



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

**Computer Law
&
Security Review**

Legal evaluation of the attacks caused by artificial intelligence-based lethal weapon systems within the context of Rome statute



Onur Sari^{a,*}, Sener Celik^b

^aOnur Sari Department of Law, Istanbul Kent University, Cihangir Mah Siraselviler Cad No 71, Beyoglu, 34433 Istanbul, Turkey

^bSener Celik Project Coordinator, Heritage Istanbul / Heritage Akademia, Turkey

ARTICLE INFO

Keywords:

AI
Autonomy
Crime of aggression
Rome statute
IP law
IT law
International law

ABSTRACT

Artificial intelligence (AI) as of the level of development reached today has become a scientific reality that is subject to study in the fields of law, political science, and other social sciences besides computer and software engineering. AI systems which perform relatively simple tasks in the early stages of the development period are expected to become fully or largely autonomous in the near future. Thanks to this, AI which includes the concepts of machine learning, deep learning, and autonomy, has begun to play an important role in producing and using smart arms. However, questions about AI-Based Lethal Weapon Systems (AILWS) and attacks that can be carried out by such systems have not been fully answered under legal aspect. More particularly, it is a controversial issue who will be responsible for the actions that an AILWS has committed. In this article, we discussed whether AILWS can commit offense in the context of the Rome Statute, examined the applicable law regarding the responsibility of AILWS, and tried to assess whether these systems can be held responsible in the context of international law, crime of aggression, and individual responsibility. It is our finding that international legal rules including the Rome Statute can be applied regarding the responsibility for the act/crime of aggression caused by AILWS. However, no matter how advanced the cognitive capacity of an AI software, it will not be possible to resort to the personal responsibility of this kind of system since it has no legal personality at all. In such a case, responsibility will remain with the actors who design, produce, and use the system. Last but not least, since no AILWS software does have specific codes of conduct that can make legal and ethical reasonings for today, at the end of the study it was recommended that states and non-governmental organizations together with manufacturers should constitute the necessary ethical rules written in software programs to prevent these systems from unlawful acts and to develop mechanisms that would restrain AI from working outside human control.

1. Introduction

Artificial Intelligence (AI) is a concept related fundamentally to information technology area, yet it has been widely discussed almost in all fields of social sciences in recent decades. Other functions accompanying AI such as machine learn-

* Corresponding author.

E-mail address: dronursari@outlook.com (O. Sari).

ing and autonomy are broadly covered in these disciplines. It is expected that these concepts would be subject to various researches in different areas of law, particularly international law. Developments in AI-Based Lethal Weapon Systems (AILWS) inevitably led the technology to new debates on act of aggression, prohibition of force, and the right to self-defense in international law. What is surprising is that the number of studies examining AI's legal responsibility in the aforementioned areas – the act of aggression in particular – is quite few. Even though there are important studies in the literature that examine the relationship between law and technology from different perspectives, researches dealing with autonomous weapon systems in terms of act of aggression and the prohibition of the use of force relating to AI systems are not adequate. This study will try to point out this deficiency.

The basic question the answer of which is sought in the article is who will be responsible for an attack that can be carried out by AILWS. It is controversial who can be accountable for damages caused by an autonomous system used by military purposes. Within the framework of liability theory some analytical evaluations were made on human responsibility in the context of product liability, fault based liability, strict liability, and vicarious liability.¹ According to a view examining the problem in the context of United Nations regulations and European Union law claims that the person who has programmed the system can be held responsible for any damage caused by an AI system.² Two of the most important regulations associated with this approach are United Nations Convention on the Use of Electronic Communications in International Contracts and Regulating Emerging Robotic Technologies in Europe, also known as Robolaw: Robotics Facing Law and Ethics. Apart from these, the reports of the UN Group of Governmental Experts on Lethal Autonomous Weapons Systems and the Tallinn Manual are among the legal regulations that need to be taken into account about responsibility problem. On this occasion, it should be reminded that the complications of liability have been dealt with in a number of researches – although not specifically within the context of act of aggression.

In this article, we examined the act of aggression that can be carried out by AILWS in terms of jurisdiction of Rome Statute of the International Criminal Court (Rome Statute). We have dealt predominantly with act/crime of aggression in the context of Statute Article 8 *bis* with particular emphasis on *rationa persone* and *ratione loci* and the subjective element of crime. The article is consisted of four parts. In the first part, we looked at the definition and content of the AI concept. Machine learning and autonomy are also included in this part as they are closely related to the subject. In the second part, we examined the applicable law that may be relevant for AI-

based systems. This chapter, addresses international criminal law and law of armed conflict (LOAC). In the third part, we briefly examined the prohibition of the use of force and the act of aggression together with the individual criminal responsibility, while in the next part assessed how actions carried out by AILWS can be evaluated within the framework of Rome Statute Article 8 *bis*. Responsibility in terms of *ratione loci* and *ratione persone* and the mental element of crime are also the subjects of this part. In the last part, we discussed the legal/ethical measures that can be taken against the risks which may arise from an AILWS attack. As a result, we concluded that despite the technological advances and human-like cognitive characteristic features of autonomous weapon systems, these military products can not be considered as legal persons; therefore, under Statute Article 8 *bis*, people who design and use such systems for defensive and/or offensive purposes will have individual criminal responsibility in case of an armed attack launched by AI-based weapons.

2. Artificial intelligence and machine learning

Alan Turing asked for the first time in the seminal review titled *Computing Machinery and Intelligence* in 1950, "Can machines think?"³ With this question, the foundation was laid for a major change that would result in the birth of AI. Having a lot of ups and downs historically, AI could come up to its current position through a long and difficult process. In this course, the workshop held at Dartmouth College in 1950 can be considered as the first important attempt for the development process of AI. The Dartmouth Research Project defined AI as, "making a machine behave in ways that would be called intelligent if a human were so behaving."⁴ In the post-Dartmouth period, AI has become a very popular scientific field of study through which it is discussed whether machines that can act like human beings can be built both for civil and military purposes. However, AI's popularity as an applied science began to wane in the early 1960s due to the difficulty of accessing large memory capacity.⁵ In the course of the debate on the critical report submitted by James Lighthill to the British Science Research Council in 1974, the US Congress either found that the budget allocated to research was high which would cause a clear halt in AI studies.⁶

It was not possible to restart researches in AI until Japan entered a general technological leap period in the 1980s. In this decade, Japan has allocated a \$850 million budget for the 5th generation computers that can translate texts, ana-

¹ Herbert Zech, 'Liability for Autonomous Systems: Tackling Specific Risks of Modern IT' Sebastian Lohsse, Reiner Schulze, Dirk Staudenmayer (eds.), *Liability for Robotics and in the Internet of Things*, (Nomos Hart, 2018) '(forthcoming)' <<https://ssrn.com/abstract=3195676> or <http://dx.doi.org/10.2139/ssrn.3195676>> accessed 14 May 2020.

² Paulius Čerka, Jurgita Grigiene and Gintare Sirbikyte, 'Liability for Damages Caused by Artificial Intelligence' (2015) 31 *Computer Law and Security Review* 376. 387.

³ Alan Mathison Turing, 'I.—Computing Machinery And Intelligence' (1950) 59 *Mind* 433.

⁴ John McCarthy and others, 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence' (2006) 27 (4) *AI Magazine* 12.

⁵ Council of Europe, 'History Of Artificial Intelligence' (*Artificial Intelligence*, 2020) <<https://www.coe.int/en/web/artificial-intelligence/history-of-ai>> accessed 27 May 2020.

⁶ Michael Haenlein and Andreas Kaplan, 'A Brief History Of Artificial Intelligence: On The Past, Present, And Future Of Artificial Intelligence' (2019) 61 (4) *California Management Review* 5.

lyze images, and communicate with humans.⁷ In response to this move from Japan, the United States increased funds by appointing the Defense Advanced Research Projects Agency (DARPA) for AI studies which paved the way for AI to regain international popularity.⁸ In the same period, John Hopfield and David Rumelhart popularized deep learning technics while Edward Feigenbaum developed expert systems in computer science.⁹ Despite the large budgets allocated to the AI studies between 1982 and 1990, AI could not reach the desired targets determined both by companies, scientific communities, and governments. Popular interest in AI has started to decline – especially because of black box effect and programming difficulties.¹⁰ However, although AI studies have been deprived of high budgets and have been discredited research programs have continued.¹¹ With the developments in computational power and data storage in the 1990s, AI has started to be able to perform complex tasks.¹² In this sense, it may not be wrong to describe the period from the early 1990s to the present as the new rise of AI.

2.1. Artificial intelligence as a conceptual framework

There is no universally accepted definition of AI.¹³ According to Raso and others, AI is an umbrella term that includes a variety of computational techniques and associated processes.¹⁴ Noyes, on the other hand, defined AI as a broad set of methods, algorithms, and technologies that make software ‘smart’ as a way that may seem human-like to an outside observer.¹⁵

In another explanation, AI is more of a system that thinks, acts, and has its own rationality.¹⁶ In these aspects, it is different from computer algorithms.¹⁷ Based on these definition attempts, we can describe AI as a unique software that can learn and act like a human being with a complex cognitive awareness.

Whatever the definition is made, the truth is that AI’s nature and characteristic – if there any – is changing with innovation management and deployment of resources in technology. It was precisely for this reason that to create an accepted universal definition is for AI did not work well. However, in terms of making a classification related to the evolution of the system, AI can be analyzed in three types of stages: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Super Intelligence (ASI).¹⁸ Unable to autonomously solve problems in other areas, ANI applies only to specific fields¹⁹ performing certain tasks such as playing strategy games, language translation, self-driving vehicles, and image recognition.²⁰ By contrast to ANI, AGI represents ‘Human-Level AIs’ – computers smart as humans in every aspect and capable of performing all intellectual tasks that humans can.²¹ AGI is a type – or stage – of AI that can do almost all tasks that a person can do rather than performing a single task.²² For this feature, it is assumed that AGI should have a general knowledge to perform the assigned tasks.²³ As a third type, ASI means an intellect that is much smarter than the best human brains practically in every field including scientific creativity, general wisdom, and social skills.²⁴ Systems in

⁷ Gil Press, ‘A Very Short History Of Artificial Intelligence (AI)’ (*Forbes*, 30 December 2016) <<https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/>> accessed 02 Feb 2020.

⁸ Haenlein and Kaplan, (n 6).

⁹ Rockwell Anyoha, ‘The History Of Artificial Intelligence - Science In The News’ (*Science in the News*, 28 August 2017) <<http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>> accessed 27 Feb 2020.

