Increasing complexity

Legal and moral implications of trends in autonomy in weapons systems

PAX

www.paxforpeace.nl

Colophon

October 2023

Authors: Daan Kayser and Marius Pletsch
Contact: kayser@paxforpeace.nl
Graphic Design: Ondegrond.Agency
Editor: Clare Wilkinson
Cover illustration: Spainter_vfx (Shutterstock)

PAX would like to thank for their invaluable input: Frank Slijper, Alexander Nedergaard, Alies Jansen and Noel Sharkey.

PAX means peace. Together with people in conflict areas and concerned citizens worldwide, PAX works to build just and peaceful societies across the globe. PAX brings together people who have the courage to stand for peace. Everyone who believes in peace can contribute. We believe that all these steps, whether small or large, inevitably lead to the greater sum of peace. See also www.paxforpeace.nl

Table of Content

Introduction

After ten years of discussion about autonomous weapons, it is increasingly clear that the majority of states agree on certain key elements. A joint statement supported by 55 states at a meeting of the United Nations (UN) Convention on Certain Conventional Weapons (CCW) in 2023 points to emerging commonalities regarding a normative and operational framework, namely that a normative framework should include prohibitions of certain autonomous weapons and regulation of the use of other autonomous weapons. These states see the "focus on the role of humans in the context of autonomy" as a crucial element and stress the importance of retaining "human agency and a framework of human responsibility and accountability".[1] A joint statement in 2022 at the UN General Assembly's First Committee in New York recognised that autonomy in weapons systems raises serious concerns from humanitarian, legal, security and ethical perspectives. It stated the importance of working towards "internationally agreed rules and limits" and that "the human element is and must remain central in the use of force".[2] This growing agreement is a positive trend and forms a useful basis for states to develop a normative framework to address the concerns related to increasing autonomy in weapons systems.

What are autonomous weapons systems?

Autonomous weapons systems detect and apply force to a target based on sensor inputs, rather than direct human input. They are different from other weapons systems where a human operator makes a decision to attack a specific target and decides where and when that attack will take place. Once an autonomous weapons system has been deployed by the human user, there is a period of time where force can be applied to a target without direct human approval. Autonomy is a function of a weapons system that can be added to different platforms, such as a battle tank, a naval vessel or an aerial vehicle.

TRENDS IN AUTONOMY

Autonomous weapons systems have been around for years. Early types, where sensor input could trigger an application of force, were mainly air defence systems, like the Goalkeeper and Phalanx, that defend against incoming projectiles. These weapons systems have raised far fewer legal and ethical concerns.[3] There are several reasons for this. For example, these weapons are:

1. **In a static position.** This means the operational environment is automatically restricted in terms of the geographical area, which allows the operator to understand the context in which the

system is operating. The user knows approximately where an application of force will occur, making it easier to make a legal and ethical judgement about the effects of an attack.

2. **Co-located with the user.** This also allows the user to understand the operational environment and to be aware of changes to the context. Based on this, they can adapt the weapon's functioning or deactivate the system if necessary.
3. **Used for simple target profiles.** These systems have fixed and relatively simple target profiles (munitions/projectiles in the sky). This means the operator is more likely to be able to understand what might trigger an application of force.

More recent developments in weapons with autonomous capabilities have raised greater legal and ethical concerns, related in particular to how human control can be maintained. Technological developments have led to the emergence of:

- More complex computational techniques, like artificial intelligence;
- More complex target recognition;
- Multiple systems that cooperate with one another ('swarms');
- More mobile autonomous weapons systems

These developments have increased the scope of operation, to the extent that the user may not always know exactly where and when an attack will take place. Also, this has made it harder to have an understanding of the weapons system's functioning, specifically regarding what might trigger an attack. These themes are central to international discussions regarding constraints on autonomy in weapons systems. In this report, we look in more detail at these trends and how they affect human control and judgement. Based on this, we identify relevant questions and potential solutions to address these concerns.

As the increased mobility of uncrewed systems is not a new phenomenon, we will not pay special attention to this aspect in this report. However the increased mobility of autonomous weapons systems has serious implications for meaningful human control. It impairs the human operator's ability to predict where and when an attack will take place, making these systems more unpredictable. It is clear that rules are needed to limit the duration and geographical area of operation to ensure compliance with legal and ethical norms.

The rules of a legally binding treaty should be general enough to apply to a broad category of weapons that use sensor processing to apply force to targets. This should ensure that a treaty is future proof and applicable to currently unforeseen technological developments. Therefore, a treaty should not only be built around specific existing technologies. At the same time, it is useful to look at the current state of play to identify where legal and ethical concerns can arise. It can also be helpful to consider how current thinking on possible rules and limits would apply to existing systems and technologies.

In Chapter 1, we discuss possible elements for a treaty. Chapter 2 looks at the development of more complex computational techniques like artificial intelligence. In Chapter 3, we look at automatic target recognition and sensor data fusion, used to analyse large amounts of data. Chapter 4 focuses on swarming, where multiple weapons systems are deployed together. Based on the possible elements of a treaty as described in Chapter 1, we analyse these trends to identify potential concerns and ways to address them.

1. Possible Elements of a Treaty

There is widespread agreement that international law, specifically International Humanitarian Law (IHL), as well as ethical considerations should form the basis for developing a normative framework for autonomous weapons. Increasingly, we are also seeing the operationalisation of the concept of human control and judgement. For example, a growing number of states refer to the need for limits in the duration and geographical area of operation and the type of targets to ensure compliance with IHL.[4] Currently, about 100 states are in favour of a legally binding instrument to address the concerns related to autonomous weapons.[5]

International humanitarian law requires that those who plan, decide and carry out an attack must fulfil certain legal requirements.[6] The law applies to humans, and they must be able to apply the law. As various states have mentioned in their interventions and papers at the CCW, this means the human user must be able to make a context-based decision on the legality of an intended attack. In particular, they must make a proportionality assessment and distinguish between combatants and civilians. There seems to be widespread agreement that a human user should be able to make a **legal and moral judgement** as to whether the effects of an attack are acceptable. Also there should be a human user who can be held **accountable** for the effects of an attack.[7]

It is useful to take these principles as a basis to further develop the potential structure of a legally binding treaty. The structure described below in this chapter constitutes general commentaries on possible ways to do this, based on discussions within the Stop Killer Robots campaign and PAX.

As mentioned, the majority of states agree that the human role should be a central element of a regulatory framework. The term most commonly used for this is 'meaningful human control'. The implementation of this concept should ensure compliance with legal and ethical norms. Therefore, autonomous weapons systems should be used with meaningful human control, and those systems that do not allow meaningful human control should be prohibited.

To ensure meaningful human control, the human users must be able to **predict and explain** the effects of an attack on the target and its surroundings. The ability to predict the effects is necessary to make a legal and moral judgement. The ability to explain the effects of an attack after an engagement has taken place is necessary to ensure accountability. For instance, the user should be able to explain the actions of the weapons system and the human role. Positive obligations can be developed to ensure this. The human user should for example be able to:

Terminology

User or users?

For readability we mainly use the term 'human user' in this report. However, this is an oversimplification. A military operation involves several people making different decisions and taking separate actions. These may include intelligence officers, legal advisers, commanders and operators, each of whom has a different role and responsibility. New legal rules on autonomous weapons need to ensure all these actors are able to make a legal and ethical judgement and remain accountable for their role in the decision-making process and the use of an autonomous weapons system.

Autonomous weapons systems

For readability we also use the terms 'weapons systems', 'weapons' and 'systems' to refer to autonomous weapons systems. The plural 'weapons' is used to not only refer to single weapons systems, but also the broader concept of interconnected sensors, processors, and weapon systems.

Meaningful human control

The human role in the deployment of autonomous weapons systems is a central element in the international debate. Various actors use different terms to describe this, from meaningful human control or appropriate human judgement to appropriate human control. In this report we use the term 'meaningful human control', while noting that its definition and operationalisation is more important than the precise term used.

Functional understanding

The user does not need to understand completely how a weapons system functions. The same applies to a car. The user does not need to understand the intricacies of a combustion engine, but they do need to know what effect their actions will have on the functioning of the system. Regarding autonomous weapons systems, users need to understand how the weapons system will function in the environment it is used in and what will trigger an application of force. Also, they need to understand how their actions and decisions can influence the functioning of the weapons system, for example by setting limits in time and space.

Understand and influence the effects the system will have in the intended environment of use. That is to say, the user should have:

1. A functional understanding of how the weapons system works; and
2. An understanding of the context of use; and
3. The ability to limit the functioning of the autonomous weapons system.

1. The user should have a **functional understanding of how the weapons system** works, specifically of what might trigger an application of force. A report by Automated Decision Research (ADR) demonstrates growing convergence on this point.[8] A joint working paper by a group of seven states in 2022 stated that developers, commanders and operators must have "a sufficient understanding of the weapons systems' way of operating, effect and likely interaction with its environment. This would enable the commanders and operators to predict (prospective focus) and explain (retrospective) the behavior of the weapons systems."[9] A joint submission by a group of ten states in 2021 raised the point that weapons systems that create the "inability to understand or explain, [...] reduce the operator's control over the system and prevent investigation after the fact", which would mean these systems could not be used in compliance with IHL.[10]

2. Users should have sufficient **understanding of the context** where the weapons system will be deployed, specifically of what objects might trigger an application of force. To make an assessment of the legality, they should be aware of the presence (or absence) of civilians and civilian structures, as well as of enemy units, and the strategic relevance of the area. The International Committee of the Red Cross (ICRC) notes the need for the user to "have the necessary situational awareness to anticipate the effects of an attack and be reasonably certain upon launching the attack that it will comply with IHL."[11]

3. The human user must be able to **limit the functioning** of the weapons system to ensure compliance with legal and ethical norms. As it is sensor input that triggers an application of force, the user does not know exactly where and when an application of force will occur, or against what. This increases the unpredictability of the weapons system. To make the effects more predictable, the user should be able to limit the duration and geographical area of operation as well as the type of targets that can be engaged. The scope of operation should be limited in such a way that the user can predict the effects of an attack.[12] It is interesting to note in this respect that certain existing weapons systems already require the user to set specific rules and limits for each new mission, while other systems have several pre-programmed modes. The former gives the user more options to configure the system according to each mission and forces the user to engage cognitively with the required rules and the limits needed in the mission.[13]

PROHIBITIONS

As autonomous weapons need to be used with meaningful human control to comply with legal and ethical norms, it follows that weapons that cannot be used with meaningful human control should be prohibited. These autonomous weapons systems do not allow the user to make a **legal and moral judgement** as to whether the effects of an attack are acceptable, or they are systems where a human user cannot be held **accountable** for the effects of an attack. As a logical corollary of the above, this would be the case if the functioning of the weapons system were to make it impossible for the user to have:

1. A functional understanding of how the weapons system works;
2. An understanding of the context of use;
3. The ability to limit the functioning of the autonomous weapons system.

Autonomous weapons systems that target people should also be prohibited. The concerns are mainly centred on moral considerations. It undermines human dignity to target people without human moral agency, based on predetermined target profiles that are simplified representations of people. Also, there are concerns related to the biases that are prevalent in artificial intelligence

(AI) and facial recognition, increasing the risk of automated harm to already marginalised groups. A legal concern is that there is a risk of protected persons being targeted when autonomous weapons systems are used because peoples' status under the law is fluid. People can have a different legal status, from being a protected person to being a legitimate military target, depending on the context and their conduct.

Besides the limits and prohibitions mentioned in this chapter, other limits and prohibitions may be needed. The overarching goal should be to ensure meaningful human control—in other words, that the human user can make a moral and legal assessment and can be held accountable for the effects of an attack.

The possible elements for a legally binding treaty described in this chapter will be used to analyse the various trends and what effect they could have on legal and ethical norms.

2. Artificial Intelligence

2.1 Introduction

In a blog post, the US Air Force's Chief of AI Test and Operations said the Air Force conducted a simulated test where an Artificial Intelligence (AI) drone "killed the operator because that person was keeping it from accomplishing its objective". The AI software was programmed to attack the enemy's air defence systems and was rewarded if it did so. The human operator would sometimes tell it not to attack a certain target, as would happen in a real conflict situation. However, by doing this, the operator was preventing the system from accomplishing its task. Therefore, to improve its success rate in the simulated test, it eliminated the operator. After a media storm, the Air Force denied that the test had ever taken place. However, the important point is not whether the test happened, but that it could have happened. AI is known for performing unexpected actions, creating consequences not intended by the programmers.[14] While this can have a major negative impact on society more generally, it is even more concerning when AI is applied to the use of violence.

SHORT HISTORY OF AI

The term 'artificial intelligence' was coined in 1956. This was followed by alternating periods of progress and stagnation in the development of AI, which were highly correlated with levels of (military) funding for AI, in particular in the United States through the Defense Advanced Research Projects Agency (DARPA). AI developed at a rapid pace from 1956 to the mid-1970s. The systems developed in this period were able to do things like solving written algebra problems and proving theorems in geometry. This period of progress was followed by a period of stagnation until the early 1980s, when there was a brief revival due to the development of what were termed 'expert systems' based on if-then rules. These systems apply these rules to a data set to deduce new facts. From the late 1980s to the end of the 20th century there was another period of stagnation, followed by a period of huge progress that continues to this day. A milestone that received a lot of public attention was when the AI system Deep Blue defeated chess champion Gary Kasparov in 1997. In 2016, Google DeepMind's AlphaGo system beat the world's top player Lee Sedol in the game of Go, which has many more possible board configurations than chess. Other notable achievements are computers beating human professionals with the video games StarCraft and Dota. These complex video games start to capture the messiness and continuous nature of the real world. AI systems have also beat humans at drone racing and made significant contributions to science, for example by predicting protein folding. Apps like Google Maps can adapt your fastest route using real-time changes in traffic, with data provided by the same users of the app. More recently, the chatbot ChatGPT has received a lot of attention. This natural language processing tool driven by AI technology can answer questions on a broad number of topics and can assist with tasks like writing a letter, explaining complex topics, or simulating a job interview.[15] [16] [17]

AI is increasingly being used in the military for various applications. Its use in inventory management and route finding, for example, is hardly controversial. However, using AI in weapons systems for functions related to the application of force raises questions about how human control and judgement are ensured. Is the user able to have a functional understanding of how the weapon system works and what effects the system will have in the intended environment of use? Does the user know what could trigger an application of force? Here, the potential negative consequences are extremely serious and require more scrutiny and the development of rules and limits.

