The commission of crimes using autonomous weapon systems: Issues of causation and attribution

Vasilios Tzoufis*1, Nikolaos Petropoulos1

Abstract - Today’s military takes advantage of technological progress primarily through developments in the war industry. Of particular importance is the development and use of advanced autonomous weapon systems, which are changing the way conflicts are fought and redefining the basic strategies on the modern battlefield. The use of these systems raises ethical, philosophical, and legal questions about the relationship of causality and attribution that must exist between an act and its criminal result.

Die Begehung von Straftaten durch autonome Waffensysteme: Fragen der Kausalität und Zurechnung

Zusammenfassung - Das moderne Militär macht sich den technologischen Fortschritt in erster Linie über Entwicklungen in der Rüstungsindustrie zunutze. Von besonderer Bedeutung sind dabei die Entwicklung und der Einsatz fortschrittlicher autonomer Waffensysteme, die die Art und Weise der Konfliktführung verändern und die grundlegenden Strategien auf modernen Schlachtfeldern neu definieren. Der Einsatz dieser Systeme wirft jedoch ethische, philosophische und rechtliche Fragen zum Verhältnis von Kausalität und Zurechnung auf, das zwischen einer Handlung und den strafrechtlichen Folgen bestehen muss.

Keywords - autonomous weapons systems, attribution, causality, criminal law, international humanitarian law

Introduction

One of the basic features of the functioning of the free economy is the production and distribution of as many goods (consumer products) as possible, which is easily deduced from our social experience. Rapid technological development maximizes the potential for increasing productivity through automation. This process and the need for reducing – if not eliminating – risks has led to the research and application of robotics and artificial intelligence systems, whose use in the production process in everyday life raises ethical, philosophical, and legal dilemmas. The use of the most modern applications of technological development was bound to involve the war industry, a branch of industry of vital importance not only for the economy and the balance of trade of many states but also for the defense and the security of states. Modern armies broadly use the exploitation of technology by the war industry: of particular importance is the production and use of advanced autonomous weapon systems (AWS), which are changing the way conflicts are conducted and redefining the basic concepts of strategy in the modern battlefield. With the use of these systems, the human factor takes a back seat, as they are characterized by a greater or lesser degree of autonomy, i.e., the ability to respond to uncertain situations by independently composing and selecting among different courses of action to accomplish goals based on knowledge and a contextual understanding of the world, itself, and the situation (STO 2020, p. 16). This reality poses significant challenges to the violations of international humanitarian law (IHL), given that the affirmation of causality and imputation according to traditional criminal theory becomes more complex than the corresponding vio-
lations that occur under conventional weapon systems. For instance, what happens when a system that guides a weapon or that makes a decision on the choice of target or the circumstances of engaging the target violates IHL, for example, by not distinguishing between civilians and combatants or striking the target without having taken precautions to protect civilians or even violates the principle of proportionality, because the expected harm to civilians, or civilian property, is excessive in relation to the expected military advantage? Who will be punished as an IHL offender in this case? What will also happen if an autonomous weapon system already has the specifications from its manufacturer to hit targets in violation of IHL? Can the military user or the manufacturer be punished as the perpetrator? (Anderson and Waxman 2013, pp. 16–17).

**Autonomous weapons systems and the risks of violating international humanitarian law**

It is commonplace that there is a need for a specific institutional framework for their use, given that the use of AWS has become widespread in recent years: they have been systematically used in warfare in Iraq and Afghanistan (Wall and Monahan 2011, p. 240), as well as in the fight against international terrorism (Ambos and Alkatout 2012, p. 342). These weapons were used for targeted operations and to support conventional battle tactics. However, the war in Ukraine gave a concrete example of the battlefield in the 21st century: widespread use of AWS, attacks against the military (and unfortunately civilian) targets exclusively using drones, precision weapons with built-in cameras, transmitting images to the military user in real-time, more effective target reconnaissance, precision strikes deep into enemy territory, etc. (Thompson 2024). Indeed, it can be argued that “more than any conflict in human history, the fighting in Ukraine is a war of drones” (Mozur and Hopkins 2023).

The use of these systems has already raised several scientific and ethical/philosophical issues already plaguing the literature. The issue concerns states, international organizations, and military alliances, and various attempts have already been made to define the essential characteristics of these systems. Despite the diversity of definitions, which have been given either within the framework of the NATO Alliance or by individual countries (STO 2020, p. 16; Department of Defense 2023, pp. 15–18; Defence Ethics Committee 2021, pp. 3–5; People’s Republic of China 2022, ICRC 2014, pp. 61–64; 2021, pp. 5–6) one can identify the essential characteristics of these systems, which to critical functions such as identification, classification, interception and engagement within time and space limits and under conditions, [b)] which include technical safeguards or intrinsic characteristics to prevent failures, misuse and relinquishment by the command of two essential duties, namely situation assessment and reporting. […] In other words, a PALWS is a lethal weapon system whose decision-making functions are defined according to the robotics meaning of decision-making autonomy, i.e., within a specific framework of action” (Defence Ethics Committee 2021, p. 4). At the same time: “A PALWS cannot take lethal initiatives that would result in it altering its functional scope” (Defence Ethics Committee 2021, p. 4).