¹⁰ Council of Europe, ‘History Of Artificial Intelligence’ (*Artificial Intelligence*, 2020) <<https://www.coe.int/en/web/artificial-intelligence/history-of-ai>> accessed 27 May 2020. As Bathaee asserts, the Black Box Problem generally can be defined as an inability to fully understand an AI’s decision-making process and the inability to predict the AI’s decisions or outputs. In this sense, the black box can be defined as an uncertainty and complexity. For Bathaee’s assessment see: Yavar Bathaee, ‘The Artificial Intelligence Black Box And The Failure of Intent And Causation’, (2018) 31 (2) *Harvard Journal of Law & Technology*, 879.

¹¹ Anyoha, (n 9).

¹² Organisation for Economic Co-operation and Development (OECD) ‘Artificial Intelligence in Society’, (Paris, 2019) OECD Publishing, Paris, <<https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm>> accessed 28 March 2020, 20.

¹³ *ibid* 22.

¹⁴ Filippo Raso and others, ‘Artificial Intelligence & Human Rights: Opportunities & Risks’ (2018) *The Berkman Klein Center for Internet & Society Research Publication Series*, Research Publication No.2018/6, 10.

¹⁵ Katherine Noyes, ‘5 Things You Need To Know About A.I.: Cognitive, Neural And Deep, Oh My!’ (*Computerworld*, 3 March 2016) <<https://www.computerworld.com/article/3040563/5-things-you-need-to-know-about-ai-cognitive-neural-and-deep-oh-my.html>> accessed 27 May 2020

¹⁶ Čerka, Grigiene and Sirbikyte, (n 2) 387.

¹⁷ Čerka, Grigiene and Sirbikyte, (n 2) 387.

¹⁸ The doctrine also distinguishes as Weak AI vs Strong AI. Weak AI is identified with the concept of Artificial Narrow Intelligence, while Strong AI is identified with the concept of Artificial General Intelligence. See; Kathleen Walch, ‘Rethinking Weak Vs. Strong AI’ (*Forbes*, 4 October 2019) <<https://www.forbes.com/sites/cognitiveworld/2019/10/04/rethinking-weak-vs-strong-ai/>> accessed 27 May 2020.

¹⁹ Andreas Kaplan and Michael Haenlein, ‘Siri, Siri, In My Hand: Who’S The Fairest In The Land? On The Interpretations, Illustrations, And Implications Of Artificial Intelligence’ (2019) 62 (1) *Business Horizons* 15.

²⁰ National Science and Technology Council: Committee on Technology, ‘Preparing for the Future of Artificial Intelligence’ (Obama Whitehouse archives.gov, 2020). <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 28 May 2020, 7.

²¹ Gonenc Gurkaynak, Ilay Yilmaz and Gunes Haksever, ‘Stifling Artificial Intelligence: Human Perils’ (2016) 32 (5) *Computer Law & Security Review* 749.

²² Michael Chui, ‘Artificial Intelligence The Next Digital Frontier?’ (2017) <<https://www.mckinsey.com/~media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx>> accessed 28 May 2020.

²³ Jonathan Grudin and Richard Jacques, ‘Chatbots, Humbots, and the Quest for Artificial General Intelligence’ [2019] *Conference on Human Factors in Computing Systems - Proceedings*.

²⁴ Nick Bostrom, ‘How Long Before Superintelligence?’ (Nickbostrom, 25 October 1998) <<https://nickbostrom.com/superintelligence.html>> accessed 28 May 2020.

the ASI stage are theoretically thought to have superhuman capacities which include the ability to perform all kinds of tasks in many different areas and to solve various given problems quickly.²⁵ It is predicted that ASI systems will be faster than AGI and can be designed to perform almost any task, yet it is agreed that AI systems which are in practice today are still in the ANI stage, not in AGI or ASI.²⁶ Researches in the AGI stage are ongoing and it seems unlikely that ASI-based systems will be designed, manufactured and put into service in the near future.

Partly based on non-military technological achievements to date, developments in AI have inevitably started to have impact the defense industry at the beginning of the 21st century. It would not be an exaggeration to claim that in the last decade offensive and/or defensive swarm technology (collective behavior of decentralized, self-organized natural or artificial systems), natural or artificial autonomous and semi-autonomous weapons, training exercises, and logistics activities can be carried out via existing or potential systems based on AI.²⁷ Phalanx Close-In Weapon Systems (CIWS) of US Navy, Patriot air defense system of US Army/Air Force, Israel defence Forces' Iron Dome air defense system, and autonomous KMAX helicopters used by several states including United States can fly on pre-programmed routes are already operating in various degrees of autonomy.²⁸ Countries' utilization of these and similar AI systems used for strategic defense and tactical security objectives has always been a controversial issue. Against the view that supports the use of such technology systems on the battlefield, there are also opposite views. According to the supporting view, especially with the use of autonomous robots, soldiers' lives can be protected and war costs can be reduced.²⁹ Also, it is claimed that collateral damage can be reduced by using less ammo with AI.³⁰ It is also considered that friend-or-foe and combatant-noncombatant distinction can be done more effectively with AI-based systems. In this context, it is hypothesized that civil and military losses that are experienced in armed conflicts due to human error can be prevented since autonomous weapons have

the potential to perform even better than human beings.³¹ Besides for lethal use, autonomous weapons have the capacity to be used in non-lethal activities. Among them, logistics, maintenance, base operations, veterans' healthcare, medical assistance, personnel management can be counted.³² On the other hand, the biggest concern of those who are against autonomous weapons is that these systems are devoid of compassion and empathy that only humans have. The fact that AI practices do not possess human-specific qualities such as conscience and judgment is considered to be the most important reasons of these concerns.³³ A second criticism is that defense systems using AI programs will increase the probability of an asymmetric war between states and non-state actors. It is very likely that cheap and simple AIs working with replicable software will be used by non-state actors in armed conflicts.³⁴

As machines getting more intelligent, many questions arise regarding their potential harmful impact on human societies.³⁵ It is difficult to predict in which direction and at what speed AI will evolve. Empowered by deep learning (a part of machine learning methods based on artificial neural networks) AI systems can successfully complete many tasks such as detection of objects in the images, understanding the languages, and processing of speech signals.³⁶ It is also expected that AI will show an exponential development with the advances in quantum computers.³⁷ However, given the unknown factors such as the black box effect, how the systematic of thinking of AI programs in the defense sector would be in order to make decisions by themselves is still not resolved. In most machine learning models the forecasting is defined as a black box which is a process that people minds can not understand³⁸ and this makes AI's predictions problematic in terms of accountability, predictability, and transparency.

²⁵ Kaplan and Haenlein, (n 19).

²⁶ Kelley M. Saylor, 'Defense Primer: U.S. Policy On Lethal Autonomous Weapon Systems [March 27, 2019]' (Homeland Security Digital Library, 2019) <<https://www.hsdl.org/?abstract&did=823590>> accessed 28 May 2020.

²⁷ James E. Baker, 'National Security Law & Emerging Technologies: Toward a Decisional Framework-Key Takeaways from the ABA-OSU Symposium and Jirga' (2019) 15 ISJLP 65, 70. Brackets belong to the authors.

²⁸ Jeffrey S. Thurnher, 'Examining Autonomous Weapon Systems From A Law Of Armed Conflict Perspective' in Hitoshi Nasu & Robert McLaughlin (eds), *New Technologies and the Law Of Armed Conflict* (TMC Asser Press, 2014) 213.

²⁹ Ronald C Arkin, *Governing Lethal Behavior In Autonomous Robots* (1st edn, Taylor & Francis Group 2009) 55.

³⁰ National Science and Technology Council: Committee on Technology, 'Preparing for the Future of Artificial Intelligence' (Obamawhitehouse.archives.gov, 2020) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 28 May 2020, 37.

³¹ Gregory P. Noone and Diana C. Noone, 'The Debate Over Autonomous Weapons Systems', (2015) 25 (1/6) J. Int'l L. 25, 32.

³² National Science and Technology Council: Committee on Technology, 'Preparing for the Future of Artificial Intelligence' <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 28 May 2020, 37.

³³ Bonnie Lynn Docherty, 'Shaking The Foundations: The human rights implications of killer robots', (2014) Human Rights Watch.

³⁴ Wolfgang Rudischhauser, 'Autonomous Or Semi-autonomous Weapons Systems A Potential New Threat of Terrorism?' (2017) Federal Academy for Security Policy Working Paper, No. 23/2017 <https://www.baks.bund.de/sites/baks010/files/working_paper_2017_23.pdf> accessed 02 May 2020, 1.

³⁵ Maayan Menashe and Eirini Kikarea. 'The global governance of cyberspace: reimagining private actors' accountability: introduction' (2019) 8 (2) Cambridge International Law Journal, 153, 154.

³⁶ Wojciech Samek, Thomas Wiegand and Klaus-Robert Müller, 'Explainable Artificial Intelligence: Understanding, Visualizing And Interpreting Deep Learning Models.' (Cornell University arXiv, Aug 2017) < <https://arxiv.org/abs/1708.08296> > accessed 19 March 2020. Brackets belong to authors.

³⁷ Baker, (n 27), 67.

³⁸ Cynthia Rudin, 'Stop Explaining Black Box Machine Learning Models For High Stakes Decisions And Use Interpretable Models Instead' (2019) 1 Nature Machine Intelligence. 206.

2.2. Machine learning as a component of artificial intelligence

Machine learning is an essential and significant constituent nearly of all AI systems, be it civilian or military. According to the definition in the National Science and Technology Council study, modern machine learning is a statistical process that starts with a body of data and tries to derive a rule or procedure that explains the data or can predict future data.³⁹ The definition of The Royal Society in the report on the subject by comparing AI and machine learning is important in terms of revealing the difference between the two concepts. Machine learning is the technology that allows systems to learn directly from examples, data, and experience. If the broad field of AI is the science of making machines smart, then machine learning is a technology that allows computers to perform specific tasks intelligently by learning from examples.⁴⁰ These systems are targeted to have the ability to make predictions about learning and specific fields of activity through the training sets that have been created. In machine learning generally algorithms such as supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction, and learning to learn are used.⁴¹ The difference to be considered here is, as emphasized above, AI and machine learning are not synonyms. The concept of AI is more comprehensive and broader than machine learning.