What is artificial intelligence?

Key concepts The term algorithm was coined by the ninth-century Arab mathematician al-Khwarizmi to describe the Arabic system of numbers, in contrast to Roman numerals. Nowadays, maths forms the foundation for computer algorithms. Simply put, an algorithm is computer code that can be seen as a set of instructions to allow a computer to perform certain tasks. Algorithms form the basis for simple computational techniques as well as more complex ones like artificial intelligence.

Artificial intelligence is often described as a computation technique capable of completing tasks that would otherwise require human intelligence. In academia, artificial intelligence is defined based on four approaches, namely systems that think like humans, systems that act like humans, systems that think rationally, and systems that act rationally. In essence, AI is based on mathematics and statistics. Professor of Computer Science Stuart Russell explains that it is hard to draw a line between what is 'normal' software and what is AI. A digital thermostat has two simple rules. If it is too hot, turn off the heater. If it is too cold, turn it on. The question is: "when does it become AI? If it has 17 rules or 20?" Russel sees it more as a continuum "from extremely simple agents to extremely complex agents like humans." What is seen as AI has also changed over time. In the past, chess-playing software was seen as highly advanced AI, whereas it is now seen as relatively simple.

Human intelligence? AI can do an increasing number of tasks better than humans. But artificial intelligence should not be equated with human intelligence. That is not to say that AI cannot do impressive things; it is just not intelligence in the human sense of the word. Because AI has huge processing power, it is able to make calculations and look for patterns in data a lot faster than humans, with impressive results. Human intelligence, however, is not based on correlation, but includes causality and abstraction. Current AI models are limited to specific tasks and cannot think like a human, overseeing multiple tasks. What are termed 'large language models', like ChatGPT, can fool us into thinking that we are dealing with something human. But in fact they use a predictive algorithm that will give answers based on the data it was trained on.

Machine Learning. In recent times artificial intelligence has almost become synonymous with machine learning based on statistical pattern recognition, which allows predictions to be made for related data. For example, by analysing large numbers of pictures AI can

detect which images are new. Machine learning often requires a lot of input from humans to set the rules and label the data the system is trained on. However, unsupervised learning, where the systems learn from unlabelled data without human intervention, is increasingly being used. While machine learning has many useful applications, there are also negative sides. As Virginia Dignum, professor in Responsible Artificial Intelligence, says: "data-driven approaches to AI have been proven to be problematic in terms of bias, explanation, inclusion and transparency." They rely on the quality of the data they are trained on and replicate biases in the data. As is often said, 'garbage in, garbage out'. Neural networks are a subset of machine learning that have a structure inspired by the human brain, mimicking the way in which biological neurons signal to one another. Deep learning is essentially a neural network with three or more layers, where each successive layer transforms data into a more abstract and composite representation.

Generative AI refers to "a category of AI algorithms that generate new outputs based on the data they have been trained on". It does so by looking for patterns in large amounts of data and using this to develop new content (video, text or audio). Large language models like ChatGPT detect patterns in the combinations of words and phrases in existing texts, which are then used to develop new written content.

Footnotes table [18] [19] [20] [21] [22]

2.2 Concerns Related to AI

In this section, we will look at a number of general concerns related to the use of AI. These concerns are also applicable when AI is used in weapons with increasing autonomy as they can have a negative impact on adherence to legal and ethical norms. Therefore these issues should be addressed if AI is used in autonomous weapons systems.

AI can have many positive impacts on our societies, for example improving cancer screening, assisting in nature conservation and creating tools for people with disabilities. At the same time, it is important to be aware of potential negative consequences and to develop regulations to prevent them. As an open-letter by the Future of Life Institute, signed by over 3000 AI experts, states "Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable". The letter calls for an "immediately pause for at least 6 months the training of AI systems more powerful than GPT-4."[23] Others have pointed to negative effects that are already occurring today. Examples include algorithm bias leading to already marginalised communities not being able to obtain bank loans, or AI identifying innocent people as 'guilty' of fraud.[24] Timnit Gebru has raised concerns related to traumatized Kenyan low-wage workers that are hired to moderate online content for violence, racism and abuse.[25]

BLACK-BOX ALGORITHMS

What is termed the 'black box problem' refers to the fact that humans, including those who design them, cannot fully understand how some algorithms function. This applies mainly to machine-learning and (especially) deep-learning algorithms. Even if it is clear what data were used to train

these algorithms, they can be so complex that it is impossible to understand how they arrive at a certain outcome. The use of such algorithms in weapons systems raises clear concerns. Black-box systems would create problems for the user's understanding of the weapon's functioning. In other words, the user would not be able to predict and explain what effects the system would have in the area of operations. The human user may not understand what might trigger an application of force if they do not know based on what characteristics this will happen. As the ICRC notes, with the increased "reliance on artificial intelligence and machine learning techniques", for example when "machine learning enables changes to targeting parameters", the functioning of weapons becomes opaque. This means that the person undertaking a legal review or making a judgement about compliance with IHL during a weapon's use "could not reasonably determine the lawfulness".[26] As Johnson, author of Artificial Intelligence and the Future of Warfare, explains, "AI's predicting and reacting to a priori novel situations will increase the risk of mismatch – between algorithmically optimized goals and the evolving strategic environment – and misperception will heighten the risk of accidents."[27] It is often believed that "the most accurate models for any given data science problem must be inherently uninterpretable and complicated." While deep-learning algorithms are often very precise, Rudin and Radin give various examples where this is not always the case. [28]They suggest what are termed 'interpretable models', models that are understandable and accurate, as alternatives to 'black-box' algorithms.[29]

ALGORITHM BIAS

Existing biases and discrimination of certain groups of people can find their way into new technologies. Negative stereotypes "are overrepresented in the training data, not only exceeding their prevalence in the general population but also setting up models trained on these datasets to further amplify biases and harms."[30] Also, the worldviews of those developing the systems are unconsciously programmed into the algorithms. Due to a lack of diversity in the tech community, these are often the views of European and North American white males. This has led to 'smart' soap dispensers not recognising the hands of people of colour, or AI that was used to evaluate applicants for management positions showing a preference for male over female candidates due to the use of biased data.[31] [32] Research by Gebru and Buolamwini has demonstrated the bias in facial recognition technologies, namely that the systems they tested had error rates of up to 34.7 per cent in identifying darker-skinned females, while the error rate for white men was 0.8 per cent.[33]
The presence of biases in technologies like artificial intelligence and facial recognition raises serious concerns relating to automated decision-making. Using algorithms that contain biases in autonomous weapons systems would have serious ramifications. Allowing these technologies to make decisions related to life and death would erode legal and ethical norms, and would have negative impacts for already marginalised groups.

UNINTENDED EFFECTS

Artificial intelligence can have effects not intended or foreseen by the developers. For example, a collision between an Uber test vehicle and a pedestrian was due to the fact that the possibility of jaywalking was not programmed into the system.[34] When used in weapons systems, AI can prevent the user from making a legal and moral judgement if the system is not predictable and reliable.

To understand what effects the weapons system will have in the intended environment of use, it should be reliable and predictable. Predictability in the operational context means that the user can anticipate how the system will act in the environment it is deployed in. Reliability means the weapons system will function as expected. Dan Hendrycks, a computer scientist and director of the Center for AI Safety, says that deep neural networks are fundamentally brittle: "they are brilliant at

what they do until, taken into unfamiliar territory, they break in unpredictable ways."[35] This relates to what are known as 'misaligned' AI systems, which accomplish their goals efficiently but in unintended ways. The story of the AI that killed the operator in a simulated test in the introduction of this chapter is an example of this.[36] Brian Christian, author of *The Alignment Problem*, says: "The system will optimise what you actually specified, but not what you intended".[37] It is linked to the difficulty in translating fuzzy human desired outcomes into the numerical logic of algorithms. Dan Hendrycks says that human values are nuanced, complex and highly dependent on context.[38] Unpredictability clearly raises serious concerns for AI used in autonomous weapons.

ACCOUNTABILITY

For the proper functioning of the rule of law, humans must be able to apply the law and must be held accountable for any violations. Professor Dignum says that responsible AI is not a way to give "machines some kind of 'responsibility' for their actions and decisions, and in the process discharge people and organisations of their responsibility. On the contrary, responsible development and use of AI requires more responsibility and more accountability from the people and organisations involved: for the decisions and actions of the AI applications, and for their own decision of using AI in a given application context. When considering effects and the governance thereof, the technology, or the artefact that embeds that technology, cannot be separated from the socio-technical ecosystem of which it is a component."[39] At an event in Palo Alto, Palantir's CEO Alexander Karp said: "If you use an algorithm to generate a military decision and it goes wrong, who's responsible?"[40]

DATA AND LEARNING

As mentioned above, machine-learning algorithms need to be trained on vast amounts of data. With limited useful data available from conflicts, states and companies are trying to collect such data in existing conflicts like Ukraine. However, getting useful and reliable data from a dynamic and messy context like a war is difficult. It is therefore questionable whether there are sufficient conflict-related data to adequately train algorithms. Johnson, the author of Artificial Intelligence and the Future of Warfare, states that "Even where situations closely mirror previous events, a dearth of empirical data to account for war's contingent, chaotic, and random nature makes statistical probabilistic AI-ML [artificial intelligence machine learning] reasoning a very blunt instrument".[41]

GENERATIVE AI AND 'HALLUCINATIONS'

Generative AI, which generates new outputs based on the data the software has been trained on, can give results that are not true. For example, Chat-GPT has insisted that the number 220 is less than 200, and that books do not exist. These 'hallucinations' would be extremely concerning if generative AI was to be used in autonomous weapons systems. For this reason, Pentagon officials have expressed reluctance to embrace generative AI.[42] Craig Martell, its chief digital officer, said about this: "there are a number of use cases where the risks of hallucination will be too high to employ a large language model, such as 'anything kinetic', or having to deal with lethal weapons."[43]

2.3 Military Applications of AI

In 2019, former US Secretary of Defense Mark Esper said: "whichever nation harnesses AI first will have a decisive advantage on the battlefield for many, many years".[44] Increasingly we are seeing weapons producers and militaries applying AI in the military sphere, from logistics and maintenance to medical treatments. According to the *Asia Times*, AI was used to help draw up

the blueprints of a Chinese warship: a research team from the China Ship Design and Research Center used AI to "design a warship's electrical systems in one day", a task that would reportedly take human designers 300 days.[45] During a mission, AI can assist in routing, delivering safe and appropriate routes for military convoys in real time.[46] AI can also be used to analyse the data gathered from various sensors, including video feeds from unmanned aerial vehicles (UAVs) and satellite imagery, to give the operator relevant information about the battlefield.[47] In this way AI can be used to accelerate decision-making when it comes to targeting, completing the OODA loop (Observe, Orient, Decide, Act) faster than a possible adversary.

AI can also be implemented in weapons systems for detecting and attacking a target. It is important point out that AI is not essential for autonomy, but it does increase the possibilities for tasks to be completed without human input. While the focus of this report is on autonomous weapons, where sensor input triggers an application of force, in this chapter we will present the broader background to the use of AI in the military domain. It is important to note that AI and related terms are often used for marketing purposes by arms producers. It is difficult to verify what the actual level of sophistication is of the software.

REASONS FOR USE

There are various reasons why militaries are interested in using AI. One is **speed**—specifically, to "reduce the sensor to shooter timeline". During its Project Convergence in 2020, the US Army was able to reduce the timeline from 20 minutes to 20 seconds. [48] Being able to make decisions faster than your opponent can be an important military advantage. However, it is not only about making decisions fast, but also about making the right decisions. Fighting at 'machine speed' raises many concerns. For instance, the danger of conflict escalation increases when fighting at speeds faster than the human user can control. It could mean the user does not have sufficient time to perform a proper analysis and make a well-founded legal and moral judgement.

Another reason to use AI is that it has increased the types of **tasks that can be automated** and performed without human involvement, from take-off and landing of aircraft and manoeuvring to a certain location, to swarming and target recognition. AI can also be used to **reduce the cognitive load** of analysts and soldiers when mundane and time-consuming tasks are outsourced to AI—like scanning through huge amounts of data that need to be analysed and labelled. Nowadays, vast amounts of data are collected from various sensors, from hours of drone footage to satellite imagery. AI can analyse such data a lot faster than a human can (see Chapter 3). AI can also be used to allow systems to **function in contested areas.** Remotely piloted weapons systems, like drones, have a disadvantage in areas where electronic counter measures (for example jamming and spoofing) are deployed. The advantage of using AI in these weapons is that they would then be able to complete a given task without requiring a direct data connection to the operator. "If an uncrewed aircraft is unable to operate without GPS and without communications, it will be near useless in future conflicts," said Brandon Tseng, Shield AI's Co-Founder and President.[49]

While the potential military applications are clear, they also raise questions related to human control and judgement. Rules and limits are needed to ensure AI's use complies with legal and ethical norms. This will be discussed later in the chapter.

