain degree of security of retaliation. However, the US was allegedly unaware of the existence of the 'Dead Hand', thus invalidating the deterrent value that it was supposed to have (Borrie, 2019).

The early-warning and the command-and-control systems employed by the US and the USSR provide an idea of some ways in which AI could be effectively introduced to the nuclear environment and transform it. However, an important factor that can enhance its success is the perception of its potential by the adversary, rather than only what it can actually achieve. This means that AI-based systems could also drastically affect nuclear strategic stability, which is maintained when "adversaries lack a significant incentive to engage in provocative behaviour" (Geist and Lohn, 2018: 8). It is argued that the use of AI and the availability of AI-based ISR could fuel hostilities and lead to unintentional nuclear escalation by reducing an actor's confidence in its own second-strike capabilities, who could then consequently decide to strike first. Therefore, because of these risks that add to the ones reported above, it is deemed unlikely in the near future that AI-based technologies' implementation will subvert existing nuclear systems. Nevertheless, despite not entirely ruling out the possibility of AI supporting human decision-making (Geist and Lohn, 2018; Sankaran, 2019), it remains unclear whether this application, intended to help leaders providing a clearer picture of the context, will have a positive or negative effect on the process (Borrie, 2019).

## MODERN APPLICATIONS OF AI TO NUCLEAR TECHNOLOGY

### Tools for oversight

Another potential application of ML/AI systems to the CBRN weapons context is by exploiting them as a tool for oversight over strategic trade of dual-use goods in order to support non-proliferation efforts. Specifically, deep neural networks (DNNs) could be used to identify, recognise and classify dual-use tools with the purpose of supplementing existing mechanisms in non-proliferation. Withorne et al. (2020) conducted research to evaluate this possibility by elaborating several models and training them to recognise and classify these tools from a series of given datasets on dual-use items.

The main idea behind this project was to address the problem of non-highly specialised personnel often having to conduct trade controls, and missing potentially restricted goods. The high validation accuracy yielded by some of the models demonstrated the potential asset that the employment of such technology could add to the existing procedures. Nevertheless, these models need to be considered alongside human control, and it must be remembered that the effectiveness of these DNN models relies on the rigorousness of the datasets on which they are trained, as previously mentioned (Withorne et al., 2020).

## IMPLEMENTATION OF AI-BASED DELIVERY SYSTEMS

The effective implementation of AI in delivery systems could also alter the ways in which we know warfare. In fact, both the US and Russia are allegedly developing new torpedoes, unmanned underwater vehicles (UUVs) and other unmanned bombers, allowing them varying degrees of autonomy and arming them with CBRN weapons. The application of machine learning on these systems is expected to significantly improve their qualitative features, especially augmenting their preciseness and accuracy (Boulanin, 2019), and therefore even softening the public perception of CBRN weapons to be very inaccurate and indiscriminate (Kallenborn and Bleek, 2018). Several countries are thought to be undertaking research on machine learning to develop models that can guide hypersonic vehicles, which are not able to be driven manually because of their high speed. In the case of UUVs, autonomy is a necessary feature since underwater remote operability is impossible, such as in the case of the Russian 'Poseidon', also known as 'Status-6'. Nevertheless, at least from the US side, it seems unlikely that they will arm these highly technologically advanced devices with nuclear weapons because these are believed to necessitate specific safeguards that cannot, as of yet, be guaranteed by autonomous systems (Boulanin, 2019).

## DRONE SWARMS

Despite this, another report has focused specifically on the potential of AI-based drone swarms in relation to CBRN weapons. The most powerful nuclear actors are already pursuing this technology, but because of their intrinsic technological features, drone swarms are still a limited tool. In fact, it is technically extremely challenging to program them to behave autonomously, thus reducing the chances of non-state actors being able to produce them, and consequently limiting the risk of non-state actors to deliver CBRN weapons through drone swarms. Kallenborn and Bleek (2018) argue that drone swarms could play three different
roles when related to CBRN weapons, namely complementing, challenging and substituting them.

According to their report, the use of AI in drone swarms can complement existing delivery systems or even create new ones by increasing preciseness, endurance and survivability against traditional defences. Their "intelligence" would allow them to persistently collect relevant information and share it among its elements in order to increase chances of strike success and even be an excellent tool to defeat air, missile and anti-submarine countermeasures. Moreover, drone swarms could challenge CBRN weapons because they could help mitigate their potential threat. Drawing on the example of AI-enhanced ISR systems, drone swarms could identify and preventively destroy the enemy's CBRN stockpiles before they are armed upon bombers, or destroy bombers themselves after they are launched. They could also interfere with missile launches or prevent bombers from departing. Lastly, drone swarms could even substitute CBRN weapons as the new weapon of mass destruction. In fact, even if armed "only" with conventional weapons, they could represent an equally great threat when deployed instead of CBRN weapons because of their intrinsic characteristics. Many experts and observers have argued against drone swarms due to their potential of becoming extremely dangerous as they autonomy increases

(e.g. "Slaughterbots"; Kallenborn, 2020; Atherton, 2020). It is believed that actors could be more prone to deploying such technology because CBRN prohibitions would not apply (Kallenborn and Bleek, 2018).