When to use machine learning is essentially a matter that can help one to understand the basis of the subject. According to Shalev-Shwartz and Shai Ben-David, there are two aspects of a given problem which may be required by programs that learn and improve on the basis of their experience. These are the complexity of program and the need for adaptivity. When there are very large and complex data sets, such as astronomical data, turning medical archives into medical knowledge, weather prediction, analysis of genomic data, web search engines, and electronic commerce, the complexity requires a machine learning rather than directly programming the computers to carry out these tasks. Secondly, many tasks change over time or from one user to another. When it is essential decoding a handwritten text, detecting spam emails programs, or recognizing speeches, machine learning programs whose behavior adapts to their input data offer a solution to such issues since they are adaptive to changes in the environment they interact with.⁴²

³⁹ National Science and Technology Council: Committee on Technology, 'Preparing for the Future of Artificial Intelligence' (Obama Whitehouse Archives, 2020). <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 28 May 2020, 8.

⁴⁰ The Royal Society, 'Machine Learning The Power And Promise Of Computers That Learn By Example' (Royal Society, 2020) <<https://royalsociety.org/-/media/policy/projects/machine-learning/publications/machine-learning-report.pdf>> accessed 28 May 2020, 16.

⁴¹ See; Taiwo Oladipupo Ayodele, 'Types of Machine Learning Algorithms' in Yagang Zhang (ed) *New Advances in Machine Learning*. (IntechOpen 2010).

⁴² Shai Shalev-Shwartz and Shai Ben-David, *Understanding Machine Learning* (Cambridge University Press 2017), 21-22.

As machine learning is developing rapidly, it may be useful to briefly explain the concept of deep learning, one of its sub-fields. As we mentioned above, deep learning is a subset of machine learning where neural networks – algorithms inspired by the human brain – learn from large amounts of data.⁴³ It is thought that as much data as possible is needed to train deep learning models. The more data loaded on an AI based system, the more successful the deep learning model will be.⁴⁴ For example, virtual assistants and web technologies such as Alexa, Google Search, and Google Photos are all based on deep learning and more usage in these system occurs more accurate informations.⁴⁵ This is because these systems learn as a result of the interaction with users.

2.3. Autonomy and autonomous systems

What autonomy is or is not has always been an interesting and complex question both for military and civilian researchers. According to Clarke, human autonomy means the capacity for independent decision and action.⁴⁶ Likewise, every system that has the feature of capability to sense, decide, and act without any human intervention has a degree of autonomy.⁴⁷ It is not necessary to use AI in every autonomous system.⁴⁸ However, with the developments in autonomy and machine learning, AI applications are moving towards the point where they can quickly decide and act without human intervention.⁴⁹

Autonomous systems is generally divided into two categories as semi-autonomous and fully-autonomous ones. According to the NATO Headquarters Supreme Allied Commander directive, machines that can perform a function entirely on their own which humans are unable to intervene are often referred to as 'fully autonomous' or the systems with 'human out of the loop'.⁵⁰ In the directive, it was indicated that au-

⁴³ IBM, 'What is deep learning, and why does it matter?' (IBM, 2020) <<https://www.ibm.com/cloud/deep-learning>> accessed 28 May 2020.

⁴⁴ SAS, 'Artificial Intelligence – What It Is And Why It Matters' (SAS Insights, 2020) <https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html> accessed 28 May 2020.

⁴⁵ *ibid.*

⁴⁶ Roger Clarke, 'Why The World Wants Controls Over Artificial Intelligence' (2019) 35 (4) *Computer Law & Security Review* 423, 426.

⁴⁷ Peter M Asaro, 'How Just Could a Robot War Be?' (2008) 175 *Frontiers in Artificial Intelligence and Applications* 50, 51.

⁴⁸ It is important to note here that in some research automated systems and autonomous systems have been made to a dual distinction. As Saylor argues, automated systems are strictly rule-based, while autonomous systems exhibit artificial intelligence. For further information see, Kelley M. Saylor, 'Defense Primer: U.S. Policy On Lethal Autonomous Weapon Systems [March 27, 2019]' (Homeland Security Digital Library, 2019) <<https://www.hsdl.org/?abstract&did=823590>> accessed 28 May 2020.

⁴⁹ Virginia Dignum, 'Responsible Artificial Intelligence: Designing AI for Human Values' [2017] *ICT Discoveries* 1 <<http://hdl.handle.net/20.500.11948/2177https://www.itu.int/en/journal/001/Pages/default.aspx>> accessed 28 May 2020.

⁵⁰ Paul D. Scharre, 'The Opportunity and Challenge Of Autonomous System Autonomous Systems: Issues for Defence Pol-

tonomy implies self-government, self-management, and decision making, then, decision choices are carried out according to pre-defined conditions or system constraints and outputs are non-deterministic, or probabilistic, and dependent on detected changes in the environment which are random.⁵¹ A semi-autonomous system, on the other hand, is a system that can operate autonomously under some conditions, yet can not always complete an entire task on their own.⁵² The biggest difference between semi-autonomy and fully-autonomy stems from full human control criterion. If there is any meaningful human control in the system, then this system will not be accepted as a fully-autonomous one.⁵³

Considering the problem in terms of our research question, it is controversial whether an AI will be held responsible for damages caused by semi-autonomous and fully-autonomous systems, and whether an act of killing executed by an AI-based system is lawful or not.⁵⁴ Reminding that we will discuss these questions below, we can refer to Szpak's explanation at this stage. According to him, AI should be prohibited from making an autonomous decision about death or life, and states that delaying such a legal arrangement will have tragic consequences.⁵⁵ In semi-autonomous systems, it is possible to hold the person responsible using or producing this system. Landmines with smart sensors can detect and detonate itself can be interpreted as a kind of semi-autonomous defense system with no machine learning.⁵⁶ These landmines can explode when they perceive movement at a certain intensity but can not decide when to be active physically. Therefore, if damage occurs caused by landmines, it is possible that the person who installed the equipment could be held responsible. This approach may also be applied to other AI-based systems working under semi-autonomy principles. Since semi-autonomous AI-based arms are subject to human intervention to execute commands, user or manufacturer may be held responsible for the damages that may occur as a result of running these systems.

In terms of the scope of our study, the main complication that should be discussed at this stage is about AI systems that performing with full-autonomy. Since these systems can act without any human intervention, it is very difficult to predict how they will behave in actual and complex armed conflict zones. As mentioned above, weapons with fully-autonomous

character are not yet available.⁵⁷ In order to use this type of autonomous weapons either for defensive or offensive purposes, it seems an AI software with very high computational power must be developed, preferably in ASI stage. Even in this condition, it may well be open to discussion whether such ASI systems will obey humans and/or become harmless.⁵⁸ According to one opinion considering the damages they may cause, fully-autonomous systems which can get out of human control should never be produced.⁵⁹ Moreover if technology reaches the level to produce ASI – which is unlikely for the foreseeable future – it can be useful in banning fully-autonomous weapons in accordance with humanitarian law and other legal norms.⁶⁰ Considering the current AI technology, it does not seem possible to design, produce, and put into service any kind of AILWS that can think and act in the ASI stage with full-autonomy.

3. Applicable law on the acts of artificial intelligence

No exclusive improvements were made to the Rome Statute specifically regarding the responsibility of AI systems – or, of any other autonomous weapons – which can be used in armed conflicts. There are no binding certain regulations in the form of peremptory norms related with the responsibility of AILWS in other basic texts of international law as well. According to Clarke, applications of new technologies are generally subject to existing laws including the various forms of commercial law, particularly contractual obligations comprising express and implied terms, consumer rights laws, and copyright and patent laws. In some contexts, such as robotics, cyborg artefact, and AI software embedded in devices, product liability laws together with the tort of negligence, human rights law, anti-discrimination law, and data protection law are likely to apply.⁶¹ As specified in literature, Article 36 of the 1977 Additional Protocol I of Geneva Convention which is among the basic legal texts of LOAC could also be applied to the use of new technologies such as AILWS.⁶² Supported this approach,

icymakers' Andrew P. Williams and Paul D. Scharre, (eds), *NATO Communications and Information Agency*, (NATO Communication and Information Agency 2015), 10.

⁵¹ *ibid* 10.

⁵² Shlomo Zilberstein, 'Building Strong Semi-Autonomous Systems' (Twenty-Ninth AAI Conference On Artificial Intelligence, Austin, January 2015) 4088.

⁵³ Jarna Maria Petman, 'Autonomous Weapons Systems and International Humanitarian Law : 'Out of the Loop?' University of Helsinki The Eric Castren Institute of International Law and Human Rights Research Reports 2017, <<https://researchportal.helsinki.fi/en/publications/autonomous-weapons-systems-and-international-humanitarian-law-out>> accessed 28 May 2020, 71.

⁵⁴ Agnieszka Szpak, 'Legality Of Use And Challenges Of New Technologies In Warfare – The Use Of Autonomous Weapons In Contemporary Or Future Wars' (2019) 28 (1) *European Review* 118.

⁵⁵ *ibid* 128.

⁵⁶ Petman, (n 53).

⁵⁷ See; Kelley M. Sayler, 'Defense Primer: U.S. Policy On Lethal Autonomous Weapon Systems [March 27, 2019]' (Homeland Security Digital Library, 2019) <<https://www.hsdl.org/?abstract&did=823590>> accessed 28 May 2020.

⁵⁸ See; Nick Bostrom, 'How Long Before Superintelligence?' (Nickbostrom, 25 October 1998) <<https://nickbostrom.com/superintelligence.html>> accessed 28 May 2020.

⁵⁹ Rudischhauser, (n 34), 2.

⁶⁰ *ibid* 2.

⁶¹ Roger Clarke, 'Regulatory Alternatives For AI' (2019) 35 (4) *Computer Law & Security Review*. 398, 401-402.