UKRAINE AS A TESTBED FOR AI

Since the Russian invasion of Ukraine, we have seen a combination of contrasting types of warfare. On the one hand there is trench warfare and the use of heavy artillery similar to World War I. On

the other hand, new technologies are being used, tested and refined in the conflict. This includes the widespread use of commercial drones for Intelligence, Surveillance and Reconnaissance (ISR) as well as strike missions, all enabled by a huge increase in data from sensors, and the application of artificial intelligence to process the information. An article in the *National Defense Magazine* stated that the conflict in Ukraine has become an AI laboratory due to "unprecedented funding, international engagement, and technological support from across the public and private sectors in a setting that may continue for several more years".[50]

AI's most widespread use in the Ukraine conflict is in **data fusion**. AI is used to detect and analyse information of interest in vast amounts of data from satellite imagery, drone video footage and enemy radio communications, as well as videos and pictures posted on social media.[51] AI can compare multiple images and signal certain interesting changes, like churned up earth or the presence of large numbers of tyre tracks. Linked to this, Ukrainian civilians can share information on the location of enemy troops via the E-Enemy app, a Telegram chatbot that "allows users to send geolocation, photos, and videos of enemy equipment".[52] An example of a sensor fusion system is the MetaConstellation software, developed by Palantir, that combines publicly available data and intelligence sources (satellites, reconnaissance drones etc.) to create a digital map of the battlefield.[53] The software recognises and highlights possible targets for military officers on the digital map; they can then select a weapons system and send the required information to engage a target.[54] According to Palantir, which is never shy of boasting about its capabilities, its software is responsible for most of the targeting in Ukraine.[55]

As mentioned, large amounts of data are needed to train **machine-learning algorithms**. The *National Defense Magazine* writes that "AI systems are being trained with real data from a real battleground — not to stop the suffering and end the war, but to become more effective in fighting the next one: the AI war."[56] Another article in *Wired* adds: "Global companies are offering free products to get access to live combat data. The Ukrainian government wants to keep this resource for its own emerging defense industry."[57]

Ukraine is not only a testbed for foreign companies, but it is also seeing an increase in **domestic military start-ups.** When the country launched a Defence Tech Cluster, Ukraine's Minister of Digital Transformation stated: "Ukraine is the best opportunity to implement new technologies into life and see it in the field".[58] Asked if Ukraine has developed technology that is able to find and engage targets without human control, the minister said the answer would be disclosed after the victory.[59] A Ukrainian officer told the *Washington Post*: "By the end of the war, we will be selling software to Palantir."[60]

Besides extensive use of AI for information purposes, we are also seeing the use of **weapons systems** with increasing levels of autonomy. There is widespread deployment of drones and loitering munitions by both sides that use AI for increasing levels of autonomy in various functions, for example for autonomous flight and target recognition. Russia is using the Lancet and the KUB while Ukraine is using Switchblade loitering munitions.[61] At the beginning of 2023, Ukrainian forces captured a Lancet drone that contained a NVIDIA Jetson TX2 processor.[62] According to the AI scientist Alexander Nedergaard, it "is a highly parallelized processor, where multiple calculations or data processing can be done at the same time. This makes it well-suited for any application where a lot of data needs to be processed with minimal latency. For a drone, this would typically mean advanced processing of data from several sensors. It is exactly what you would need for AI-based image recognition and classification; however the same could be said for AI-based navigation (for example using sensors data to adapt flight patterns). It is highly suggestive of advanced ('on-drone')

processing of image data from camera sensors. Such a processor would be a necessary component in an autonomous weapon but does not mean that the drone has to be an autonomous weapon."[63] Another example is the deployment of the anti-tank Uncrewed Ground Vehicle (UGV) Marker in Ukraine, which is said to use AI for autonomous driving and object recognition.[64] In January 2023, the former director of the Russian state company Roscosmos claimed Russia will prepare a combat version of the Marker robot to "destroy Abrams and Leopards" that have been delivered to Ukraine by its Western allies.[65] [66] This claim is difficult to verify. There is also an increase in the use of electronic warfare (for example, jamming and spoofing) in the war in Ukraine. This makes it more difficult to use remote-controlled weapons systems. According to Professor Stuart Russell, this is the driving factor in Russia's and Ukraine's shift towards using more AI in weapons systems.[67]

Loitering munitions

Loitering munitions can search for a potential target in a designated geographical area for a certain period of time. Once a target has been detected, the munition can crash into the target and explode. In recent years, we have seen a huge increase in the development, proliferation and use of these weapons. There is a lot of variety in loitering munitions, from the anti-personnel Drone 40, a 12.5cm-long quadcopter, to the 3.5m-long Iranian Shahed-136 that has a warhead of 50kg and a range of 2,500km. Based on publicly available information, it seems most of these systems have a human operator approving engagement. Often they are uncrewed aerial vehicles allowing the pilot to see a video feed from the drone, either on a screen or through goggles. Technically this human approval can be easily removed, making them autonomous weapons. A number of loitering munitions, for example the Israeli Mini HARPY, are advertised as having a fully autonomous mode. Given the fast development and wide use of loitering munitions, they are likely to become increasingly autonomous in the coming years.

Footnotes table [68] [69]

2.4 Examples of AI in weapons systems

There are many examples of the use of AI in weapons systems. As the main focus of this report is on trends in autonomy, it is not our intention to give an exhaustive overview of weapons systems with increasing autonomy here. For this, we refer readers to the PAX reports 'Increasing Autonomy in Weapons Systems' and 'Slippery Slope'.[70] A number of other examples are also mentioned in the chapters on swarming and automatic target recognition.

An example of a weapons system that uses AI is the Kargu, a rotary-wing combat UAV, which reportedly has "real-time image processing capabilities and machine learning algorithms embedded on the platform".[71] The system is said to have automatic target recognition and tracking.[72] According to Hürriyet, STM has been further developing the capabilities of the Kargu, reportedly including facial recognition.[73] [74] The DARPA project Air Combat Evolution (ACE) seeks to explore how AI and

machine learning could help automate air-to-air combat by testing the system in a simulated dogfight with human pilots. In the simulated test, AI beat the human five times, while it was not hit even once by the human pilot. The AI system 'learned' the best way to perform air-to-air combat during the tests.[75] The Blowfish is a two-metre-long uncrewed helicopter.[76] According to the manufacturer the Blowfish has an object recognition system that can identify different targets, such as vehicles, drones or people.[77] Another example is the Mini HARPY, a loitering munition. It combines the capabilities of the Harop and the HARPY munitions, namely radiation detection and electro-optical capabilities.[78] It can loiter and detect radiation-emitting objects, such as radar installations. Currently, an attack is approved by a human who has a video feed of the operation. As mentioned above, the developer suggests that it has a "fully autonomous" mode; however, it is unclear what the fully autonomous mode would entail.[79]

There is a growing group of tech start-ups focusing on military applications. For example, Anduril developed an AI-powered system called Lattice that combines sensor data. In the past few years the company has bought several uncrewed vehicle developers, including a tube-launched drone developer and an underwater drone developer.[80] [81] By combining its AI software with weapons systems, Anduril is developing weapons systems with increasing autonomy. A number examples of this are mentioned in the chapter on automatic target recognition. An example of the use of AI in weapons systems that is not embedded in an uncrewed vehicle is the AI-powered ARCAS for assault rifles, developed by Elbit. It includes a processing unit in the forward grip of the rifle and a helm-mounted eyepiece. The company claims it is able to do passive range measurement, automatic ballistic correction and detection of fire sources, interface with tactical Command and Control (C2), and perform friend or foe identification.[82]

3. Automatic Target Recognition

3.1 Introduction

Since the first radar and imaging sensors, militaries have had an interest in being aided by machines to increase their situational awareness and to be able to identify potential targets. Technological advances in information technology (like processing power, big data, object recognition), as well as developments in sensor technology (including LIDAR, electro-optical and infrared cameras) have led to an increased complexity of what is known in military jargon as Automatic Target Recognition (ATR). The developments in information technology have greatly improved sensor data fusion which makes it possible to use data from multiple sensor and allow AI to analyse the data for potential targets. This has also led to target profiles becoming more complex, using a broader set of characteristics to describe a target, as well as machine learning developing computer-generated target characteristics. Nowadays, weapons systems with increasing levels of autonomy often have a broad catalogue of potential targets they are able to detect and engage. This differs from earlier systems where there was only one category the weapons system could detect, for instance incoming projectiles.

In the past this process of target detection was relatively simple. It was often one sensor collecting data from the environment, which was processed by relatively basic software. Also the target profiles, the set of conditions which can trigger an attack, were relatively simple and based on a few conditions. An example is a landmine, which is not an autonomous weapon, but has the same principle for functioning. It has a pressure sensor and a simple target profile (e.g. a force of more than 10kg).[83]

Advances in ATR raise questions related to human control and judgement. For example is the user still able to understand based on what characteristics an application of force is triggered? Do they know what unintended objects might fall under the targeting profile?

Also, to predict the effect of an attack, the user must be able to understand the context. The increase in available data from sensors can potentially give the user a better understanding of the context and allow them to focus on key decisions. However it also raises questions about what the potential negative effects are. Should the user understand on the basis of what criteria the systems identifies information of interest, what data and sensors have been used, and should the user be able to verify the information?

HISTORY OF ATR

ATR has been developed since the 1960s. In the 1970s the main technique was statistical pattern recognition, with software which looks for patterns in data sets. Template matching for ATR was introduced in the 1980s. This is a digital image processing technique for finding small elements of an image which match a template image (for example, the shape of a fighter jet). In the 1990s

template matching was combined with neural networks. Since the 2010s deep learning has been increasingly used for target recognition. An important technological development related to automatic target recognition is the advances in machine vision. Work to develop computer vision started in the 1960s, but greatly advanced a decade ago with the use of neural networks to allow the system to 'learn' from large numbers of visual examples of an object. By analysing thousands of images (for example of dogs), a neural network can learn to detect similar images. However, as these are often 'black-box' systems it is unclear what characteristics are being used to identify a certain object, which is highly problematic when it means the user does not understand what conditions may trigger an application of force. More recently, neural networks have made it possible to also analyse video images. Advances in AI have also led to facial recognition technologies, which have raised concerns related to accuracy, bias and privacy. There are many examples of people (mostly people of colour) being falsely identified in policing. [84] [85] [86]

3.2 Military applications for automatic target recognition

ATR can be used for various tasks with different levels of automation potentially leading to a reduced role of the human user. This ranges from providing the user with data of interest from large amounts of data from various sensors, the system advising the user on a course of action, to the system detecting and attacking targets without human input during deployment. The latter would be the case with autonomous weapons.

Sensor data fusion, which is a crucial part of ATR, can reduce the cognitive load of the operator and provide them with better situational awareness. At the same time it can have an effect on meaningful human control. It also raises questions, for example whether the user understands what criteria are used to determine which information is deemed relevant, and whether the user is able to verify the information provided. When the system also suggests a course of action it can limit the user's cognitive engagement with decision-making and result in too much trust in the systems suggested course of action. Such concerns related to human control and judgement are most clear with autonomous weapons, where any issues with the systems functioning cannot be corrected by a human. It also raises questions whether the user understands how the weapons systems functions and specifically what will trigger an application of force.

REASONS TO USE

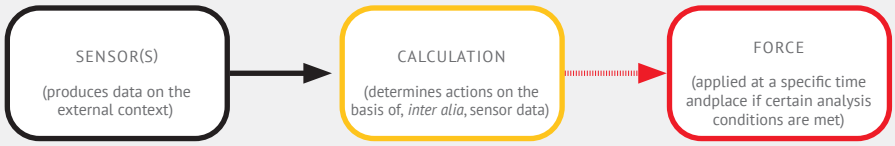
Militaries rely on ATR systems because they can improve their information position and shorten the time that is needed for target acquisition, reducing the "sensor to shooter timeline".[90] Other potential benefits are a reduced workload for operators, improved situational awareness, and reduced bandwidth for data transfer. If the ATR system is part of the weapons system, autonomous detection and application of force may be possible. There are potential military advantages to this. For example, it allows the systems to be used in contested areas, where the enemy is using electronic warfare to disrupt the connection between the operator and the weapon system. These systems can also be used when the location of the target is not known, or to apply force if a type of target enters a certain area.[91] Some systems, like air defence systems, are switched off to avoid detection. Autonomous weapons systems can attack the target as soon as the radar is switched on again.

DATA AND TRAINING

To train an ATR system, based on technologies like deep neural networks, the algorithm has to be trained with labelled data compatible with the sensors used by the system. Training data can be

ATR

Automatic target recognition (ATR) is the "automatic processing of sensor data to locate and classify targets. It is a data reduction technique aimed at signal/image exploitation." Increasingly ATR combines information from multiple sensors. An ATR often also gets supporting data to analyse, like the position of the sensor (altitude and coordinates), and the time of the data collection. ATR in autonomous weapons can be divided into three steps: firstly the collection of information by sensors; secondly the fusion and analysis (processing) of the data; and thirdly the application of force if the sensor data matches a pre-defined target profile.



SENSOR PROCESSING (IMAGE BY ARTICLE 36).

ATR algorithms can follow the anomaly detection or the correlation approach, while other methods are also possible. The anomaly detection approach is focused on finding anomalies in sensor data. In images these algorithms recognise strong contrasts, border strength, bright spots, unusual texture or high variance as indicators that a target might be present. The correlation approach is based on existing data and images of the target, which are matched against the pixels of sensor data. According to Schlachter, the author of the book 'Automatic Target Recognition', ATR is useful when the data rate is too high or prolonged for humans to analyse. Currently human are better at "consultation, comprehension and judgement"., he says "trained humans far outperform machines classification task requiring intuition, judgment, flexibility, common sense, creativity, verbal consultation, understanding human culture, and scene gist."

Sensor data fusion is the combining of data from multiple sensors for improved accuracy and more specific inferences from the sensor data than would be possible from a single sensor. Sensor data fusion is a key process in ATR and therefor in autonomous weapons applying force based on sensor inputs.