## CONCLUSION

This article has assessed how artificial intelligence could be implemented in early-warning systems and how it could change how nuclear powers operate their command-and-control. Despite some general reluctance and challenges, it is believed that AI will play a role in decision-making within the next few decades as a "trusted advisor", especially in command-and-control systems. This will save time and stress on the decision-makers who would see themselves without the pressure of having to handle an overload of information and the fear of causing countless losses. Lastly, various applications of artificial intelligence on CBRN weapons' delivery and through drone swarms have been analysed, highlighting advantages and disadvantages. In conclusion, it is clear that AI applications are diverse, and while most of them are still hypothetical, their contribution to the dimension of weapons of mass destruction is both promising and controversial.

## BIBLIOGRAPHY


- Atherton, K. D. (2020). 'Are Armed Drone Swarms A Weapon Of Mass Destruction?' Forbes. [Online] [Accessed: 4 January 2021. Available at https://www.forbes.com/sites/kelseyatherton/2020/06/29/are-armed-drone-swarms-a-weapon-of-mass-destruction/?sh=62d3122b52fb]
- Borrie, J. (2019). 'Cold War Lessons for Automation in Nuclear Weapon Systems', in Boulanin, V. (editor), The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Vol. 1.
- Boulanin, V. (2019). 'The Future of Machine Learning and Autonomy in Nuclear Weapon Systems', in Boulanin, V. (editor), The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, Vol. 1.
- Chen, H. (2018). 'Artificial Intelligence: How Would AI Subvert Future War', China National Defense News, January 2, 2018. [Online][Accessed: 4 January 2021. Available at http://www.mod.gov.cn/jmsd/2018-01/02/content_4801253.htm]
- Conn, A. (2018). 'Podcast: AI and Nuclear Weapons—Trust, Accidents, and New Risks with Paul Scharre and Mike Horowitz', Future of Life Institute, September 28, 2018. [Online][ Accessed: 4 January 2021. Available athttps://futureoflife.org/2018/09/27/podcast-ai-and-nuclear-weapons-trust-accidents-and-new-risks-with-paul-scharre-and-mike-horo witz/?cn-reloaded=1&cn-reloaded=1]
- Geist, E. and Lohn, A. J. (2018). 'How Might Artificial Intelligence Affect the Risk of Nuclear War?', (Rand Corporation: Santa Monica). [Online][Accessed: 4 January 2021. Available at https://www.rand.org/pubs/perspectives/PE296.html]
- Kallenborn, Z. and Bleek, P. C. (2018) 'Swarming Destruction: Drone Swarms and Chemical, Biological, Radiological, and Nuclear Weapons', *The Nonproliferation Review,* 25(5-6), pp. 523-543
- Kallenborn, Z. (2020). Swarms of Mass Destruction: The Case for Declaring Armed and Fully Autonomous Drone Swarms as WMD, Modern War Institute. [Online][Accessed: 4 January 2021. Available at https://mwi.usma.edu/swarms-mass-destruction-case-declaring-armed-fully-autonomous-drone-swarms-wmd/]
- Lin, P., Bekey, G. and Abney, K. (2008). 'Autonomous Military Robotics: Risk, Ethics, and Design', US Department of Navy; Office of Naval Research Report, (California Polytechnic State University: San Luis Obispo).
- Outrider Foundation (n.d.) 'Accidents, Errors, and Explosions', Outrider Post. [Online][Accessed: 4 January 2021.Available at https:// outrider.org/nuclear-weapons/timelines/ accidents-errors-and-explosions/]
- Sankaran, J. (2019). 'A Different Use for Artificial Intelligence in Nuclear Weapons Command and Control', War on the Rocks, 25 April 2019. [Online] [Accessed: 4 January 2021. Available at https://warontherocks.com/2019/04/a-different-use-for-artificial-intelligence-in-nuclear-weapons-command-and-control/]
- Scharre, P. (2019). 'Military Applications of Artificial Intelligence: Potential Risks to International Peace and Security', The Militarization of Artificial Intelligence, (New York), August 2019, pp. 11-18. [Online][Available at https://www.un.org/disarmament/the-militarization-of-artificial-intelligence/]

- Sisson, M. (2019). 'Multistakeholder Perspectives on the Potential Benefits, Risks, and Governance Options for Military Applications of Artificial Intelligence', The Militarization of Artificial Intelligence, (New York), August 2019, pp. 3-5. [Online][Available at https://www.un.org/disarmament/the-militarization-of-artificial-intelligence/]
- Stephens, M. (2019). A machine-learning revolution. [Online][Accessed: 3 January 2021. Available at https://physicsworld.com/a/a-machine-learning-revolution/]
- Stop Autonomous Weapons (2017). Slaughterbots. [Online][Accessed: 4 January 2021. Available at https://www.youtube.com/watch?v=9CO6M2HsoIA&feature=emb_title]
- Spindel, J. (2019). 'Artificial Intelligence, Nuclear Weapons, and Strategic Stability', The Militarization of Artificial Intelligence, (New York), August 2019, pp. 6-10. [Online][Available at https://www.un.org/disarmament/the-militarization-of-artificial-intelligence/]
- Withorne, J. (2020). Machine Learning Applications in Nonproliferation: Assessing Algorithmic Tools for Strengthening Strategic Trade Controls. (Washington DC: James Martin Center for Nonproliferation Studies). [Online][Accessed: 4 January 2021. Available at https://nonproliferation.org/machine-learning-applications-in-nonproliferation-assessing-algorithmic-tools-for-strengthening-strategic-trade-controls/]