⁶² Geneva Academy, 'Autonomous Weapon Systems Under International Humanitarian And Human Rights Law - The Geneva Academy Of International Humanitarian Law And Human Rights' (Geneva-academy, 2014) <https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf> accessed 29 May 2020. 17. Article 36 of the Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International

Vivek argues that although currently there are no treaties that specifically address or govern the use of autonomous weapons, there is a universal consensus that the LOAC applies to AILWS.⁶³ However, it has also been stated that the use of autonomous weapons will weaken the entire paradigm on rule of war, and is incompatible with the main principles of LOAC such as distinction and proportionality.⁶⁴ Based on these comments, it can be concluded that LOAC could be applicable law for AILWS even it may be controversial in some ways. As the Rome Statute has jurisdiction in war crimes – together with the crime of genocide, crimes against humanity, and the crime of aggression – it seems possible to apply it for the use of AILWS as well.⁶⁵ Although the relationship between the Statute and other norms of international law is not without friction, Statute relates to an area of general international law which has recently been at the centre of the controversies.⁶⁶ Since international criminal law is a field of international law that seeks to regulate the behavior of states, organizations, and individuals operating across national boundaries in commission of international crimes, Rome Statute naturally is a part of international law.⁶⁷ Moreover, Article 21 of the Rome Statute on Applicable Law indicates this relationship asserting that, “The Court shall apply where appropriate, applicable treaties and the principles and rules of international law.”⁶⁸

Two other important authoritative sources for the use of AILWS are also worth mentioning, although they may not correspond to ‘applicable law’ on the subject technically. The first of these is the set of reports prepared in 2018 and 2019 by the Group of Governmental Experts (GGE) which was created under the structure of the United Nations.⁶⁹ In its report dated 23 October 2018, GGE announced some important principles regarding lethal autonomous weapon systems under the title ‘Possible Guiding Principles’. According to one of its basis, International Humanitarian Law (IHL) – that is to say LOAC – continues to apply fully to all weapon systems including the potential development and use of lethal autonomous weapon

systems.⁷⁰ It is important to emphasize LOAC in this context as applicable law. The GGE did not elaborate on the Geneva Conventions, and, in a similar vein, did not specify which LAOC principles could be applied. However, given the basic norms of LOAC and Customary Humanitarian Law (CHL), it can be concluded that the principles of distinction, proportionality, military necessity, and unnecessary suffering are here explicitly referred to. Another principle of the GGE is that the human responsibility for decisions on the use of weapon systems must be retained since accountability can not be transferred to machines.⁷¹ This principle is interesting in that it supports the findings we reached in this article. The GGE seems to have accepted that the machines can not be held responsible for their actions, and that the human being is primarily responsible for an act of aggression that may occur, no matter how advanced AI systems are as a result of machine learning, deep learning, or other similar processes. Another principle in the report prepared by the Group includes appropriate non-physical safeguards which emphasizes cyber-security against hacking or data spoofing by terrorist groups and the risk of proliferation.⁷² GGE discusses this principle in the context of developing or acquiring new weapon systems based on emerging technologies in the area of lethal autonomous weapon systems linking AILWS to cyber attacks. However, this linking has not been made exhaustively and clearly, therefore the principle appears to be flawed and incomplete to some extent. Moreover, the framework and scope of cyber attacks and defense technologies together with military and legal measures to be taken against them are so far from being clear. An additional guiding principle is presented in the report of GGE prepared in 2019. This one is identified “Human-machine interaction which may take various forms and be implemented at various stages of the life cycle of a weapon should ensure that the potential use of weapon systems based on emerging technologies in the area of lethal autonomous weapon systems is in compliance with applicable international law, in particular International Humanitarian Law (IHL). In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapon system as a whole.”⁷³ As clearly seen, the principle emphasized that the applicable law related to autonomous weapon systems should be LOAC, while it does not make a reference directly to the Rome Statute in the context of individual responsibility. However, it is certain that war crimes such as attacks against civilians, or causing great suffering or serious injury to body or health as a result of violation of LOAC are generally under the jurisdiction of the Court which encourages us to consider the Statute as the essential legal tool.

Armed Conflicts (Additional Protocol I) 1977 states: “In the research, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is obliged to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party.”

⁶³ Vivek Sehrawat, ‘Autonomous Weapon System: Law Of Armed Conflict (LOAC) And Other Legal Challenges’ (2017) 33 (1) Computer Law & Security Review 38.

⁶⁴ Szpak, (n 54). The main principles of LOAC are distinction, proportionality, military necessity, and unnecessary suffering.

⁶⁵ Rome Statute of the International Criminal Court art 5

⁶⁶ Andreas L. Paulus, ‘Legalist Groundwork For The International Criminal Court: Commentaries On The Statute Of The International Criminal Court’ (2003) 14 (4) European Journal of International Law 843, 850-851.

⁶⁷ Legal Information Institute, ‘International Criminal Law’ (Cornell Law School LII, 2020) <https://www.law.cornell.edu/wex/international_criminal_law> accessed 28 May 2020.

⁶⁸ Rome Statute of the International Criminal Court art 21.

⁶⁹ Group of Governmental Experts is the executive body of “The High Contracting Parties” which is a party to “The Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects”, known as CCW.

⁷⁰ ‘Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems’, (23 October 2018) CCW/GGE.1/2018/3, Possible Guiding Principles, III, A, 21(a), 4.

⁷¹ Possible Guiding Principles, III, A, 21(b), 4.

⁷² Possible Guiding Principles, III, A, 21(e), 4.

⁷³ ‘Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems’, (25 September 2019) CCW/GGE.1/2019/3, III, 16, 3.

The second authoritative source that should also be mentioned about AILWS is the Tallinn Manual on The International Law Applicable to Cyber Warfare. Prepared by the International Group of Experts at the Invitation of The NATO Cooperative Cyber defence centre of Excellence, the Manual is composed to produce a guide on the law governing cyber warfare, yet it is not an arrangement exclusively created for AI-based defense systems. The title that can be associated with the AI in the document is 'software agents' that can operate autonomously. This so-called 'agent' is a computer process managed by an operating system which performs one or more tasks on behalf of a human user. Since it is possible for these software agents to operate autonomously, or to communicate and coordinate their actions with other agents in a distributed computing environment, the case is indirectly related to AI. According to the corresponding statement in the Manual, software agents are used for executing queries across distributed repositories of information available via the World Wide Web.⁷⁴ Tallin Manual, due to the purpose, nature and the scope of its preparation, does not introduce any further regulations and binding provisions regarding AI, nor does it constitute a particular and holistic framework for AILWS. However, it is important that AI-based systems need not only have a physical structure to be defined as weapons. Non-physical elements such as 'software agents' can also be considered as weapons. Although not explicitly stated in the Manual, the use of these elements not of a physical existence in a hostile manner may be interpreted as an act of aggression.

4. Prohibition of use of force in international law and crime of aggression

Use of force is a strongly regulated conduct, regardless of whether it creates an act/crime of aggression. It is indisputably a *jus cogens* that the use of force is prohibited except for legitimate self-defense. Before evaluating the acts of AILWS in terms of the Rome Statute, we will briefly mention the prohibition of the use of force, crime of aggression, and personal criminal responsibility.

4.1. Prohibition of the use of force and crime of aggression

Under certain circumstances, the use of force has been considered legitimate by well-known ethical theorists and jurists throughout the historical process. International treaties, Customary International Law (CIL), and international norms created as a result of state practices also reveal that the war is legitimate under certain conditions. According to Saint Thomas Aquinas' theory of natural law, war was just if it fulfilled three conditions. The first is to be waged by the authorities empowered to do that (*auctoritas principis*); the second is that the cause is just (*justa causa*); and the third is that the intention is correct (*intentio recta*). Although the second condition is essential, the first one gained in importance in the passage of time because war was always linked with the sovereign, thus, the right to

wage war became an integral part of sovereignty.⁷⁵ Similarly, Hugo Grotius conceptualized the legitimacy of the war as well as the rules that must be followed in the armed conflict. He was the first to point to the difference between *jus ad bellum* and *jus in bello* redressing balance between them such as between war and peace replacing and mutually excluding each other (*inter bellum et pacem nihil est medium*).⁷⁶

Preventing states from use of force or threatening to use force against each other has a long historical evolution process. After the World War I, hostile acts and tendency to engage in a war were tried to be banned by two international regulations, the Covenant of the League of Nations and Briand-Kellog Pact. In the post-war era, the victorious allies and associated powers included Article 227 in the Treaty of Versailles, which provided a potential basis to prosecute Kaiser Wilhelm II for initiating a war of aggression. Following the World War II, the victorious allies resurrected the idea of prosecuting the waging of a war of aggression resulted in trials which German and Japanese leaders faced charges of crimes against the peace alongside charges of war crimes and crimes against humanity at the Nuremberg and Tokyo Tribunals.⁷⁷ Use of force was prohibited in the process following the war within the scope of Article 2(4) in the United Nations Treaty. In other respects, criminalizing of act of aggression was only carried out at the Kampala summit in June 2010. In this context, the international core crime of aggression defined in Article 8 of the Rome Statute is the offense of using force against another state without justification under international law.⁷⁸

The objective of this article is not to make a detailed examination on act/crime of aggression which is frequently discussed in highly qualified studies in the literature, though, it is important to note two points about act/crime of aggression in this chapter. First is not all acts of aggression would cause to crimes of aggression. If an act – an attempt to use of force – were authorized by the Security Council or permitted under Article 51, it would hardly provide the predicate for an act or crime of aggression. Two requirements which must be met before an act becomes a prosecutable crime are the violation of the Charter and manifestation of this violation.⁷⁹ Accordingly, in cases where there is no clear violation of Charter and its manifestation, a crime will not occur. Secondly, in recent years the ban on the use of force has started to be discussed in the context of two different interpretations. On one side, there is the extensive approach which consists in interpreting the rule in the most flexible manner possible. In this way, doctrines such as 'preventive self-defense', the 'implicit authorization' of the Security Council, or the right of 'humanitarian

⁷⁵ Momir Milojević, 'Prohibition Of Use Of Force And Threats in International Relations' (2001) 1 (5) Law and Politics 581, 583.

⁷⁶ *ibid* 583.

⁷⁷ Matthew Gillett, 'The Anatomy of an International Crime: Aggression at the International Criminal Court.' (2013) 13 (4) International Criminal Law Review 829.

⁷⁸ Salar Abbasi, 'A Conceptual Incongruence between International Laws of Self-Defense and the International Core Crime of Aggression', (2018) 6 (1) Penn. St. J.L. & Int'l Aff 196.

⁷⁹ Michael J. Glennon, 'The blank-prose crime of aggression.' (2010) 35 (1) Yale J. Int'l L. 71, 100.

⁷⁴ NATO Cooperative Cyber Defence Centre of Excellence 'Tallinn Manual on the International Law Applicable to Cyber Warfare' (Cambridge University Press 2013), 214.

intervention' can be accepted as conforming to the rules.⁸⁰ On the other side, is an approach which can be categorized as restrictive that favors a much stricter interpretation of the prohibition making it much less likely that new exceptions will be viewed as acceptable.⁸¹ Considering the state practices, it can be assumed that both approaches will have positive and negative impacts on the protection of international peace and security. Interpreting the concept narrowly, adherence to the sovereign state principle and the tradition of not intervening may lead to severe crimes against humanity, as witnessed in Rwanda. On the contrary, preventive and pre-emptive attacks to be carried out within the framework of broad interpretation probably will cause serious damage to international law as seen in Iraq. The potential equilibrium that jurists and political decision makers have to find seems somewhere between these two interpretations.