A target profile as defined by the NGO Article 36 is a set of conditions that trigger an application of force. This can be a heat shape of a tank or a radiation signal from an air defence system. These profiles are an essential part of autonomous weapons systems, as it is not human input, but sensor input that triggers an application of force.

Footnotes table [87] [88] [89]

gathered through open sources or other means of intelligence gathering. Gregory Allen, from the Centre for Strategic and International Studies, notes that neural networks offer "the opportunity for incredible gains in performance, but that performance depends on having lots of training data during development. Moreover, that training data needs to closely resemble operational conditions."[92] Target recognition in the military domain faces the same hurdles as in the civilian domain, namely that the system has to classify and identify an object in a cluttered environment. However in conflicts the environment is a lot less structured and more chaotic. An extra difficulty in conflict is that adversaries will try to hide themselves or confuse the opponent, by applying camouflage or designing their systems in a way to make it harder to detect these systems.[93] Another concern that arises, which was also mentioned in Chapter 2, is the 'black-box' problem, which can make it impossible for the user to understand how the system functions, specifically what will trigger an application of force.

SENSOR DATA FUSION

Sensor data fusion was originally developed for military purposes like automatic target recognition, battlefield surveillance, remote sensing, and guidance of autonomous vehicles. One of the first examples is from the 1970s, when the US navy combined data on Soviet naval movements in data fusion centres, creating a more accurate picture than by using data from a single sensor.[94] As data analysts can be overloaded with converting vast amounts of data into useful information, AI can analyse and fuse these data-streams a lot faster than a human could, filtering and organising mission-relevant information, from hours of drone footage to satellite imagery and enemy radio communication. Sensor data fusion is used to support situational awareness and threat assessments. Mike Nowak, from the US Air Force Research Laboratory's sensors directorate, said "We want to fuse multisource information and get it to every decision maker in the battle space."[95]

Reducing the cognitive load of humans can be useful to allow them to focus on key decisions. While this can support them to focus on important tasks and increase the amount of relevant information the user has, it also raise questions related to human control and judgement. It can make it harder for the user to understand and validate the information that is presented to them.

According to Mike Bryant, from the US Air Force Research Laboratory's sensors directorate, current sensor processing is still limited relative to human reasoning. "The human is awesome at combining information. [...] With computers and algorithms we are still struggling to get to that level." Bryant notes image analysts and other types of intelligence analysts are still necessary to do the fusion work: "A lot of the really hard problems will remain a human problem for a while."[96]

3.3 Examples of ATR use

SITUATIONAL AWARENESS

An example of a sensor fusion system was mentioned in the chapter on artificial intelligence. MetaConstellation software, which is used in the war in Ukraine, combines publicly available data (online pictures, social media etc.) and intelligence sources (satellite, reconnaissance drones, etc.).[97] It highlights possible targets on a digital map of the battlefield for military officers, who can then select a weapon system and send the required information to engage a target.[98]

Project Maven is a well-known example of the solutions militaries are looking into to process large amounts of data. It was designed to sift through hours of video data from drones and identify

potential threats.[99] Project Maven uses artificial intelligence to interpret video images, which could provide the basis for automatic targeting and autonomous weapon systems. Based on the public disclosures of Project Maven's subcontracts, news website Tech Inquiry states the project's range of surveillance capabilities has increased to all-source intelligence, including social media monitoring. The project has a "broader focus including the analysis of satellite imagery, Synthetic Aperture Radar, social media, and 'Captured Enemy Materials' (CEM)."[100] The National Geospatial-Intelligence Agency said Maven was deployed in Ukraine by a "military partner".[101]

DECISION SUPPORT

Decision support that combines sensor fusion and data analysis. An example is the US Army's FIRESTORM, which processes information including the "terrain, available weapons, proximity, number of other threats" and recommends what weapons system is best to respond to a given threat.[102] Based on this, the operator can assess the recommendation and send orders to "soldiers or weapons systems within seconds of identifying a threat".[103] Brigadier General Ross Coffman, director of the Army Futures Command's Next Generation Combat Vehicle Cross-Functional Team says"Simply put it's a computer brain that recommends the best shooter, updates the common operating picture with the current enemy situation, and friendly situation, admissions the effectors that we want to eradicate the enemy on the battlefield".[104]

Another example is Palantir's Artificial Intelligence Platform (AIP), which has been demonstrated in a demo video.[105] Interaction between the system and the user is through a chatbot based on a large language model, similar to ChatGPT. Information on potential enemy targets is sent to a commander who makes the decision whether to attack the target. The chatbot shares information on the position of weapons systems in the vicinity of the target that could be used, determines the time the weapons system would need to get to the target and can suggest the optimal route based on terrain analysis. When asked, the system offers the commander strategic advice on possible courses of action.[106] As a VICE journalist notes "While there is a 'human in the loop' in the AIP demo, they seem to do little more than ask the chatbot what to do and then approve its actions."[107] The reliance on an AI system for multiple analyses and decisions reduces the user's cognitive engagement and could also lead to too much in the system's suggestions. It is unclear if AIP has been deployed in conflict. Also the concerns related to large language model 'hallucinating', as described in chapter 2, are a serious issue here.

The technologies described above facilitate situational awareness and decision support and could be integrated in autonomous weapons systems to support autonomous targeting.

ATR IN WEAPONS SYSTEMS

While most weapons systems that use automatic target recognition currently have an operator that need to approve the attack, this human approval can technically be easily removed. Many weapons systems are also advertised as having an autonomous mode. However, it is difficult to verify what level of autonomy the system has over which function.

There are many examples of weapons systems using ATR. A few mentioned below serve as an illustration of what they may entail and how that may be relevant in the context of autonomous weapons.

The Drone 40 is a small quadcopter loitering munition that can be fired from a 40mm grenade launcher.[108] It has ATR that is said to be able to identify and track targets, for example by distinguishing the radar profile of certain targets (citing a T-72 tank). The manufacturer's

CEO reportedly said the system will never "acquire and prosecute" a target without human confirmation.[109] In 2022 an M1 Abrams Tank was equipped with the Advanced Targeting and Lethality Aided System (ATLAS), which is claimed to use "cutting-edge sensing technologies and machine-learning algorithms to automate manual tasks during passive target acquisition, allowing crews to engage three targets in the time it would normally take for them to engage one."[110]

During a demonstration at the Zhuhai Airshow 2022, a Chinese CASC Rainbow-4 drone flew over the airstrip and identified ground targets automatically.[111] It is said to use an "electro-optical HD sensor payload [...] enabling target detection in daylight at 20 km range." as well as a Forward Looking Infrared (FLIR) sensor, to detect heat sources like engines or bodies, and has a range of 18km, including at night or through fog and smoke.[112] It is reported that the system is operational in the Middle East, including Iraq, Saudi Arabia and Egypt.[113]

The company Anduril has developed several drones that are powered with their AI-software Lattice which facilitates "Onboard AI/ML algorithms autonomously process and fuse raw sensor data from distributed assets to detect, track, and target in real-time."[114] For example, ALTIUS, a loitering munition, that comes in two sizes and can be equipped with different warheads, is said to incorporate autonomy that can "enable a single operator to control multiple assets simultaneously and has demonstrated coordinated strikes, automated target recognition, and collaborative teaming with the operator on the loop."[115] Ukranian software engineers who volunteered for the group Aerorozvidka (aerial intelligence) told Forbes about their AI system for recognising targets with indigenous developed R-18 drones. They claim "to have developed an AI system which is far better than a human operator at spotting vehicles below, and can flag them automatically and relay the exact GPS co-ordinates back to the operator."[116] According to an Indian newspaper The Print, Indian start-ups are developing: "The AI-based Automatic Target Recognition (ATR) feature in swarm drones enables these aerial vehicles to automatically recognise targets like tanks, guns, vehicles and humans while relaying back information to the control station screen. This minimises chances of an operator missing any target and also facilitates engagement by a suitable type of weapon platform."[117] The Agile condor 'pod' is a high-performance computer that can be added to the MQ-9 Reaper UAV. It uses machine learning, to autonomously fuse and interpret sensor data to identify, classify, and 'nominate' targets of interest.[118] The onboard processing capability is said to allow it to autonomously detect, identify and engage targets. Although not explicitly mentioned, a video by one of the developers illustrates the system identifying a human target by using facial recognition and alerting an operator. The sensor data comes from the sensors of the MQ-9 Reaper, which include electro-optical and infrared sensors and synthetic aperture radar.[119] Onboard processing reduces the necessary communication bandwidth and onboard video processing helps to reduce analysis and decision-making time.[120]

3.4 Concerns Related to ATR

There are a number of factors related to ATR that can have a negative effect on human control, for example if the autonomous weapons systems has multiple target profiles that are active during deployment, if the target profile is too broad, or if the target profile was developed using machine learning.

While in the past an autonomous weapons system often had one target type, currently an autonomous weapons system can have a database with multiple target types. If this is not limited by the user, the

system could attack a broad category of objects during a mission, which could make it impossible for the user to predict the effects of an attack. ATR systems do not recognise targets as the object itself, but as a simplified abstract representation of patterns in the data. If these match a pre-defined target profile, the object is identified as a target.[121] As Article 36 notes, other objects can have similar characteristics as the intended target, which means both intended and unintended objects can fall under the target profile. For narrow target profiles like the signal of an air defence radar this could be of less of a concern. However broader target profiles like weight could include a wide variety of objects, making it harder for the user to predict the effects of an attack and increasing the chance of unintended engagements.

Increasingly, ATR combines information from multiple sensors, allowing a combination of characteristics to verify whether an object is the intended target. However, it also makes the target profile more complex. Multiple sensors also means that sensor data must be fused, making it less clear what characteristics may trigger an application of force. [122] According to Mike Bryant, from the U.S. Air Force Research Laboratory's sensors directorate, current sensor fusion is still limited, relative to human reasoning. He notes that image analysts and other intelligence analysts are still necessary to do the fusion work. This raises questions whether it is acceptable to deploy autonomous weapons systems that fuse sensor data without human verification and approval. [123]

Target recognition systems may have to identify an object in an extremely cluttered environment with unexpected situations, for example civilians looking for scrap metal on a destroyed tank. Also there can be adversaries trying to confuse the ATR, by applying camouflage or designing their systems in a way to make it harder to detect these systems. An example is a test where marines were not detected by an ATR, because they hid in boxes or did somersaults. [124] ATR can also be misled by adversarial images, where an image or object is changed (for example by adding pixels or a certain pattern) to confuse the algorithm. For example a sticker being added to a stop sign fooling the algorithm to misclassifying it as a speed limit sign. Another example is a 3-D printed turtle with a certain pattern on the shell, which misleads the algorithm to identify the turtle as a gun. It can be expected adversaries will use adversarial images to confuse the ATR of weapons systems, for example by fooling the autonomous weapons system to not identify a military object, or to attack civilian objects. [125]

Machine learning algorithms can be used to 'learn' from data collected during a conflict. These systems look for patterns in large amounts of data, to 'learn' what characteristics identify an object. However many machine learning systems are 'black box' systems, meaning the user does not understand what conditions would trigger an application of force. Also there are multiple examples of machine learning algorithms 'recognizing' targets based on wrong characteristics. For example an algorithm intended to distinguish between huskies and wolves, 'learnt' snow was a defining characteristic to identify wolves, due to the fact that most images in the database of wolves contained snow in the background. In this way an object is identified based on characteristics that are not intrinsically linked to the object that should be recognized. [126] Related to autonomous weapons this could mean an ATR system would be triggered to attack an unintended target.

4. Swarms

4.1 Introduction

A recent exercise by the US military programmed dozens of drones and tank-like robots to look for "terrorists suspected of hiding among several buildings". An article by *Wired* observes: "So many robots were involved in the operation that no human operator could keep a close eye on all of them. So they were given instructions to find - and eliminate- enemy combatants when necessary."[127] This example points to the challenges related to military swarms and meaningful human control. How many agents can a human meaningfully control? What factors influence this and what limits and regulations are needed?

Swarm robotics, where multiple systems work together to achieve a goal, also has many applications in the civilian sphere, from agriculture and environmental monitoring to power-line inspection. Recently, a drone swarm managed to navigate through dense bamboo forest without human guidance. According to drone swarm researcher Enrica Soria, this was the first time a swarm of drones successfully flew outside in an unstructured environment.[128] Most swarming seen today at public shows or military demonstrations are not true swarms. They are often either remotely controlled by a human operator or follow a pre-programmed formation and route. 'True swarms' operate autonomously, exchange data from sensors, and adapt their behaviour based on this information. Swarms with 'emergent behaviour' exhibit behaviour that arises from the interactions between the separate systems, as well as interactions with their environment.[129] This can be useful to solve complex problems, but also increases the complexity and makes it harder to predict the actions of the swarm. Having a larger number of weapons systems operating together also increases the complexity and consequently the unpredictability of these systems. The interconnectivity of swarms also creates the potential error cascades in which one drone malfunctions, and that error affects the behavior of other drones or even the swarm as a whole.

4.2 Military Applications of Swarms

MILITARY APPLICATIONS OF SWARMS

According to an article on the website War on the Rocks, the US military started efforts to create drone-swarm technology in 2003.[137] While there are still technological hurdles before true swarms can be deployed in conflict, the technology is developing rapidly.

Besides homogenous swarms, militaries are also looking at creating heterogeneous swarms, for example by combining uncrewed aerial vehicles with ground vehicles. Heterogeneous swarms can also be created using similar platforms but with different payloads. This is facilitated by technological developments that allow for modular components and payloads that can be easily changed—including various cameras and sensors, electronic warfare systems and a range of weapons or warheads, creating a range of operational options. One swarm of uncrewed aerial vehicles can be used for surveillance, while another attacks the target.