4.2. Personal criminal responsibility

One of the most important features of the Rome Statute is that it regulates personal criminal liability. Under the Kampala amendments to the Statute, the crime of aggression in international law is a concept which refers to the criminal responsibility of an individual; and, individual responsibility for the crime of aggression under the Statute is dependent on a determination that a state has committed an act of aggression.⁸² Article 25(1) says the Court shall have jurisdiction over natural persons pursuant to this Statute.⁸³ Additionally, Article 25(2) asserts that a person who commits a crime within the jurisdiction of the Court shall be individually responsible and liable for punishment in accordance with this Statute.⁸⁴ If a person commits such a crime jointly with another or through another person; orders, solicits or induces the commission of such a crime; aids, abets or otherwise assists in its commission or its attempted commission; and, in any other way contributes to the commission or attempted commission of such a crime by a group of persons acting with a common purpose; then, this person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the Court.⁸⁵

It can be concluded from these articles that Statute imposes four different types of individual responsibility. The first is the responsibility for the committing of a crime; second is the responsibility for ordering and investigation; third is the responsibility of assistance; fourth is the responsibility of contributing to a group crime. Statute is truly innovative and revolutionary in that it defines these separate types of responsibility in detail and judges each one. However, due to the scope and limitation of our work, we will not examine this feature

of the Statute within the scope of different types of individual responsibility further.

5. Assessment of the acts executed by AILWS within the framework of article 8 bis of Rome Status

In the next decade(s), the qualifications of AILWS may well be increase with technological innovations, which can cause the severe increase in act/crime of aggression performed by autonomous systems in armed conflicts. In this case, it will be necessary to assess whether the action in question was the act/crime of aggression examined in the context of Article 8 bis of the Statute. The definition of attack was included in Article 8 bis of the Statute which includes the regulation of the act of aggression, and it was concluded that the specific acts to be described below would constitute a crime of aggression regardless of whether or not a war was declared. According to the definition in Article 8 bis-1, crime of aggression means "the planning, preparation, initiation or execution, by a person in a position effectively to exercise control over or to direct the political or military action of a State, of an act of aggression which, by its character, gravity and scale, constitutes a manifest violation of the Charter of the United Nations."⁸⁶

As known, crime consists of two elements as criminal act (*actus reus*) and criminal intent (*mens rea*), hence, it can assume that the existence of these two elements is necessary for acceptance that an offense has been committed by an AILWS. However, there will be no direct criminal responsibility for AI systems since these systems have no mental ability like humans which may correspond to *mens rea* element. Due to the facts that international criminal tribunals' generally accepted limit to jurisdiction to natural persons; and that machines could not experience suffering or apprehend or learn from punishment, autonomous weapons are not punishable in the context of international criminal law.⁸⁷ There is no responsibility for these type of smart arms since neither domestic nor international laws recognize AI systems as legal persons.⁸⁸

The definition of the aggression is also regulated in Article 8 bis-2 in the Statute as "the use of armed force by a State against the sovereignty, territorial integrity or political independence of another State, or in any other manner inconsistent with the Charter of the United Nations".⁸⁹ Given the capabilities of the AI technologies described above, it is predictable that a specific military software constituted by machine learning or deep learning is likely to process definitive acts of planning, preparation, initiation or execution, at least theoretically. It is not a negligible possibility that an autonomous weapon with the ability of self-target assessment and self-decision making can begin to execute deliberate attacks on designated military or civilian targets especially in complex conflict zones. Additionally, according to the Article 8 bis 2(a), "The invasion or attack by the armed forces of a State

⁸⁰ Olivier Corten, 'The Controversies Over The Customary Prohibition On The Use Of Force: A Methodological Debate' (2005) 16 (5) European Journal of International Law 803, 803.

⁸¹ *ibid* 803.

⁸² Dapo Akande and Antonios Tzanakopoulos, 'The International Court of Justice and the Concept of Aggression' in Claus Krefß and Stefan Barriga, (eds) *The Crime of Aggression: A Commentary* (Cambridge University Press, 2016).

⁸³ Rome Statute of the International Criminal Court art 25(1).

⁸⁴ Rome Statute of the International Criminal Court art 25(2).

⁸⁵ Rome Statute of the International Criminal Court art 25(3).

⁸⁶ Rome Statute of the International Criminal Court art 8 bis.

⁸⁷ Docherty Bonnie Lynn, 'Mind The Gap : The Lack Of Accountability For Killer Robots' (2015) Human Rights Watch 18, 19.

⁸⁸ Gurkaynak, Yilmaz and Haksever (n 21).

⁸⁹ Rome Statute of the International Criminal Court art 8 bis-2.

of the territory of another State, or any military occupation” is also corresponds to an act of aggression.⁹⁰ However, due to software deficiencies in current AI technology and lack of necessary fund of knowledge in machine learning, annexation of and/or attack to a sovereign territory by an AILWS seems unrealistic for the present. Therefore, it would not be plausible to make a decisive assumption that an act of aggression may occur within the context of the Article 8 bis 2(a) – unless fully-autonomous systems based on AIS described above are developed and put into military service in the defense inventory of a state. Nevertheless, it should be underlined that a state can deploy some autonomous weapons to damage the territorial integrity of another state by destroying the material elements to a certain degree. In this case, the state that embodies autonomous weapons should act in accordance with the *jus ad bellum* rules where the armed attack of this state can be considered an act of aggression.⁹¹

Another criminal act is the bombardment of the territory of a state described in Article 8 bis 2(b)(v) by another state expressed as “Bombardment by the armed forces of a State against the territory of another State or the use of any weapons by a State against the territory of another State.”⁹² Such a bombardment can be done with AI-based arms such as the weapon system known as Cannon-Delivered Area Effects Munition (C-DAEM) which can make targeting without human intervention and detection is able to execute an action⁹³ that may be considered a war crime. Such an action would be essentially within the scope of the provision specified in that article. According to subclause ‘c’ of the same article, “the blockade of the ports or coasts of a State by the armed forces of another State”⁹⁴; and to subclause ‘d’, “an attack by the armed forces of a State on the land, sea or air forces, or marine and air fleets of another State” will be regarded as an act of aggression either.⁹⁵ Consisted of robot submarines that can be loaded with 12 torpedos, the CLAWS system of US Navy can be used for blockade purposes or attack missions against an enemy.⁹⁶ According to subclause ‘d,’ any military operation

that could fall into this category will be considered as an act of aggression.⁹⁷

In the subclause ‘e’ of the article, it was stated that violations of arms transfer agreements between the two states in the status of sender and recipient will also be considered as act of aggression. Under this provision, the presence of armed force or the use of armed force in violation of the agreement is considered as act of aggression despite the end of the agreement.⁹⁸ This can be particularly important for Unmanned Aerial Vehicles (UAVs, or drones) belong to a state which was deployed on the territory of another, since such UAVs may violate the Status – at least hypothetically – acting on their own or with a given remote command. Considering the content of the subclause ‘e’ of the Article 8 bis, such an act probably will be evaluated in the context of act of aggression. In a similar way, air defense operations executed by ground-to-air missiles such as Phased Array Tracking Radar to Intercept on Target (Patriot) or S-400 Triumf can be defined as AI-based attacks since they designed to act autonomously in long-range/high-altitude armed engagements with no human intervention – although they lack necessary AI software architecture that could consider them AILWS, and count as a defensive weapons rather than offensive.

The subclause ‘f’ of the article regulates the act of aggression when a state accepts the armed forces of another state and allows the accepted armed forces to attack a third state.⁹⁹ Such an attack can also be considered an act of aggression in case of AI-based attack. In Article 8 bis 2(g), it is explained that armed bands, groups, irregulars or mercenaries acting on behalf of a state carry out armed force actions against a state at the level of the above actions can be considered as an act of aggression.¹⁰⁰ In terms of AILWS, a weapon system based on this type of software will actually and certainly not be considered an armed band, group, or irregular and/or mercenary. The reason for this is that AI software can not be considered as a human, or a group of people, given contemporary technological achievements and current legal regulations. AI, as shown above, is a kind of intelligence demonstrated by machines that can be programmed by humans, regardless of whether in ANI, AGI, or ASI stages. However, if armed bands, groups, irregulars, or mercenaries execute actions using AI technology, an act of aggression may occur for certain in the context of Article 8 bis 2(g). It is common to see that these kind of attacks based on UAV technology are carried out by terrorist organizations within asymmetric warfare conditions. Information in open sources shows that Islamic State of Iraq and Syria (ISIS) attacked with UAVs by installing small bombs and grenades to the targets in Syria and Iraq,¹⁰¹ although those weapons are not products of state-of-the-art technology based on AI. Given the intent and ability of illegal non-state actors to access nuclear weapons and other advanced technology defense

⁹⁰ Rome Statute of the International Criminal Court art 8 bis (a).

⁹¹ Geneva Academy, ‘Autonomous Weapon Systems Under International Humanitarian And Human Rights Law - The Geneva Academy Of International Humanitarian Law And Human Rights’ (Geneva-academy, 2014). < https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf > accessed 29 May 2020, 10.

⁹² Rome Statute of the International Criminal Court art 8 bis (b).

⁹³ For more information about C-DAEM weapon systems; Priya Dialani, ‘AI Missiles Will Be Developed By The US Army To Locate Their Own Targets’ (Analytics Insight, 21 August 2019). <https://www.analyticsinsight.net/ai-missiles-will-be-developed-by-the-us-army-to-locate-their-own-targets/> accessed 10 May 2020.

⁹⁴ Rome Statute of the International Criminal Court art 8 bis (c).

⁹⁵ Rome Statute of the International Criminal Court art 8 bis (d).

⁹⁶ The US Navy has not revealed what CLAWS stands for as of the date this article was written. According to the information in open sources, the project is directed by the Office of Naval Research that is responsible for the science and technology programs of the US Navy and it has been allocated 26 million USD in 2020.

⁹⁷ Rome Statute of the International Criminal Court art 8 bis (d).

⁹⁸ Rome Statute of the International Criminal Court art 8 bis (e).

⁹⁹ Rome Statute of the International Criminal Court art 8 bis (f).

¹⁰⁰ Rome Statute of the International Criminal Court art 8 bis (g).