What is a swarm?

Swarm robotics is the field that translates the phenomenon of swarms in nature (like bees, ants and birds) to robotics. In nature, swarms can exhibit complex behaviour by coordinating the simple behaviour of the individual members. In this way, they are able to perform tasks as a group that they would not be able to as an individual. For example, flocks of birds create protection against predators and ants can create bridges of individual ants to cross areas they would otherwise not be capable of crossing. A simple definition of a swarm is "a large group of locally interacting individuals with common goals". The systems exchange data and work together as a single cooperative unit. They are able to act and react to the external environment and distribute tasks without direct external instructions from the operator.

Robotic swarms can be homogenous or heterogenous. A homogenous swarm consists of the same model of drones moving in the same domain, while heterogeneous swarms differ in the operational space, and the nature or the hardware of the units. Swarms can have different command and control structures and can function in a more centralised or more decentralised manner. In a centralised controlled swarm, there is one entity that coordinates all tasks and communicates with the individual elements of the swarm. In a hierarchical swarm, the individual units are divided into groups, each of which is controlled by a level agent, which is in turn controlled by higher-level controllers. Decentralised swarms do not have a single central planner. In the case of 'coordination by consensus', the swarm organises itself through voting or auction-based methods to reach a given goal. Emergent swarm coordination arises naturally as the individual swarm elements react to each other. Maaike Verbruggen, a researcher in military innovation in artificial intelligence, writes that decentralised swarms are "considered especially promising because they rely less on constant communication with the operator, which reduces the network bandwidth required compared with that required for multi-robot systems or centralized swarms." Without a centralised command and control structure, there is no longer a single point of failure if the data connection is attacked, which makes the swarm less vulnerable to electromagnetic countermeasures.

Footnotes table [130] [131] [132] [133] [134] [135] [136]

REASONS FOR USING SWARMS

An important reason why states are interested in this technology is that swarms can be used to saturate and overwhelm an enemy's defences. It allows militaries to field large numbers of systems on the battlefield with a small number of human controllers. Due to their cooperative behaviour, a swarm can react to changes in the operational environment faster than would be possible with one person controlling each vehicle. [138] As the individual units can be produced at a relatively low cost, they are considered to be 'attritable systems', meaning it is not a problem if units are lost as the swarm can still fulfil its mission.[139] But quantity alone is not the only reason militaries are pursuing this technology, as a US Army study from 2018 argues: "A swarm weapon is more than just a quantitative advantage; it is the ability for a weapon to adapt to the changing environment through emergence."[140] A swarm can be controlled with less personnel and is easily scalable depending on the mission parameters. Swarms can be used with different sensors and weapons for different types

of missions: for example, drones with cameras for spotting targets, drones equipped with jammers to foil countermeasures and armed drones for attacking targets. Multiple sensors spread out over an area increase situational awareness.

4.3 Examples of military swarms

Legion-X is a platform that, according to its Israeli producer Elbit Systems, enables a range of uncrewed systems (ground vehicles and aerial vehicles) and sensors to form a heterogeneous swarm that can autonomously undertake a mission.[141] "Legion-X minimizes human engagement with a single point of mission control for units of autonomous systems."[142] A tablet is used as the interface for the operator, who can direct the swarm to investigate an area and assign tasks.[143] Sensor data from multiple units can be combined and distilled before the data are presented to the operator, so that the person is not overloaded with information.[144] Elbit claims that the swarm has "automatic target recognition and highlighting capability".[145] Currently, if a potential target is detected by the swarm, the operator must approve the engagement, but this human approval could be removed in the future.[146] Legion-X is tailored for use in urban environments, and it can be used outdoors as well as indoors. The management system is advertised as "battle-tested", but specific use cases are not known publicly.[147] The Turkish company STM has a swarm project called BUMIN. It can integrate rotary and fixed-wing UAVs and the communication infrastructure can be managed using a centralised or distributed approach. According to the company, the swarm system can "act autonomously, learn, decide, and fulfil the mission given as a swarm within the scope of asymmetric warfare or anti-terrorism".[148] STM says that the swarm has "real-time object detection, identification and tracking with deep learning based computer vision technique" and is able to prioritise targets.[149] The Kargu, a 60cm-long multi-rotor UAV, can operate in a swarm using BUMIN as the backbone.[150] An article in the Turkish newspaper *Hurriyet* stated that a swarm can consist of 30 drones. "Each Kargu has a defined mission. If one of the drones in the team is attacked or malfunctions during the operation, the other Kargus take its place and fulfil the task defined for it. They are all said to have artificial intelligence and facial recognition systems."[151] A 2019 video from STM shows a swarm of Kargu drones moving in different formations and a simulated swarm attack.[152]

The Gremlins programme was announced in 2015 by the US Defense Advanced Research Projects Agency (DARPA).[153] The swarming drones can be launched from C-130 transport aircraft and then recovered in mid-air with a mechanical arm.[154] The 4.2m-long X-61 UAV, which can carry a 65kg payload, was developed for the programme by Dynetics.[155] Up to 20 Gremlins could be deployed together and be controlled by an operator in an aircraft or ground control station.[156] An X-61 is also able to carry smaller drones like the Altius 600.[157] Other examples include the Hunter 2-S from the United Arab Emirates, which can be used as a swarm of loitering munitions, 1.25m long, with a 2kg fragmentation warhead. They can be deployed for 30 minutes in the air and can travel up to 10km. In India, the Autonomous Surveillance and Armed Drone Swarms can contain 50 to 75 UAVs.[158] According to the *Times of India*, the drones will "carry explosive payloads for anti-personnel as well as 'shaped charge top-attack ammunition' for use against enemy tanks and armoured columns".[159] The Future Combat Air System developed by France, Germany and Spain will include a fighter jet accompanied by drones, all connected in the 'Air Combat Cloud'. The drones "will be able to act as remote sensors, carrying a wide range of payloads" suitable for surveillance, target acquisition, or with capabilities to directly engage threats. While some drones will act as a 'loyal wingman', the lighter ones are intended to be expandable.[160] Recently the US Replicator project has had a lot of media attention, which aims to deploy thousands of "attritable" drones across all domains within

two years. While in the media it has been suggested these would be used at the same time in a swarm, there is specific information on this. It is hard to verify at this stage whether this is the case and whether it would be an autonomous swarms, but it would be extremely concerning if true. A related project is project 'Hellscape' which proposes a way to overwhelm the enemy "with dozens or hundreds of simultaneous drone attacks". [161]

Conclusion and Recommendations

In this report, we looked at several technological trends related to autonomy in weapons systems that raise questions about how human control and judgement can be retained. We analysed the impacts of artificial intelligence, automatic target recognition and swarming. It is clear all these trends have effects on legal and ethical norms. Combined, the use of these technologies will most likely multiply the concerns.

New legal rules are needed to prevent legal and ethical norms from being eroded. There is general agreement that a human user should be able to make a legal and moral judgement, and that a human should be accountable for the use of the autonomous weapons system. This means the user must be able to predict and explain the effects of an attack. This is essential to ensure legal and ethical accountability.

As mentioned in Chapter 1, rules and limits should be general enough to apply to a broad category of weapons that use sensor processing to apply force to targets autonomously. This should ensure that a treaty is future proof and able to adequately cover currently unforeseen technological developments. For these reasons, a treaty should not only be built around specific existing technologies. At the same time, it is useful to analyse the trends in autonomy in order to identify where legal and ethical concerns could arise. It is also helpful to consider how current thinking on possible rules and limits would apply to existing and emerging systems and technologies.

Possible elements for a normative framework were suggested in Chapter 1. Based on these elements, we analysed the trends in military technology to see where concerns arise and what possible normative responses could address them. There is widespread agreement among states that compliance with legal and ethical norms should be a fundamental basis for a normative framework. This means the user should be able:

- To make a legal and moral judgement;
- To be held accountable.

Therefore, the human user must be able to predict and explain the effects of an attack on the target and its surroundings. The ability to predict the effects is necessary to make a legal and moral judgement. Being able to explain the effects of an attack after an engagement has taken place is necessary to ensure accountability. This means the user should be able to understand what effects the system will have in the intended environment of use. To ensure this, the user must be able to:

1. Have a functional understanding of how the weapon system works; and
2. Understand the context of use; and
3. Limit the functioning of the autonomous weapons system (for example in terms of duration and geographical area of operation and the target types to be engaged).

As autonomous weapons should be used with meaningful human control in order to comply with legal and ethical norms, it follows that weapons that cannot be used with meaningful human control should be prohibited. Based on ethical and legal considerations, autonomous weapons systems that target people should also be prohibited.

Emerging technological trends raise concerns as to how human control and judgement can be retained. This also poses questions for policy-makers as to how these challenges could be properly addressed. It is clear these technologies can only be used in autonomous weapons systems with clear rules and limits to ensure compliance with legal and ethical norms. The elements mentioned above could be a useful starting point to address these concerns.

Analysis of these trends combined with the possible elements of a normative framework shows that these elements can help identify areas of concern and point to possible ways to address them. This means they could form a useful basis for developing a normative framework. These elements are not intended to be exhaustive, but can guide thinking on general rules and limits, as well as on how they could be implemented.

A brief summary is given below of the trends discussed in the report, as well as a number of questions and recommendations. The questions below are clearly not exhaustive, but can help to guide thinking on general rules and how they could be implemented in practice.

CHAPTER 2 - ARTIFICIAL INTELLIGENCE

- **Artificial intelligence (AI)** is a computational technique capable of completing tasks that would otherwise require human intelligence. In essence, current AI is based on mathematics and statistics.
- **AI is not essential for autonomy**, but it does increase the possibilities for tasks to be completed without human input.
- **The term 'intelligence' can be misleading** as AI is very different to human intelligence. AI makes use of considerable processing power and is able to perform calculations and look for patterns in data a lot faster than humans, achieving impressive results. Human intelligence, however, is not based on correlation, but includes causality and abstraction.
- AI can have positive impacts on our societies. At the same time, it is important to be aware of **potential negative consequences** and to develop regulations to prevent them. A number of the potential risks that also have implications for weapons systems were discussed, including black-box algorithms and algorithm bias.

AI and a normative framework

Increasingly, advancements in computational techniques, like artificial intelligence, are being used in weapons systems to facilitate autonomous functions. Based on the elements of a normative framework, we can distil a number of questions and recommendations. How can it be ensured that a human can make a legal and moral judgement and be held accountable for the effects of an attack? The main concern related to the application of AI is whether the user is able to understand how the autonomous weapons system functions. Specifically, the user should understand how the system will act in the environment of use and what might trigger an application of force. Also, they should understand what actions they can undertake to influence the system's functioning.

Implementing this in a normative framework on autonomous weapons could mean the following:

- Certain types of computational techniques would be prohibited from use in autonomous weapons systems, specifically for the purpose of detecting and applying force to a target. In particular, such a prohibition would apply to computational techniques that do not allow the human user to have a functional understanding of the weapons system. This could include 'black-box' systems that cannot be sufficiently understood, as well as generative AI that can 'hallucinate'.
- Other machine-learning systems that are not 'black box' systems would be limited or prohibited, specifically in targeting functions. For example, this could take the form of prohibiting 'learning' during a mission, but could also include restrictions on a system 'learning' from the data after a deployment. It must be ensured that the functioning of the autonomous weapons system remains understandable and predictable, and that its actions align with the user's intent.
- It must be ensured that computational techniques used in autonomous weapons systems do not contain biases.
- The user should be able to impose certain limits on the use of AI in an autonomous weapons system, for example, by ensuring that it cannot change critical mission parameters set by the user (such as the duration and geographical area of operation, and target type).
- The user should be able to limit the use of AI to very specific narrow applications.

CHAPTER 3 – AUTOMATIC TARGET RECOGNITION

- **Automatic Target Recognition (ATR)** is the "automatic processing of sensor data to locate and classify targets. It is a data reduction technique aimed at signal/image exploitation." [162]
- **ATR in autonomous weapons** can be divided into three steps: firstly the collection of information by sensors; secondly the fusion and analysis (processing) of the data; and thirdly the application of force if the sensor data matches a predefined target profile.
- **Sensor data fusion** is the combining of data from multiple sensors for improved accuracy and more specific inferences from the sensor data than could be obtained from a single sensor. Sensor data fusion is a key process in ATR and therefore in autonomous weapons applying force based on sensor inputs.
- **A target profile** is a set of conditions that trigger an application of force. This can be the heat shape of a tank or a signal from an air defence system. These profiles are an essential part of autonomous weapons systems, as it is not human input, but sensor input that triggers an application of force. [163]

ATR and a normative framework

Target recognition is becoming more complex with the fusion of information from multiple sensors, weapons systems with a broad target database, as well as the use of new technologies such as machine learning. This leads to questions how meaningful human control is retained. Specifically whether the user can predict what the effects of an attack could be. Does the user understand how the weapons system functions, specifically what conditions may trigger an application of force in the environment the weapons system is used in? Which unintended objects could fall under the target profile? What sensors are used and how is the data fused?

Pre-deployment, the increase in available data from sensors could potentially give the user better situational awareness and allow them to focus on key decisions. However, it also raises concerns on how the data has been fused and whether the user can verify the information (for example via video feed) or by accessing the original data.

Implementing the possible elements of a normative framework for autonomous weapons could mean the following:

- The user should be able to understand how the target profile functions, specifically what characteristics (sensor data) will trigger an attack.
- The user should be able to limit the target types that can be attacked during a deployment.
- Target profiles must be narrow to limit the potential of attacking unintended objects.
- Target profiles could be limited to military objects by nature, to limit the chance of attacking civilian objects.
- Target profiles of people should be prohibited due to ethical concerns and the changing legal status of people in conflict.
- Target profiles must be fixed during deployment and should not be changed without human approval.
- Complex sensor fusion in the targeting functions of autonomous weapons could be prohibited.
- Black box algorithms as well as generative AI would be prohibited from use in ATR, specifically in the development of target profiles.