¹⁰¹ Eric Schmitt, ‘Pentagon Tests Lasers And Nets To Combat A Vexing Foe: ISIS Drones’ (Nytimes, 2017) <<https://www.nytimes.com/2017/09/23/world/middleeast/isis-drones-pentagon-experiments.html>> accessed 28 May 2020.

systems, it would not be an inaccurate assessment to assume that the attacks ISIS launched can be handled also by other armed bands, groups, irregulars, or mercenaries using highly developed AI capabilities in foreseeable future. It is obvious that with the spread of AI technology it will be cost-efficient effort to acquire the essential software and hardware to develop and use them in defensive and/or offensive military purposes. If a threat assessment is made in the context of a clear and present danger related to AILWS, it is probable that these developments may lead to increased use of autonomous vehicles by terrorist organizations such as ISIS which occasionally began to employ UAVs at the down of this decade.¹⁰²

5.1. Responsibility in terms of *ratione personae* and *ratione loci*

The Rome Statute allows us to make definite implications in terms of *ratione personae* and *ratione loci* regarding a possible attack by an AILWS. The most important regulation for an AI system under the *ratione personae* is in Rome Statute Article 25(1). The ICC has the power to judge only natural persons.¹⁰³ Article 25(2) emphasized personal responsibility in terms of punishment with the phrase “person who commits a crime shall be individually responsible and liable for punishment in accordance with this Statute”.¹⁰⁴ It is known that organizations also were tried in Nuremberg Crime Trials – one of the forerunners of the Court – and these were accepted as criminal.¹⁰⁵ In the Rome Statute however, although different opinions on the matter were put forward during the establishment of the Court, it was accepted that only real persons, not organizations could be tried.

In order to hold an AI system accountable, a legal personality must be given to it by legal authorities,¹⁰⁶ yet, such a personality has not been given to any AI program for today, neither in military nor civil area. Nevertheless, there is also an opinion emphasises that a legal personality should be given to AI considering its characteristics of intelligence, autonomous decision-making, the ability to learn from their own experience, memory, planning, complexity, formality and the ability to manipulate structures.¹⁰⁷ However, even in this case an AI system may not be judged in the context of the Rome Statute since only real persons – human beings responsible for their behaviors/misbehaviors – can be subject to jurisdiction. Regardless of the technological evolutionary stages described above, an AI system is primarily and essentially a product, a software built on hardware in the simplest terms. Just because in the near future an AI that might have situational aware-

ness can be able to think and act like a human does not mean it may be held accountable for its actions. It is evident that it would not make any sense to accept an AI software or hardware as a human and to imprison, or try to re-encode it to create deterrence and signatures for other software systems.¹⁰⁸ Even autonomous weapon systems with AI abilities, though capable of learning, can not do anything what the human beings who created them do not want them to do.¹⁰⁹ Therefore, a ‘morale agent’ will be required to be held responsible for damages caused by an AI system.¹¹⁰ This agent can be an AI software programmer, or user given comprehensive technological conditions.¹¹¹ However, it is also necessary to determine whether the violation of an LOAC principle is caused by the operator/user, or the programmer. Criminal liability of the person using any weapon system with appropriate skills may arise due to the programmer’s own error.¹¹²

It is not impossible to hold the software programmer responsible for the act of aggression caused by an AI system. As most IT experts and software engineers may agree on, the principle limitation of AI is that it learns from the data.¹¹³ There is no other way in which knowledge can be incorporated unless a revolutionary artificial awareness and completely innovative learning process emerges. If an AILWS makes an incorrect assessment, the main reason for this probably will be the error in the datasets that the system has learned from. For this reason, the responsibility will fall on the software developer who provides training from the unsuitable, erroneous, or deficient datasets. The other group that might be responsible is the users. If an act of aggression occurs as a result of the user’s intentional command, it will be almost inevitable to hold the user responsible. In this context, it is certain that the learning process and the datasets to which AI is subject to will be important.¹¹⁴ If machine learning process is performed

¹⁰² Gregory C Allen and Taniel Chan, *Artificial Intelligence And National Security* (1st edn, Cambridge, MA: Belfer Center for Science and International Affairs 2017).

¹⁰³ Rome Statute of the International Criminal Court art 25(1).

¹⁰⁴ Rome Statute of the International Criminal Court art 25(2).

¹⁰⁵ The Avalon Project, ‘Judgment : The Accused Organizations’ (Avalon Yale Law School Lillian Goldman Law Library, 2020) <<https://avalon.law.yale.edu/imt/judorg.asp>> accessed 28 May 2020.

¹⁰⁶ Čerka, Grigiene and Sirbikyte, (n 2) 383.

¹⁰⁷ Paulius Čerka, Jurgita Grigiene and Gintare Sirbikyte, ‘Is It Possible To Grant Legal Personality To Artificial Intelligence Software Systems?’ (2017) 33 (5) *Computer Law & Security Review* 685.

¹⁰⁸ See Christopher J. Spinelli, ‘The Rise of Robots: The Military’s Use of Autonomous Lethal Force’ (unpublished thesis, Air War College, Air University 2015).

¹⁰⁹ Marco Sassòli, ‘Autonomous weapons and international humanitarian law: Advantages, open technical questions and legal issues to be clarified’ (2014) 90 *International Law Studies/Naval War College*. 308, 338.

¹¹⁰ See Kathleen McKendrick, ‘Banning Autonomous Weapons Is Not The Answer’ (Chatham House, 2018) <<https://www.chathamhouse.org/expert/comment/banning-autonomous-weapons-not-answer>> accessed 28 May 2020.

¹¹¹ In the 2018 expert report of ICRC it is said that, in solving the problem who will be responsible for the damages caused by AI “assign responsibility to the operator or commander who authorizes the activation of the autonomous weapon system or programmers and manufacturers, in case of malfunction”. International Committee of the Red Cross, ‘Ethics And Autonomous Weapon Systems: An Ethical Basis For Human Control?’ (International Committee of the Red Cross, 2018) <<https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>> accessed 28 May 2020, 11.

¹¹² Sassòli, (n 110), 325.

¹¹³ SAS, ‘Artificial Intelligence – What It Is And Why It Matters’ (SAS Insights, 2020) <https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html> accessed 28 May 2020.

¹¹⁴ In the doctrine, it is stated that big data, the algorithms and software should learn human rights during machine learning and

with the inappropriate datasets – and, as a result of this process if an act of aggression are executed – it may be the responsibility of experts who determine the inputs in datasets and upload the training.

With a similar approach to the legal arrangement of *ratione personae*, ICC jurisdiction is constituted within the scope of *ratione loci* in various articles of the Rome Statute. As it is known, jurisdiction of the Court is valid only for the states that are party to the Rome Statute,¹¹⁵ and the Court is authorized even if the crime was committed on a ship or aircraft registered in the States parties.¹¹⁶ A citizen of a state party will be subject to the jurisdiction of the Court, regardless of whether the crime is committed on or outside the territory of the state party. On the other hand, a state that is not a party to the Statute can also apply to the Court and accept that its jurisdiction over act of aggression,¹¹⁷ as seen in the Ivory and Ukraine examples.¹¹⁸ Furthermore, as required by The Charter of UN Chapter VII, the Security Council may also request the Court's prosecutor to open an investigation on the grounds that an act of aggression has been committed. In this case, even if the state is not a party to the Statute or does not recognize the Court's jurisdiction, it may start a prosecution against any actor for whom there is strong evidence of committing a crime, as witnessed in the The Sudan case which was referred by the Security Council to the Court in 2005.¹¹⁹ Assuming that an AILWS will operate in the hardware on which its software is installed, it is clear that a military AI program designed for defensive or offensive purposes is considered to be located in the territory of the country in which the hardware in question is located. Therefore, the location of AI will not pose a problem of uncertainty which may be risen from its characteristic features resulting from being a non-material entity. The Statute will be valid in the country where an AILWS is installed, regardless of software's qualifications.

5.2. Mental element of crime

Article 30(1) the Statute asserts, "Unless otherwise provided, a person shall be criminally responsible and liable for punishment for a crime within the jurisdiction of the Court only if the material elements are committed with intent and knowledge".¹²⁰ As stated here, to prove the existence of an act of aggression, an act must be enforced deliberately and intentionally. As mentioned above, a software's criminal responsibility may not occur because it is accepted that mental ability can not be found in any AI system. Therefore, it is quite difficult to assume that intention or information could be present in

training. See Themistoklis Tzimas, 'The Need for an International Treaty for AI from the Perspective of Human Rights' (2019) 4 *Scientia Moralitas Journal* 73, 80.

¹¹⁵ Rome Statute of the International Criminal Court art 12 (1).

¹¹⁶ Rome Statute of the International Criminal Court art 12 (2)(a).

¹¹⁷ Rome Statute of the International Criminal Court art 12 (3).

¹¹⁸ 'Côte D'Ivoire' Coalition For The International Criminal Court' (Coalition for The ICC, 2020) <<http://www.coalitionfortheicc.org/country/cote-divoire>> accessed 28 May 2020.

¹¹⁹ See *Darfur, Sudan, Situation in Darfur, Sudan*, ICC-02/05 (<https://www.icc-cpi.int/darfur>).

¹²⁰ Rome Statute of the International Criminal Court art 30(1).

the action that an AI system performs. The responsibility lies with the user, operator, and manufacturer of the software. An AI system performs the task assigned by the user, or operator, or follows the instructions coded by its manufacturer. These actors will have an impact on the existence of a weapon system performing with the help of an AI software in the act of aggression. Inevitably, this will increase the criminal responsibility of these actors.

There is always the possibility that AI systems can act outside their scheduled tasks for unpredictable reasons. According to authorities on LOAC and criminal responsibility in armed conflict like ICRC, such unwanted acts can be caused by software errors and system flaws, human cognitive bias, in-built algorithmic bias, normal accidents where there is no clear error, deliberate hacking, and spoofing or cyber attacks.¹²¹ It is clear that the intention can not exist in such a situation. On the other hand, it can be assumed that AILWS may cause an act due to negligence in the context of act of aggression. However, since the act of aggression is deliberately committed and the intent and knowledge condition will not coexist negligence, an act that arise in this way will not be considered an act of aggression. Similarly, in the Statute, the element of a mistake of fact is considered as another factor that can remove or limit the individual responsibility. According to Article 32(1), "A mistake of fact shall be a ground for excluding criminal responsibility only if it negates the mental element required by the crime."¹²² In this context, if a fully-autonomous UAV with highly developed AI capability causes a damage by mistake attacking another state's ground target, aircraft, or ship, this action will not be considered as an act of aggression since there is no intent or knowledge. The important point here is that the mistake must be reasonable. But not to be forgotten, even an unreasonable mistake of fact can at least at times provide a defense if it negates a required specific intent.¹²³

6. Measures that can be taken regarding the actions of AILWS

Given the possible dramatic consequences of act of aggression which could include huge losses of life and property, it is clear that the manufacturer or user will need to take the necessary measures to prevent probable lethal errors of an AILWS software. First of all, a number of preventive measures can be taken within the existing international law regulations. In this context, the Geneva Conventions constitute a remarkable example in that they compose a significant reference that new types of weapons must comply with humanitarian law. States

¹²¹ International Committee of the Red Cross, 'Ethics And Autonomous Weapon Systems: An Ethical Basis For Human Control?' (International Committee of the Red Cross, 2018) <<https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>> accessed 28 May 2020, 14.

¹²² Rome Statute of the International Criminal Court art 32(1).

¹²³ John McCurley, 'Is "Mistake of Fact" a Defense to Criminal Charges?' (Nolo, 2020) <<https://www.nolo.com/legal-encyclopedia/is-mistake-fact-defense-criminal-charges.html>> accessed 28 May 2020.

would also rely on both CIL and IHL for making law in autonomous weapon systems which will be used in space as well as in earth.¹²⁴

Secondly, the precaution can be taken to determine the specific ethic rules for AI programs and code them into the related software. In this sense, it is obvious that experts working on machine learning, deep learning, and other related research fields will need to consider morality together with rule of law when designing software. The program should be written on the basis of the set of ethical rules and values which have been obeyed and interiorized for centuries in war based on *jus in bello*. Ethics, also called moral philosophy, means the discipline concerned with what is morally good and bad and morally right and wrong.¹²⁵ In the same vein, robo-ethics describes the ethics and morals of robotics, the science of robots capturing the integration of ethics into AI and algorithms.¹²⁶ An AI system that includes ethical rules in its software will probably reduce the risk of an act of aggression that may be resulted from a misuse of a fully-autonomous weapon. The proposed system can be compared to software that prevents allied forces from accidentally hitting each other's military units during an armed conflict called Identification Friend or Foe (IFF). It is already possible to distinguish one's own units from enemy's with current defense technology with the help of IFF. A similar application can ensure the systems that will work with AI comply with ethical rules. Norms and principles of LOAC such as distinction can be coded and taught in this system. The possibility of weapon systems using AI software to act against non-combatants during a military operation can be minimized to the extent possible through the instrument of ethical improvement. A C-DAEM unit described above coded in accordance with ethical rules will find targets itself and it will boast algorithms able to separate between friends and foes in complex combat zones that could contain civilians alongside military troops.¹²⁷ In this context, the most important and difficult process in determining ethical codes will probably be to bring states together convincing them the necessity of working together – for instance, to form an international commission. In this process, international governing bodies like United Nations, NATO, International Monetary Fund and/or the World Bank can work on identifying the potential consequences of autonomous lethal use of force and creating international norms in this field.¹²⁸ The rules that would emerge as a result of such an effort could even be transformed into a set of standard codes that would enable any military AI system of lethal capacity to learn and comply.

The third measure concerns the creation of a specific legal framework for weapon systems using AI. As technology becomes more widespread and warfare becomes increasingly possible even in space, it may not be practical to apply only traditional legal principles like the ones in LOAC. Therefore, it may be appropriate to make some new arrangements for the responsibility of AI.¹²⁹ In this process, it will be an important development to create a moral code for AI manufacturers and enable AI programmers to create AI-based systems in accordance with these accepted codes and requirements.¹³⁰ Thus, the AI programmer will have to comply with certain standards when designing the system, and will be responsible for non-standard design and production. Arrangements can be made on the basis of compliance with ethical values and the privacy principle.¹³¹ It is also clear that some technical standards need to be established for safe, trustworthy, secure development, testing and implementation of AI technologies.¹³² Thanks to these standards, any system performing with the help of AI software can be made more secure. Standards covering all institutions working on AI research, invention, innovation, dissemination and application should be regulated by responsible institutions.¹³³ It is beneficial for these standards to be based on issues such as making risk assessment on the positive and negative aspects of AI implementation, ensuring human control, human safety and wellbeing, determining obligations, and establishing a system that is compatible with human values.¹³⁴ Software programmer's standards regarding act of aggression can be regulated based on the rules of law, especially LOAC principles. In Vivek's words, 'maintaining the AI arms race without regulatory rules is a bad idea', and it is useful for all nations developing AI and autonomous weapon systems to link the use of such weapons to a legal framework.¹³⁵ United States, United Kingdom, China, the European Union, and some private entities have been able to create certain legal frameworks and declaratory documents, yet, their efforts in this regard has been described as mild or primary.¹³⁶

Finally, some mechanisms can be developed in lethal weapon systems that can disable AI activity completely when it is needed. It is important and necessary for researchers to develop such a mechanism that can shut down the whole system to prevent an accidental and/or unauthorized attack. A fail-safe software such as the applications that stop the entire operation in nuclear power plants can be installed in fully-autonomous weapons working with machine learning, so that the system can be entirely turned off if it is out of control in some way. Considering today's technological achievements, to

¹²⁴ Patrick van Esch and others, 'Autonomous Weapon Systems: Is A Space Warfare Manual Required?' (2017) 33 (3) Computer Law & Security Review 382.

¹²⁵ Peter Singer 'Ethics', (*Encyclopædia Britannica* January 2020) <<https://www.britannica.com/topic/ethics-philosophy>> accessed 28 March 2020.

¹²⁶ Julia M. Ptaschunder, 'Artificial Intelligence Evolution: On The Virtue Of Killing In The Artificial Age' (2018) 4 *Scientia Moralitas-International Journal of Multidisciplinary Research* 51.

¹²⁷ Luke Dormehl, 'U.S. Army's A.I. Missiles Will Not Stop, Ever, Until They Hit Their Target' (*Digital Trends*, 15 August 2019) <<https://www.digitaltrends.com/cool-tech/ai-missile-c-daem-army/>> accessed 28 May 2020.

¹²⁸ Spinelli, (n 109).

¹²⁹ Van Esch and others, (n 125).

¹³⁰ Haenlein and Kaplan, (n 6).

¹³¹ Rachel L. Finn and David Wright, 'Unmanned Aircraft Systems: Surveillance, Ethics And Privacy In Civil Applications' (2012) 28 (2) *Computer Law & Security Review* 184, 191.

¹³² The-White-House, 'American Artificial Intelligence Initiative: Year One Annual Report' 36 <<https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>>. accessed 13 February 2020, 15.

¹³³ Roger Clarke, 'Principles And Business Processes For Responsible AI' (2019) 35 (4) *Computer Law & Security Review* 410, 416.

¹³⁴ *ibid* 416.

¹³⁵ Sehrawat, (n 64).

¹³⁶ Tzimas, (n 115), 78.

create such an advanced fail-safe software may not be an impossible task neither for defense industry nor private companies in information sector.

7. Conclusion

The use of AI in fully-autonomous weapon systems performing under complex processes such as machine learning is turning into a global phenomenon. With the advent of quantum computers, advances in big data, and the developments of computer technology in general, AILWS are expected to become more common especially in the countries with available technology. Since many systems are at early research phase today, it would be unrealistic to mention the existence of a human-like AI, hence, any machine learning-based autonomous system seems to remain a human-dependent technology in the foreseeable future.

Any kind of fully-autonomous weapon system that is completely closed to human intervention would/should not mean that human responsibility will be ceased, if such a system harms a human being or property. According to the Rome Statute Article 8 bis, it is clear that an AI system examined in this study – regardless of its intelligence level and/or development stage – can not be held responsible *per se* for an act of aggression. Weapon systems with highly developed AI softwares do not have a criminal capacity in the context of international law, in particular the Rome Statute. Also, AI systems can not intentionally commit crimes, since they do not have the mental ability will which means that the perpetrator of a possible act/crime of aggression to be held responsible will be the producer or user.

Given the harm an AILWS may cause to people and property, it is beyond doubt that comprehensive measures must be taken to prevent the act of aggression which will be caused by an autonomous weapon. First of all, governments should regulate the use of smart arms by applying the existing international law. Second, all AI systems should be taught legal and ethical rules in accordance with LOAC principles and peremptory norms on the use of force in international law. Third, there should be a legal framework for weapon systems based on AI including a moral code for manufacturers. Finally, it may be a necessary precaution to develop fail-safe mechanisms that can render AI systems dysfunctional when needed. If all these measures were elaborated by governments, international organizations, and other responsible non-state actors – at least in the form of a soft law on machine learning and autonomous weapons – it would be possible to minimize the risks caused by military AI systems in complex combat zones where high-technology seems will be used more and more in near future.

Declaration of Competing Interest

None.

REFERENCES

- Abbasi, S. (2018). 'A conceptual incongruence between international laws of self-defense and the international core crime of aggression', 6 (1), PENN. ST. J.L. & INT'L AFF 179.
- Akande D, Tzanakopoulos A. The international court of justice and the concept of aggression. In: Kreß Claus, Barriga Stefan, editors. *The Crime of Aggression: a Commentary*. Cambridge University Press; 2016.
- Allen G, Chan T. *Artificial Intelligence And National Security*. first ed. Cambridge, MA: Belfer Center for Science and International Affairs; 2017.
- Anyoha, R. (2020). 'The history of artificial intelligence – science in the news' (Science in the News, 28 August 2017) <<http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>>accessed 27 Feb
- Arkin RC. *Governing Lethal Behavior In Autonomous Robots*. first ed. Taylor & Francis Group; 2009.
- Asaro MP. How just could a robot war be? *Front. Artif. Intell. Appl.* 2008;175:50.
- Ayodele TO. Types of machine learning algorithms. In: Zhang Yagang, editor. *New Advances in Machine Learning*. IntechOpen; 2010.
- Baker EJ. National security law & emerging technologies: toward a decisional framework-key takeaways from the aba-osu symposium and Jirga. *ISJLP* 2019;15:65.
- Bathae Y. The artificial intelligence black box and the failure of intent and causation. *Harv. J. Law Technol.* 2018;31(2):879 *Harvard Journal of Law & Technology*.
- Bostrom, N. (1998). 'How long before superintelligence?' (Nickbostrom, 25 October 1998) <<https://nickbostrom.com/superintelligence.html>>accessed 28 May
- Chui, M. (2017). 'Artificial intelligence the next digital frontier?.' <<https://www.mckinsey.com/~media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx>>accessed 28 May (2020).
- Clarke R. Principles and business processes for responsible AI. *Comput. Law Secur. Rev.* 2019;35(4):410.
- Clarke R. Regulatory alternatives for AI. *Comput. Law Secur. Rev.* 2019;35(4):398.
- Clarke R. Why the world wants controls over artificial intelligence. *Comput. Law Secur. Rev.* 2019;35(4):423.
- Council of Europe, 'History of artificial intelligence' (Artificial Intelligence, (2020)). <<https://www.coe.int/en/web/artificial-intelligence/history-of-ai>>accessed 27 May 2020.
- Corten O. The Controversies Over The Customary Prohibition On The Use Of Force: a Methodological Debate. *European Journal of International Law* 2005;16(5):803.
- "Côte D'ivoire' Coalition for the international criminal court' (Coalition for The ICC, (2020)). <<http://www.coalitionfortheicc.org/country/cote-divoire>>accessed 28 May (2020).
- Cynthia R. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* 2019;1:206.
- Čerka P, Grigienė J, Širbikytė G. Is It Possible to grant legal personality to artificial intelligence software systems? *Comput. Law Secur. Rev.* 2017;33(5):685.