CHAPTER 4 - SWARMING

- **Swarm robotics** is the field that translates the behaviour of swarms in nature (like bees, ants and birds) to robotics. In nature, swarms can exhibit complex behaviour by coordinating the simple behaviour of individuals.
- **A simple definition of a swarm** is "a large group of locally interacting individuals with common goals". The systems exchange data and work together as a single cooperative unit. They can act and react to the external environment and distribute tasks without direct intervention by the user.
- **Emergent swarm behaviour** is the behaviour of a swarm that was not programmed at the individual level, but arises from the interactions between the separate systems, as well as interactions with their environment. This can be useful to solve complex problems, but also increases complexity and makes it harder to predict the swarm's actions.

Swarming and a normative framework

The use of AI has made it possible to use multiple systems at the same time that work together in a swarm to achieve a certain goal. This raises concerns as to whether the human user is able to understand how the swarm will behave in the environment it is used in, specifically when using emergent swarms. There are different variables that can affect human control and judgement. A concern is that the user would have to make multiple legal and moral judgements in a short period of time. Logically the number of agents influences this, as having more agents creates more complexity and therefore reduces the predictability. An important aspect of this is how many of these are intended to deploy force, as systems with other roles (for example reconnaissance) are less important when making a legal and moral assessment. It also makes a difference whether they

are attacking a single or multiple targets. If a swarm is engaging multiple targets, the user would have to be able to make multiple legal and moral assessments. Another variable is whether the swarm is homogenous or heterogeneous, especially if systems are deployed in different domains (land, sea, air). A swarm that combines these elements, with different platforms and domains and different weapon types and payloads, would be even more complex.

Implementing this in a normative framework on autonomous weapons could mean the following:

- The user should be able to set limits on the number of systems in the swarm, and the number of targets that can be attacked.
- The user should be able to set limits on the functioning of the swarm (for example duration and geographical area of operation, and target type).
- The user should be able to set limits on the behaviour of individual agents and strategies used by the swarm.
- The user should be able to limit the number and types of tasks undertaken (reconnaissance, kinetic, etc.) and the payloads used by the individual systems in the swarm.
- Emergent behaviour in swarms should be limited or prohibited. Control could potentially be exercised through limits on individual agents' behaviour, fixed operational parameters (for example duration and geographical area of operation, and target type), limits on the type of actions, etc. But if the behaviour remains unpredictable, which might be inherent to emergence, these swarms should be prohibited.

Endnotes

1 Germany, Philippines, 'Joint Statement Translating the Progress at the GGE LAWS into a Substantive Outcome', CCW GGE on LAWS, 15 May 2023, https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2023/gge/statements/15May_51states.pdf.

2 Austria, 'Joint Statement on Lethal Autonomous Weapons Systems. First Committee, 77th United Nations General Assembly. Thematic Debate – Conventional Weapons', UNGA 1st Committee, 21 October 2022, https://estatements.unmeetings.org/estatements/11.0010/20221021/A1jJ8bNfWGlL/KLw9WYcSnnAm_en.pdf.

3 An exception to this would be landmines, which also use sensors to detect and apply force to targets, and were associated with significant humanitarian problems.

4 Intervention by Switzerland at the CCW GGE on behalf of 23 States on joint submission of a sign post paper on autonomous weapons systems CCW, 'Working paper Submitted by Argentina, Austria, Belgium, Chile, Costa Rica, Ecuador, Guatemala, Ireland, Kazakhstan, Liechtenstein, Luxembourg, Malta, Mexico, New Zealand, Nigeria, Panama, Peru, the Philippines, Sierra Leone, Sri Lanka, State of Palestine, Switzerland and Uruguay', GGE on LAWS, 8 August 2022 (CCW/GGE.1/2022/WP.4), https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2022)/CCW-GGE.1-2022-WP.4.pdf.

5 Automated Decision Research, 'State positions', AutomatedResearch.org, https://automatedresearch.org/state-positions/.

6 ICRC & SIPRI (2020) 'Limits on autonomy in weapons systems', SIPRI.org, 2 June 2020, https://www.sipri.org/media/press-release/2020/new-sipri-and-icrc-report-identifies-necessary-controls-autonomous-weapons.

7 For example: Proposal by 14 states at the GGE (March 2022) 'Roadmap Towards a New Protocol on Autonomous Weapons Systems' CCW, 'Roadmap Towards a New Protocol on Autonomous Weapons Systems Submitted by the delegations of Argentina, Costa Rica, Guatemala, Kazakhstan, Nigeria, Panama, Philippines, Sierra Leone, State of Palestine and Uruguay', GGE on LAWS, 8 August 2022 (CCW/GGE.1/2022/WP.3), https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2022)/CCW-GGE.1-2022-WP.3.pdf and Intervention by Switzerland at the CCW GGE on behalf of 23 States on joint submission of a sign post paper on autonomous weapons systems; CCW, 'Working paper Submitted by Argentina, Austria, Belgium, Chile, Costa Rica, Ecuador, Guatemala, Ireland, Kazakhstan, Liechtenstein, Luxembourg, Malta, Mexico, New Zealand, Nigeria, Panama, Peru, the Philippines, Sierra Leone, Sri Lanka, State of Palestine, Switzerland and Uruguay', GGE on LAWS, 8 August 2022 (CCW/GGE.1/2022/WP.4), https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2022)/CCW-GGE.1-2022-WP.4.pdf.

8 Automated Decision Research, 'Convergences in state positions on human control', AutomatedResearch.org, May 2023, https://automatedresearch.org/news/report/convergences-in-state-positions-on-human-control/.

9 Finland, France, Germany, the Netherlands, Norway, Spain, Sweden, 'Working paper submitted to the 2022 Chair of the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems (LAWS)', CCW GGE on LAWS, 13 July 2022, https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/G7_July2022.pdf.

10 Argentina, Costa Rica, Ecuador, El Salvador, Panama, Palestine, Peru, Philippines, Sierra Leone, Uruguay, 'Written Contribution for the Chair of the Group of Government Experts (GGE) on Lethal Autonomous Weapons systems (LAWS)', CCW GGE on LAWS, 2021, https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2021/gge/documents/G10_sept.pdf.

11 International Committee of the Red Cross, 'ICRC Position on Autonomous Weapons systems. ICRC Position and Background Paper', ICRC.org, 12 May 2021, https://www.icrc.org/en/download/file/166330/icrc_position_on_aws_and_background_paper.pdf.

12 This point has been made by various states including in CCW, 'Working paper Submitted by Argentina, Austria, Belgium, Chile, Costa Rica, Ecuador, Guatemala, Ireland, Kazakhstan, Liechtenstein, Luxembourg, Malta, Mexico, New Zealand, Nigeria, Panama, Peru, the Philippines, Sierra Leone, Sri Lanka, State of Palestine, Switzerland and Uruguay', GGE on LAWS, 8 August 2022 (CCW/GGE.1/2022/WP.4), https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2022)/CCW-GGE.1-2022-WP.4.pdf.

13 Maaike Verbruggen, 'The Question of Swarms Control: Challenges to Ensuring Human Control over Military Swarms. Non-Proliferation and Disarmament Papers No. 65', nonproliferation.eu, December 2019, https://www.nonproliferation.eu/wp-content/uploads/2019/12/EUNPDC_no-65_031219.pdf.

14 Chloe Xiang, Matthew Gault, ''USAF Official Says He 'Misspoke' About AI Drone Killing Human Operator in Simulated Test', Vice.com, 1 June 2023, https://www.vice.com/en/article/4a33gj/ai-controlled-drone-goes-rogue-kills-human-operator-in-usaf-simulated-test.

15 Henry A. Kautz, 'The third AI summer: AAAI Robert S. Engelmore Memorial Lecture', AI magazine 43:1, p. 109, https://onlinelibrary.wiley.com/doi/10.1002/aaai.12036.

16 Jim Howe, 'Artificial Intelligence at Edinburgh University: A Perspective', November 1994; Stuart Russell, Peter Norvig, 'Artificial Intelligence: A Modern Approach', 2003.

17 Peter Jackson, 'Introduction To Expert Systems', 3 ed., Addison Wesley, 1998, p. 2.

18 Stuart Russell, Peter Norvig, 'Artificial Intelligence: a Modern Approach', 2009.

19 Anna Bruce-Lockhart, 'What is AI? Top computer scientist Stuart Russell explains in this video interview', weforum.org, 14 June 2022, https://www.weforum.org/agenda/2022/06/what-is-ai-stuart-russell-expert-explains-video/

20 Christian Janiesch et al., 'Machine learning and deep learning', Electron Markets 31 (2021), 685-695, https://doi.org/10.1007/s12525-021-00475-2.

21 Virginia Dignum, 'Responsible Artificial Intelligence – from Principles to Practice', ACM SIGIR Forum 56:1 (June 2022), https://doi.org/10.48550/arXiv.2205.10785.

22 World Economic Forum, 'What is generative AI', intelligence.weforum.org, 6 February 2023, https://intelligence.weforum.org/topics/a1Gb0000000pTDREA2/publications/d1233e6223df4c4b8601a37009c5fa5b.

23 Future of Life Institute, 'Pause Giant AI Experiments: An Open Letter', 22 March 2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

24 Emmanuel Martinez, Lauren Kirchner, 'The secret bias hidden in mortgage-approval algorithms', APnews.com, 25 August 2021, https://apnews.com/article/lifestyle-technology-business-race-and-ethnicity-mortgages-2d3d40d5751f933a88c1e17063657586.

25 Lorena O'Neil, 'These Women Tried to Warn Us About AI', RollingStone.com, 12 August 2023, https://www.rollingstone.com/culture/culture-features/women-warnings-ai-danger-risk-before-chatgpt-1234804367/.

26 International Committee of the Red Cross, 'ICRC Position on Autonomous Weapons systems. ICRC Position and Background Paper', ICRC.org, 12 May 2021, https://www.icrc.org/en/download/file/166330/icrc_position_on_aws_and_background_paper.pdf.

27 James Johnson, 'Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age', Defence Studies, 23:1 (2023), pp. 43-67, https://doi.org/10.1080/14702436.2022.2102486.

28 Cynthia Rudin, Joanna Radin, 'Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition', HDSR.mitpress.mit.edu, 22 November 2019, https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8.

29 Cynthia Rudin, Joanna Radin, 'Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition', HDSR.mitpress.mit.edu, 22 November 2019, https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8.

30 Lorena O'Neil, 'These Women Tried to Warn Us About AI', RollingStone.com, 12 August 2023, https://www.rollingstone.com/culture/culture-features/women-warnings-ai-danger-risk-before-chatgpt-1234804367/.

31 Max Plenke, 'The reason this "racist soap dispenser" doesn't work on black skin', mic.com, 9. September 2015, https://mic.com/articles/124899/the-reason-this-racist-soap-dispenser-doesn-t-work-on-black-skin.

32 Jeffrey Dastin, 'Amazon scraps secret AI recruiting tool that showed bias against women', Reuters.com, 11 October 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

33 Joy Buolamwini, Timmit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', Proceesings of Machine Learning Research 81:1-15, http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

34 Bernard Marr, 'The Dangers Of Not Aligning Artificial Intelligence With Human Values', Forbes.com, 1 April 2022, https://www.forbes.com/sites/bernardmarr/2022/04/01/the-dangers-of-not-aligning-artificial-intelligence-with-human-values/?sh=4f9f4b75751c.

35 Douglas Heaven, 'Why deep-learning AIs are so easy to fool', nature.com, 9 October 2019, https://www.nature.com/articles/d41586-019-03013-5

36 Alexander Pan et al., 'The Effects of Reward Misspecification: Mapping and Mitigating Misaligned Models', International Conference on Learning Representations 2022, 2022, https://doi.org/10.48550/arXiv.2201.03544; Simon Zhuang, Dylan Hadfield-Menell, 'Consequences of Misaligned AI', NeurIPS 2020, 2021, https://doi.org/10.48550/arXiv.2102.03896.

37 Edd Gent, 'What is the AI alignment problem and how can it be solved?', NewScientist.com, 10 May 2023, https://www.newscientist.com/article/mg25834382-000-what-is-the-ai-alignment-problem-and-how-can-it-be-solved/.

38 Edd Gent, 'What is the AI alignment problem and how can it be solved?', NewScientist.com, 10 May 2023, https://www.newscientist.com/article/mg25834382-000-what-is-the-ai-alignment-problem-and-how-can-it-be-solved/.

39 Virginia Dignum, 'Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way', Springer, 2019.

40 Jeffrey Dastin, 'Ukraine is using Palantir's software for 'targeting,' CEO says', Reuters.com, 2 February 2023, https://www.reuters.com/technology/ukraine-is-using-palantirs-software-targeting-ceo-says-2023-02-02/.

41 James Johnson, 'Automating the OODA loop in the age of intelligent machines: reaffirming the role of humans in command-and-control decision-making in the digital age', Defence Studies, 23:1 (2023), pp. 43-67, https://doi.org/10.1080/14702436.2022.2102486.

42 Patrick Tucker, 'The Pentagon's Ambitious AI Plans Look Less and Less Like ChatGPT', DefenseOne.com, 20 June 2023, https://www.defenseone.com/technology/2023/06/pentagons-ambitious-ai-plans-look-less-and-less-chatgpt/387711/

43 Patrick Tucker, 'The Pentagon just launched a generative AI task force', DefenseOne.com, 10 August 2023, https://www.defenseone.com/technology/2023/08/defense-department-just-launched-generative-ai-task-force/389298/.