- Dialani, P. (2019))'AI missiles will be developed by the US army to locate their own targets' (Analytics Insight, <https://www.analyticsinsight.net/ai-missiles-will-be-developed-by-the-us-army-to-locate-their-own-targets/> accessed 10 May 2020
- Dignum, V. (2017). 'Responsible artificial intelligence: designing AI for human values' [2017]ICT Discoveries 1 <<http://hdl.handle.net/20.500.11948/2177https://www.itu.int/en/journal/001/Pages/default.aspx>> accessed 28 May
- Docherty BN. *Mind the gap: the lack of accountability for killer robots*. Hum. Rights Watch 2015.
- Docherty LB. *Shaking the foundations: the human rights implications of killer robots*. Hum. Rights Watch 2014.
- Dormehl, L. (2020). 'U.S. army's A.I. missiles will not stop, ever, until they hit their target' (Digital Trends, 15 August 2019) <<https://www.digitaltrends.com/cool-tech/ai-missile-c-daem-army/>> accessed 28 May
- Finn LR, Wright D. *Unmanned aircraft systems: surveillance, ethics and privacy in civil applications*. Comput. Law Secur. Rev. 2012;28(2):184.
- Geneva Academy, 'Autonomous weapon systems under international humanitarian and human rights law - the geneva academy of international humanitarian law and human rights' (Geneva-academy, 2014) <<https://geneva-academy.ch/research/our-project/detail/9-autonomous-weapon-systems-under-international-humanitarian-and-human-rights-law>> accessed 29 May (2020).
- Gillett M. *The anatomy of an international crime: aggression at the international criminal court*. Int. Criminal Law Rev. 2013;13(4):829.
- Glennon JM. *The blank-prose crime of aggression*. Yale J. Int'l L. 2010;35(1):71.
- Grudin J, Jacques R. *Chatbots, humbots, and the quest for artificial general intelligence*. 19th. Conference on Human Factors in Computing Systems Glasgow May, 2019.
- Gurkaynak G, Yilmaz I, Haksever G. *Stifling artificial intelligence: human perils*. Comput. Law Secur. Rev. 2016;32(5):749.
- Haenlein M, Kaplan A. *A brief history of artificial intelligence: on the past, present, and future of artificial intelligence*. Calif. Manag. Rev. 2019;61(4):5.
- IBM, 'What is deep learning, and why does it matter?' (Ibm, 2020) <<https://www.ibm.com/cloud/deep-learning>> accessed 28 May (2020).
- International Committee of the Red Cross, 'Ethics and autonomous weapon systems: an ethical basis for human control?' (International Committee of the Red Cross, 2018) <<https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>> accessed 28 May (2020).
- John McCurley, 'Is "Mistake of Fact" a defense to criminal charges?' (Nolo, 2020) <<https://www.nolo.com/legal-encyclopedia/is-mistake-fact-defense-criminal-charges.html>> accessed 28 May (2020).
- Kaplan A, Haenlein M, 'Siri Siri. In my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus. Horiz. 2019;62(1):15.
- Legal Information Institute, 'International criminal law' (Cornell Law School LII, 2020) <https://www.law.cornell.edu/wex/international_criminal_law> accessed 28 May (2020).
- McCarthy J. *A proposal for the dartmouth summer research project on artificial intelligence*. AI Mag. 2006;27(4):12.
- McKendrick, K. (2018). 'Banning autonomous weapons is not the answer' (Chatham House, 2018) <<https://www.chathamhouse.org/expert/comment/banning-autonomous-weapons-not-answer>> accessed 28 May
- Menashe M, Kikarea E. *The global governance of cyberspace: reimagining private actors' accountability: introduction*. Camb. Int. Law J. 2019;8(2):153.
- Milojević M. *Prohibition of use of force and threats in international relations*. Law Polit. 2001;1(5):581.
- National Science and Technology Council: Committee on Technology, 'Preparing for the future of artificial intelligence' (Obamawhitehouse.archives.gov, 2020) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 28 May (2020).
- NATO Cooperative Cyber Defence Centre of Excellence 'Tallinn manual on the international law applicable to cyber warfare' (Cambridge University Press (2013)).
- Noone GP, Noone DC. *The debate over autonomous weapons systems*. J. Int'l L. 2015;25(1/6):25.
- Noyes, K. (2016). '5 things you need to know about a.i.: cognitive, neural and deep, oh my!' (Computerworld, 3 March 2016) <<https://www.computerworld.com/article/3040563/5-things-you-need-to-know-about-ai-cognitive-neural-and-deep-oh-my.html>> accessed 27 May
- Organisation for Economic Co-operation and Development (OECD) 'Artificial Intelligence in Society', (Paris, 2019) OECD Publishing, Paris, <https://doi.org/10.1787/eedfee77-en>. <<https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm>> accessed 28 March (2020).
- Paulus LA. *Legalist groundwork for the international criminal court: commentaries on the statute of the international criminal court*. Eur. J. Int. Law 2003;14(4):843.
- Petman JM. 'Autonomous weapons systems and international humanitarian law: 'out of the loop?'. University of Helsinki The Eric Castren Institute of International Law and Human Rights Research Reports, 2017.
- Press, G. (2020). 'A very short history of artificial intelligence (AI)' (Forbes, 30 December 2016) <<https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/>> accessed 02 Feb
- Puaschunder MJ. *Artificial intelligence evolution: on the virtue of killing in the artificial age*. Sci. Moralitas-Int. J. Multidiscip. Res. 2018;4:51.
- Raso, F. (2018).and others, 'Artificial intelligence & human rights: opportunities & risks' The Berkman Klein Center for Internet & Society Research Publication Series, Research Publication No.2018/6.
- Rudischhauser, W. (2017). 'Autonomous or semi-autonomous weapons systems a potential new threat of terrorism?' Federal Academy for Security Policy Working Paper, No. 23/2017 <https://www.baks.bund.de/sites/baks010/files/working_paper_2017_23.pdf> accessed 02 May 2020.
- Samek, W., Wiegand, T., & Müller, K.R. (2017). 'Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models.' (Cornell University arXiv, Aug 2017) <<https://arxiv.org/abs/1708.08296>> accessed 19 March
- SAS, 'Artificial intelligence – what it is and why it matters' (SAS Insights, (2020)). <https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html> accessed 28 May (2020).
- Sassòli M. *Autonomous weapons and international humanitarian law: advantages, open technical questions and legal issues to be clarified*. Int. Law Stud./Nav. War Coll. 2014;90:308.
- Sayler, K.M. (2019))'Defense primer: U.S. policy on lethal autonomous weapon systems [March 27, 2019]' (Homeland

- Security Digital Library,
<<https://www.hsdl.org/?abstract&did=823590>>accessed 28 May (2020).
- Scharre PD. *The opportunity and challenge of autonomous system autonomous systems: issues for defence policymakers*. In: Williams Andrew P, Scharre Paul D, editors. *NATO Communications and Information Agency. NATO Communicatin and Information Agency*; 2015.
- Schmitt, E. (2017))'Pentagon tests lasers and nets to combat a vexing foe: ISIS drones' (Nytimes,.
<<https://www.nytimes.com/2017/09/23/world/middleeast/isis-drones-pentagon-experiments.html>>accessed 28 May (2020).
- Sehrawat V. *Autonomous weapon system: law of armed conflict (LOAC) and other legal challenges*. *Comput. Law Secur. Rev.* 2017;33(1):38.
- Shalev-Shwartz S, Ben-David S. *Understanding Machine Learning*. Cambridge University Press; 2017.
- Singer, P. (2020). 'Ethics', (Encyclopædia Britannica January 2020)
<<https://www.britannica.com/topic/ethics-philosophy>> accessed 28 March 28
- Spinelli, C.J. (2015))'The rise of robots: the military's use of autonomous lethal force' (unpublished thesis, Air War College, Air University.
- Szpak A. *Legality of use and challenges of new technologies in warfare – the use of autonomous weapons in contemporary or future wars*. *European Review* 2019;28(1):118.
- The Avalon Project, 'Judgment: the accused organizations' (Avalon Yale Law School Lillian Goldman Law Library, 2020)
<<https://avalon.law.yale.edu/imt/judorg.asp>>accessed 28 May (2020).
- The Royal Society, 'Machine learning the power and promise of computers that learn by example' (Royal Society, 2020)
<<https://royalsociety.org/-/media/policy/projects/machine-learning/publications/machine-learning-report.pdf>>accessed 28 May (2020).
- The-White-House, 'American artificial intelligence initiative: year one annual report' 36 <<https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>>. accessed 13 February (2020).
- Thurnher SJ. *Examining autonomous weapon systems from a law of armed conflict perspective*. In: Nasu Hitoshi, McLaughlin Robert, editors. *New Technologies and the Law of Armed Conflict*. TMC Asser Press; 2014.
- Turing AM. I.—Computing machinery and intelligence. *Mind* 1950;59:433.
- Tzimas T. *The need for an international treaty for AI from the perspective of human rights*. *Sci. Moralitas J.* 2019;4:73.
- Van Esch P. *Autonomous weapon systems: is a space warfare manual required?* *Comput. Law Secur. Rev.* 2017;33(3):382.
- Walch, K. (2019). 'Rethinking weak vs. strong AI' (Forbes, 4 October 2019) <<https://www.forbes.com/sites/cognitiveworld/2019/10/04/rethinking-weak-vs-strong-ai/>>accessed 27 May
- Zech H. *Liability for autonomous systems: tackling specific risks of modern IT*. In: Lohsse Sebastian, Schulze Reiner, Staudenmayer Dirk, editors. *Liability for Robotics and in the Internet of Things*. Nomos Hart; 2018 (forthcoming)'
<<https://ssrn.com/abstract=3195676> or <http://dx.doi.org/10.2139/ssrn.3195676>>accessed 14 May 2020.
- Zilberstein S. *Building strong semi-autonomous systems*. Twenty-Ninth AAAI Conference On Artificial Intelligence, 2015.