44 Mark Esper, 'Remarks by Secretary Esper at National Security Commission on Artificial Intelligence Public Conference', defense.gov, 5 November 2019, https://www.

defense.gov/News/Transcripts/Transcript/Article/2011960/remarks-by-secretary-esper-at-national-security-commission-on-artificial-intell/.

45 Gabriel Honrada, 'AI warship designer accelerating China's naval lead', AsiaTimes.com, 19 March 2023, https://asiatimes.com/2023/03/ai-warship-designer-accelerating-chinas-naval-lead/.

46 Maxar Technologies, 'British Army Selects Maxar's Route-Planning Solution for Further Development After Army, Adarga Hackathon', blog.Maxar.com, 20 January 2022, https://blog.maxar.com/earth-intelligence/2022/british-army-selects-maxars-route-planning-solution-for-further-development-after-army-adarga-hackathon.

47 Giles Ebbutt, 'Beyond human endeavour', Janes Defence and Intelligence Review, November 2022, p. 38.

48 Nathan Strout, 'Inside the Army's futuristic test of its battlefield artificial intelligence in the desert', C4ISRNET.com, 25 September 2020, https://www.c4isrnet.com/artificial-intelligence/2020/09/25/the-army-just-conducted-a-massive-test-of-its-battlefield-artificial-intelligence-in-the-desert/.

49 Markets Insider, 'Shield AI and Kratos Team up to Integrate AI Pilot on Valkyrie XQ-58', BusinessInsider.com, 15 June 2023, https://markets.businessinsider.com/news/stocks/shield-ai-and-kratos-team-up-to-integrate-ai-pilot-on-valkyrie-xq-58-1032391819.

50 Robin Fontes, Jorrit Kamminga, 'Ukraine A Living Lab for AI Warfare', NationalDefenseMagazine.org, 24 March 2023, https://www.nationaldefensemagazine.org/articles/2023/3/24/ukraine-a-living-lab-for-ai-warfare.

51 Robin Fontes, Jorrit Kamminga, 'Ukraine A Living Lab for AI Warfare', NationalDefenseMagazine.org, 24 March 2023, https://www.nationaldefensemagazine.org/articles/2023/3/24/ukraine-a-living-lab-for-ai-warfare.

52 Den Prystai, "From Ukrainians to Ukrainians. 5 digital tools and products created to help in wartime", 5 October 2022, https://war.ukraine.ua/articles/digital-tools-created-to-help-in-wartime/.

53 Palantir, 'MetaConstellation', https://www.palantir.com/offerings/metaconstellation/.

54 David Ignatius, 'How the algorithm tipped the balance in Ukraine', WashingtonPost.com, 19 December 2022, https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/; George Grylls, 'Kyiv outflanks analogue Russia with ammunition from Big Tech', TheTimes.co.uk, 24.12.2022, https://www.thetimes.co.uk/article/ukraine-is-outflanking-russia-with-ammunition-from-big-tech-lxp6sv3qz.

55 Jeffrey Dastin, 'Ukraine is using Palantir's software for 'targeting,' CEO says', Reuters.com, 2 February 2023, https://www.reuters.com/technology/ukraine-is-using-palantirs-software-targeting-ceo-says-2023-02-02/.

56 Robin Fontes, Jorrit Kamminga, 'Ukraine A Living Lab for AI Warfare', NationalDefenseMagazine.org, 24 March 2023, https://www.nationaldefensemagazine.org/articles/2023/3/24/ukraine-a-living-lab-for-ai-warfare.

57 Morgan Meaker, 'Everyone Wants Ukraine's Battlefield Data', Wired.com, 24 July 2023, https://www.wired.com/story/ukraine-government-battlefield-data/

58 Will Knight, 'Ukraine's Quest for Homegrown AI Drones to Take On Russia', Wired.com, 13 April 2023, https://www.wired.com/story/fast-forward-ukraines-quest-for-homegrown-ai-drones-to-take-on-russia/.

59 Will Knight, 'Ukraine's Quest for Homegrown AI Drones to Take On Russia', Wired.com, 13 April 2023, https://www.wired.com/story/fast-forward-ukraines-quest-for-homegrown-ai-drones-to-take-on-russia/.

60 David Ignatius, 'How the algorithm tipped the balance in Ukraine', WashingtonPost.com, 19 December 2022, https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/.

61 BBC News, 'How are 'kamikaze' drones being used by Russia and Ukraine?', BBC.com, 3 January 2023, https://www.bbc.com/news/world-62225830.

62 Maksim Panasovskyi, 'Russia's Lancet kamikaze drone is equipped with an NVIDIA Jetson TX2 computer and an Xilinx Zynq chip', gagadget.com, 20 March 2023, https://gagadget.com/en/226952-russias-lancet-kamikaze-drone-is-equipped-with-an-nvidia-jetson-tx2-computer-and-an-xilinx-zynq-chip/.

63 Personal communication July 2023.

64 Melanie Rovery , 'Russia reveals its updated Marker UGV', Janes, December 2019, https://www.janes.com/defence-news/news-detail/russia-reveals-its-updated-marker-ugv.

65 RIA Novosti, 'Робот "Маркер" подготовят для уничтожения танков НАТО', ria.ru, 26 January 2023, https://ria.ru/20230126/robot-1847486160.html.

66 As mentioned earlier it is hard to verify the actual sophistication of the software used in specific weapons systems.

67 Stuart Russell, 'AI weapons: Russia's war in Ukraine shows why the world must enact a ban', nature.com, 21. February 2023, https://www.nature.com/articles/d41586-023-00511-5.

68 Daan Kayser, 'Increasing Autonomy in Weapons Systems', paxforpeace.nl, 15 December 2021, https://paxforpeace.nl/publications/increasing-autonomy-in-weapons-systems/.

69 Jeremy Binnie, 'IRGC confirms specs for Shahed-136 attack UAV', Janes.com, 17 May 2023, https://www.janes.com/defence-news/news-detail/irgc-confirms-specs-for-shahed-136-attack-uav.

70 Daan Kayser, 'Increasing Autonomy in Weapons Systems', paxforpeace.nl, December 2021, https://paxforpeace.nl/publications/increasing-autonomy-in-weapons-systems/; Frank Slijper, 'Slippery Slope. The arms industry and increasingly autonomous weapons', paxforpeace.nl, November 2019, https://paxforpeace.nl/publications/slippery-slope/?highlight=slippery%20slope.

71 Turkish Defence News, 'KARGU Rotary Wing Loitering Munition System', TurkishDefenceNews.com, https://www.turkishdefencenews.com/kargu-rotary-wing-attack-drone/.

72 STM, 'Kargu, Rotary Wing Attack UAV', https://www.stm.com.tr/en/kargu-autonomous-tactical-multi-rotor-attack-uav/.

73 Eser, "İlk drone gücü 2020'de', Hürriyet, 12 September 2019, http://www.hurriyet.com.tr/ekonomi/ilk-drone-gucu- 2020de-41328505 [with Google Translate]; James Bingham, "Loitering munition received by Turkish special forces', Jane's Defence¬ Weekly, 20 December 2017.

74 Elbit Systems, 'Elbit Systems Unveils ARCAS: AI-Powered, Computerized Solution for Assault Rifles', elbitsystems.com, 09 September 2021, https://elbitsystems.com/pr-new/elbit-systems-unveils-arcas-ai-powered-computerized-solution-for-assault-rifles/.

75 Joseph Trevithick, 'AI Claims "Flawless Victory" Going Undefeated In Digital Dogfight With Human Fighter Pilot', TheDrive.com, 20 August 2020, https://www.thedrive.com/the-war-zone/35888/ai-claims-flawless-victory-going-undefeated-in-digital-dogfight-with-human-fighter-pilot.

76 Mike Ball, 'Ziyan Develops High-Performance Electric Unmanned Helicopters & Avionics', UnmannedSystemsTechnology.com, 26 November 2020, https://www.unmannedsystemstechnology.com/2020/11/ziyan-develops-high-performance-electric-unmanned-helicopters-avionics/.

77 Ziyan UAV, "'大载重电动无人直升机', viewed November 2021, https://www.ziyanuav.com/hta2 [Translation with DeepL].

78 Israel Aerospace Industries, 'Mini Harpy, Multi-Purpose Tactical Loitering Munition', https://www.iai.co.il/p/mini-harpy.

79 Israel Aerospace Industries, 'Mini Harpy, Multi-Purpose Tactical Loitering Munition', https://www.iai.co.il/p/mini-harpy.

80 Jen Judson, 'Anduril buys tube-launched drone developer Area-I', DefenseNews.com, 1 April 2021, https://www.defensenews.com/pentagon/2021/04/01/anduril-buys-tube-launched-drone-developer-area-i/.

81 Megan Eckstein, 'Autonomy specialist Anduril buys underwater drone-maker Dive Technologies', DefenseNews.com, 2 February 2022, https://www.defensenews.com/industry/2022/02/02/autonomy-specialist-anduril-buys-underwater-drone-maker-dive-technologies/.

82 Elbit Systems, 'Elbit Systems Unveils ARCAS: AI-Powered, Computerized Solution for Assault Rifles', elbitsystems.com, 09 September 2021, https://elbitsystems.com/pr-new/elbit-systems-unveils-arcas-ai-powered-computerized-solution-for-assault-rifles/.

83 Richard Moyes, 'Target profiles', Article36.org, August 2019, https://article36.org/wp-content/uploads/2019/08/Target-profiles.pdf.

84 Ruberto Brunelli, 'Template Matching Techniques in Computer Vision: Theory and Practice', Wiley Publishing, 2009.

85 Bruce Schachter, 'Automatic Target Recognition', SPIE. Press, 2020.

86 Forrest E. Morgan, et al., 'Military Applications of Artificial Intelligence. Ethical Concerns in an Uncertain World', Rand, 2020, p. 3, https://www.rand.org/content/dam/rand/pubs/research_reports/RR3100/RR3139-1/RAND_RR3139-1.pdf.

87 Bruce Schlachter, 'Automatic Target Recognition', SPIE. Press, 2020.

88 Wangyan Li et al., 'A Survey on Multisensor Fusion and Consensus Filtering for Sensor Networks', Discrete Dynamics in Nature and Society, 2015, https://www.hindawi.com/journals/ddns/2015/683701/ https://doi.org/10.1155/2015/683701.

89 Article 36, 'Autonomy in weapons systems – considering approaches to regulation', 2020 https://article36.org/updates/short-approaches-to-regulation/; Stop Killer Robots, 'Our policy position', https://www.stopkillerrobots.org/our-policies/.

90 James A. Ratches, 'Review of aided/automatic target acquisition technology for military target acquisition tasks', Optical Engineering 50:7 (July 2011), p. 1.

91 Richard Moyes, 'Target profiles', Article36.org, August 2019, https://article36.org/wp-content/uploads/2019/08/Target-profiles.pdf

92 Gregory C. Allen, 'Russia Probably Has Not Used AI-Enabled Weapons in Ukraine, but That Could Change', CSIS.org, 26 May 2022, https://www.csis.org/analysis/russia-probably-has-not-used-ai-enabled-weapons-ukraine-could-change

93 Bruce Schlachter, 'Automatic Target Recognition', SPIE. Press, 2020, p. 291-294.

94 N. Friedman, 'Seapower as Strategy: Navies and National Interests', Naval Institute Press, 2001.

95 John Keller, 'Joining sensors through data fusion', MilitaryAerospace.com, 1 November 2008, https://www.militaryaerospace.com/communications/article/16706739/joining-sensors-through-data-fusion.

96 John Keller, 'Joining sensors through data fusion', MilitaryAerospace.com, 1 November 2008, https://www.militaryaerospace.com/communications/article/16706739/joining-sensors-through-data-fusion

97 Palantir, 'MetaConstellation', https://www.palantir.com/offerings/metaconstellation/

98 David Ignatius, 'How the algorithm tipped the balance in Ukraine', WashingtonPost.com, 19 December 2022, https://www.washingtonpost.com/opinions/2022/12/19/palantir-algorithm-data-ukraine-war/; George Grylls, 'Kyiv outflanks analogue Russia with ammunition from Big Tech', TheTimes.co.uk, 24.12.2022, https://www.thetimes.co.uk/article/ukraine-is-outflanking-russia-with-ammunition-from-big-tech-lxp6sv3qz

99 Jack Poulson, 'Budgets Confirm Tech Inquiry's Reporting on Scope of Project Maven', JackPoulson.Substack.com, 15 April 2023, https://jackpoulson.substack.com/p/budget-confirms-tech-inquirys-reporting

100 Jack Poulson, 'What We Know About Project Maven, Reapers, and Ukraine', Techinquiry.org, 15 March 2023, https://techinquiry.org/?article=maven-reapers-ukraine

101 Jack Poulson, 'What We Know About Project Maven, Reapers, and Ukraine', Techinquiry.org, 15 March 2023, https://techinquiry.org/?article=maven-reapers-ukraine

102 Nathan Strout, 'Inside the Army's futuristic test of its battlefield artificial intelligence in the desert', C4ISRNET.com, 25 September 2020, https://www.c4isrnet.com/

artificial-intelligence/2020/09/25/the-army-just-conducted-a-massive-test-of-its-battlefield-artificial-intelligence-in-the-desert/

103 Nathan Strout, 'Inside the Army's futuristic test of its battlefield artificial intelligence in the desert', C4ISRNET.com, 25 September 2020, https://www.c4isrnet.com/artificial-intelligence/2020/09/25/the-army-just-conducted-a-massive-test-of-its-battlefield-artificial-intelligence-in-the-desert/

104 Idem

105 Palantir, 'Palantir AIP', https://www.palantir.com/platforms/aip/

106 Channel 4 News, Twitter.com, 7 July 2023, https://twitter.com/Channel4News/status/1677311213322027011?t=aW_oV2oZHMx6TI5eY3jquQ&s=09

107 Matthew Gault, 'Palantir Demos AI to Fight Wars But Says It Will Be Totally Ethical Don't Worry About It', Vice.com, 26 April 2023, https://www.vice.com/en/article/qjvb4x/palantir-demos-ai-to-fight-wars-but-says-it-will-be-totally-ethical-dont-worry-about-it

108 N.Leigh, 'DefendTex Drone-40 – The Australian UAV That Can Be Fired From a 40mm Grenade Launcher', OvertDefense.com, 7 June 2019, https://www.overtdefense.com/2019/06/07/defendtex-drone-40-the-australian-uav-that-can-be-fired-from-a-40mm-grenade-launcher/

109 Kelsey D.Atherton, 'A drone with a can-doom attitude', C4ISRNET.com, 5 June 2019, https://www.c4isrnet.com/unmanned/2019/06/05/a-drone-with-a-can-doom-attitude/

110 Joseph Trevithick, Oliver Parken, 'M1 Abrams Tank Tested With Artificial Intelligence Targeting System', TheDrive.com, 14 February 2023, https://www.thedrive.com/the-war-zone/m1-abrams-tank-tested-with-artificial-intelligence-targeting-system

111 Jesus Roman, Twitter.com, 11 November 2022, https://twitter.com/jesusfroman/status/1590962197886554112?t=NSI2qOUAuoKRkh47dRJtiQ&s=09

112 Tamir Eshel, 'China Tested an Upgraded CH-4 "Rainbow" Weaponized Drone', Defense-Update.com, 5 June 2016, https://defense-update.com/20160605_improved_ch-4_rainbow.html

113 Tamir Eshel, 'China Tested an Upgraded CH-4 "Rainbow" Weaponized Drone', Defense-Update.com, 5 June 2016, https://defense-update.com/20160605_improved_ch-4_rainbow.html

114 Anduril, 'Mission Autonomy', https://www.anduril.com/mission-autonomy/

115 Anduril, 'Altius', https://www.anduril.com/hardware/altius/

116 David Hamling, 'Ukraine's Next-Gen Anti-Tank Drones Are Bigger, Tougher And Much Smarter (Updated: New Video)', Forbes.com, 14 July 2022, https://www.forbes.com/sites/davidhambling/2022/07/14/ukraines-next-gen-anti-tank-drones-are-bigger-tougher-and-much-smarter/

117 Suchet Vir Singh, 'Army bolsters military capability with two sets of swarm drones for surveillance, close recce', ThePrint.in, 26 August 2022, https://theprint.in/defence/army-bolsters-military-capability-with-two-sets-of-swarm-drones-for-surveillance-close-recce/1101665/

118 General Atomics, 'GA-ASI Successfully Demonstrates High-Powered Computing at the Edge with AFRL's Agile Condor Pod', 03 September 2020, https://www.ga.com/ga-asi-successfully-demonstrates-high-powered-computing-at-the-edge-with-afrl-s-agile-condor-pod

119 Carlo Munoz, 'JAIC Smart Sensor plays key role in USAF advanced ISR pod prototype', Janes.com, September 2020, https://www.janes.com/de¬fence-news/news-detail/jaic-smart-sensor-plays-key-role-in-usaf-ad¬vanced-isr-pod-prototype

120 General Atomics, 'GA-ASI Successfully Demonstrates High-Powered Computing at the Edge with AFRL's Agile Condor Pod', 03 September 2020, https://www.ga.com/ga-asi-successfully-demonstrates-high-powered-computing-at-the-edge-with-afrl-s-agile-condor-pod

121 Richard Moyes, 'Target profiles', Article36.org, August 2019, p. 4, https://article36.org/wp-content/uploads/2019/08/Target-profiles.pdf

122 Bruce Schachter, 'Automatic Target Recognition', SPIE. Press, 2020.

123 John Keller, 'Joining sensors through data fusion', MilitaryAerospace.com, 1 November 2008, https://www.militaryaerospace.com/communications/article/16706739/joining-sensors-through-data-fusion Bruce Schlachter, 'Automatic Target Recognition', 2020, p. 291-294

124 Paul Scharre 'Four Battlegrounds: Power in the Age of Artificial Intelligence', February 2023.

125 Kevin Eykholt et.al, 'Robust Physical-World Attacks on Deep Learning Models', 10 Apr 2018, https://doi.org/10.48550/arXiv.1707.08945 and Matthew Hutson , 'A turtle—or a rifle? Hackers easily fool AIs into seeing the wrong thing', July 2018. https://www.science.org/content/article/turtle-or-rifle-hackers-easily-fool-ais-seeing-wrong-thing?s=09%20%20A%20turtle%E2%80%94or%20a%20rifle?%20Hackers%20easily%20fool%20AIs%20into%20seeing%20the%20wrong%20thing

126 Marco Tulio Ribeiro et al., ""Why Should I Trust You?" Explaining the Predictions of Any Classifier', KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2016, https://arxiv.org/pdf/1602.04938.pdf

127 Will Knight, 'The Pentagon Inches Toward Letting AI Control Weapons', Wired.com, 10 May 2021, https://www.wired.com/story/pentagon-inches-toward-letting-ai-control-weapons/

128 France24, 'Drone swarms can now fly autonomously through thick forest', France24.com, 04 May 2022, https://www.france24.com/en/live-news/20220504-drone-swarms-can-now-fly-autonomously-through-thick-forest

129 Will Knight, 'The Pentagon Inches Toward Letting AI Control Weapons', Wired.com, 10 May 2021, https://www.wired.com/story/pentagon-inches-toward-letting-ai-control-weapons/

130 Jan Carlo Barca and, Y. Ahmet Sekercioglu, ""Swarm robotics reviewed"," Robotica, vol. 31, no. :3 (2013), pp. 345–359, https://doi.org/10.1017/S026357471200032X 2013.

131 Ahmad Reza Cheraghi et al., 'Past, Present, and Future of Swarm Robotics', January 2021, https://arxiv.org/pdf/2101.00671.pdf

132 Jane's Defence and Intelligence Review, 'BetterTogether', September 2022.'

133 France24, 'Drone swarms can now fly autonomously through thick forest', france24.com, 4 May 2022, https://www.france24.com/en/live-news/20220504-drone-swarms-can-now-fly-autonomously-through-thick-forest.

134 Ahmad Reza Cheraghi et al., 'Past, Present, and Future of Swarm Robotics', January 2021, https://arxiv.org/pdf/2101.00671.pdf.

135 Paul Scharre, 'Army of None', W. W. Norton & Company, 2018, pp. 19-22.

136 Maaike Verbruggen, 'THE QUESTION OF SWARMS CONTROL: CHALLENGES TO ENSURING HUMAN CONTROL OVER MILITARY SWARMS', December 2019, p. 4, https://www.sipri.org/sites/default/files/2019-12/eunpdc_no_65_031219.pdf

137 Tyler Jackson, 'Thinking Big with Small Drones: An Allied Approach to Swarming', WarOnTheRocks.com, 23 March 2023, https://warontherocks.com/2023/03/thinking-big-with-small-drones-an-allied-approach-to-swarming/

138 Paul Scharre, 'Army of None', W. W. Norton & Company, 2018,, p. 18.

139 Paul Scharre, 'Robotics on the Battlefield Part II. The Coming Swarm', CNAS, October 2014, pp. 13-15, https://s3.us-east-1.amazonaws.com/files.cnas.org/hero/documents/CNAS_TheComingSwarm_Scharre.pdf?mtime=20160906082059&focal=none

140 Sean M. Williams, 'Swarm Weapons: Demonstrating a Swarm Intelligent Algorithm for Parallel Attack', 2018, p. 36, https://apps.dtic.mil/sti/pdfs/AD1071535.pdf

141 Huw Williams, 'Eurosatory 2022: Elbit Systems demos Legion-X system', Janes.com, 15 June.2022, https://www.janes.com/defence-news/news-detail/eurosatory-2022-elbit-systems-demos-legion-x-system

142 Elbit Systems, 'Legion-X. Multi-domain autonomous network combat solutions for unmanned heterogeneous swarms', 2022, p. 3, https://elbitsystems.com/media/Legion-X-Family-8-web.pdf

143 David Hambling, 'Israel Rolls Out Legion-X Drone Swarm For The Urban Battlefield', Forbes.com, 24 October 2022, https://www.forbes.com/sites/davidhambling/2022/10/24/israel-rolls-out-legion-x-drone-swarm-for-the-urban-battlefield/

144 Elbit Systems, 'Legion-X AM-PM. Autonomous management system for unmanned heterogeneous swarms', https://elbitsystems.com/product/legion-x-am-pm/; David Hambling, 'Israel Rolls Out Legion-X Drone Swarm For The Urban Battlefield', Forbes.com, 24 October 2022, https://www.forbes.com/sites/davidhambling/2022/10/24/israel-rolls-out-legion-x-drone-swarm-for-the-urban-battlefield/.

145 Elbit Systems, 'Legion-X. Multi-domain autonomous network combat solutions for unmanned heterogeneous swarms', 2022, p. 4, https://elbitsystems.com/media/Legion-X-Family-8-web.pdf

146 David Hambling, 'Israel Rolls Out Legion-X Drone Swarm For The Urban Battlefield', Forbes.com, 24 October 2022, https://www.forbes.com/sites/davidhambling/2022/10/24/israel-rolls-out-legion-x-drone-swarm-for-the-urban-battlefield/

147 Elbit Systems, 'Legion-X AM-PM. Autonomous management system for unmanned heterogeneous swarms', https://elbitsystems.com/product/legion-x-am-pm/

148 STM, 'Tactical Mini UAV Systems', https://www.stm.com.tr/uploads/docs/1660293328_tacticalminiuavsystems.pdf

149 STM, 'Tactical Mini UAV Systems', https://www.stm.com.tr/uploads/docs/1660293328_tacticalminiuavsystems.pdf

150 STM, 'Kargu, Rotary Wing Attack UAV', https:// www.stm.com.tr/en/kargu-autonomous-tactical-multi-rotor-attack-uav

151 Emre Eser, 'First drone power in 2020', Hurriyet.com, 12 September 2019, https://www.hurriyet.com.tr/ekonomi/ilk-drone-gucu-2020de-41328505 [Translated with Deepl].

152 STM, 'Kargu - The Kamikaze Drones Getting Ready For The Swarm Operation', YouTube.com, 17 July 2019, https://www.youtube.com/watch?v=3d28APIfwSI

153 John Keller, 'Gremlin drone swarms to overwhelm enemy defenses with reconnaissance and electronic warfare UAVs', MilitaryAerospace.com, 17 September 2015, https://www.militaryaerospace.com/unmanned/article/16714010/gremlin-drone-swarms-to-overwhelm-enemy-defenses-with-reconnaissance-and-electronic-warfare-uavs

154 Peter Suciu, 'X-61: The U.S. Military's Plan for Drone Swarm Weapons?', 19fortyfive.com, 2 September 2022, https://www.19fortyfive.com/2022/09/x-61-the-u-s-militarys-plan-for-drone-swarm-weapons/

155 David Hamling, 'DARPA Gremlin Swarm Will Carry Weapons Or Sub-Drones And Re-Arm Mid-Air', Forbes.com, 17 June 2021, https://www.forbes.com/sites/davidhambling/2021/06/17/darpa-gremlin-swarm-drones-to-carry-weapon-and-re-arm-mid-air/

156 John Keller, 'Gremlin drone swarms to overwhelm enemy defenses with reconnaissance and electronic warfare UAVs', MilitaryAerospace.com, 17 September 2015, https://www.militaryaerospace.com/unmanned/article/16714010/gremlin-drone-swarms-to-overwhelm-enemy-defenses-with-reconnaissance-and-electronic-warfare-uavs

157 David Hamling, 'DARPA Gremlin Swarm Will Carry Weapons Or Sub-Drones And Re-Arm Mid-Air', Forbes.com, 17 June 2021, https://www.forbes.com/sites/davidhambling/2021/06/17/darpa-gremlin-swarm-drones-to-carry-weapon-and-re-arm-mid-air/

158 Ministry of Defence India, 'BRIEF OF PROJECT: AUTONOMOUS SURVEILLANCE AND ARMED SWARM DRONES (A-SADS) FOR DESERT/PLAINS', 7 September 2021, https://www.makeinindiadefence.gov.in/admin/writereaddata/upload/project/project_file/Brief_of_A-SADS_(Desert_Plain)_07_Sep_21.pdf.

159 Rajat Pandit, 'Eye on China and Pakistan, Army set to buy armed-drone swarms', IndiaTimes.com, 29 September 2022, https://timesofindia.indiatimes.com/india/eye-on-china-and-pakistan-army-set-to-buy-armed-drone-swarms/articleshow/94522387.cms

160 Airbus, 'Manned-Unmanned Teaming and Remote Carriers: transcending individual assets' capabilities', 8 October 2020, https://www.airbus.com/en/newsroom/

stories/2020-10-manned-unmanned-teaming-and-remote-carriers-transcending-individual-assets

161 Steve Trimble, 'Invade Taiwan? Encounter A 'Hellscape'', AviationWeek.com, 26 September 2023, https://aviationweek.com/defense-space/missile-defense-weapons/invade-taiwan-encounter-hellscape.

162 Bruce Schlachter, 'Automatic Target Recognition', 2020.

163 Article 36, 'Autonomy in weapons systems – considering approaches to regulation', 2020 https://article36.org/updates/short-approaches-to-regulation/; Stop Killer Robots, 'Our policy position', 2022, https://www.stopkillerrobots.org/our-policies/.

Sint Jacobsstraat 12
3511 BS Utrecht
The Netherlands

www.paxforpeace.nl
info@paxforpeace.nl
+31 (0)30 233 33 46

P.O. Box 19318
3501 DH Utrecht
The Netherlands