In any war conflict, the fundamental humanitarian problem is the risk of violation of basic principles of the IHL, such as...
the principle of distinction between civilians and combatants, as enshrined in Additional Protocol I (UN 1977) to the Geneva Convention, Articles 48, 51 and 52, according to which attacks against civilians and civilian objects are prohibited in the conduct of combat operations, and such as the principle of proportionality, as enshrined in Additional Protocol I (UN 1977) to the Geneva Convention, Articles 51 (5) (b), 57 (2) (a) (iii), and 57 (2), and 52, which prohibits attacks which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive about the concrete and direct military advantage anticipated.

The problem of the use of autonomous weapons systems is compounded by the fact that they are now available to non-state entities.

Therefore, in line with what has been mentioned above, attribution of liability for the violation of the above principles of IHL does not encounter particular problems in cases of the use of partially autonomous lethal weapon systems because the relevant issues of causation and attribution of liability in cases of violation of IHL through the use of these systems (PALWS) will be treated in the same way as the violation through conventional weapon systems due to the active role in a more significant or to a lesser extent of the human factor. However, as we will see below, causation and attribution of an act to the attribution of the perpetrator require a radical change in the interpretative tools of traditional criminal theory when the violation is caused by a lethal weapon system (LAWS), the autonomy of which precludes initial or subsequent user intervention. How will responsibility be attributed to a military man who detonates an autonomous weapon system, which the manufacturer already programs not to distinguish between civilians and combatants (impossibility of human intervention even at the initial stage)? How can responsibility be attributed to a military man who fires a weapon system in order to hit a military target but then finds out that civilians are going to be hit, which he could not foresee beforehand, while the autonomous weapon system does not allow him to intervene afterward, e.g., to modify its course, etc. (impossibility of human intervention in the next – after detonation – stage)?

Despite the optimistic view that has been expressed about their use, according to which a modern ‘smart’ weapon system can protect civilians or limit their losses and therefore, army robots may be more humane on the battlefield than humans (Arkin 2010, pp. 334–335), the sad reality of today’s wars confirms the opposite, as in all wars civilian casualties are significant. Therefore, it is perfectly justified to be concerned about the continuous production of weapons capable of killing or injuring civilians en masse and to accept a dystopian approach to the near future with armed robots running amok (Guha and Galliott 2023, p. 64) against civilians.

On the other hand, the use of autonomous weapon systems has made it possible to conduct military operations from afar without the local presence of military personnel involved in the operation, making it challenging to investigate responsibilities in cases of violation of the IHL during the execution of a combat operation. Are states or their secret services exploiting the potential of this technology; it is now a fact that several military operations of states or their secret services use the technology of modern autonomous weapon systems (AWS) to carry out “targeted killings” (Wall and Monahan 2011, p. 242), or “preventive wars” (Fabiano 2023, pp. 214–215), a practice that may involve people who do not have combatant status, but even suspects (Dombrowsky 2013, p. 236) and is therefore criticized for its opposition to IHL (Ambos and Alkatout 2012, p. 359). Not to mention that the Israeli High Court of Justice considers this practice justified only under stringent conditions (Rabi and Plaw 2020, pp. 227–231).

It should also be pointed out that the problem of the use of autonomous weapons systems is compounded by the fact that they are now available to non-state entities, where prosecution is highly doubtful in the absence of judicial institutions. According to the United Nations, a non-state actor (NSA) is defined as a group of people that has “the potential to employ arms in the use of force to achieve political, ideological or economic objectives; [but] are not within the military structures of states, state-alliances or intergovernmental organizations; and are not under the control of the state(s) in which they operate” (Rutherford 2015, pp. 76–77).

Although NSAs include different categories of armed groups, terrorist groups are perceived as the most notorious and dangerous to both states and civilians, given the widespread use of weapons by them in the international environment. What is particularly problematic about the NSAs activities is an apparent lack of enforcement in the current global judicial system to apply international humanitarian law to NSAs (Meron 2011, pp. 32–35). In other words, although in the context of international law, the International Criminal Court is perceived as the only competent judicial authority that could prosecute perpetrators of war crimes, including NSAs. However, the existing mechanisms and procedures do not appear effective in formulating the execution of the court in this context – i.e. in the case of NSAs (Olásolo 2005, pp. 101–103). Of course, states do enjoy the right to bring
perpetrators to justice in their territory when NSAs violate the respective provisions of national law (Porter 2010, p. 198).

Over the last decade or so, several non-state actors across the globe have incorporated drones into their operations with at least five of them operating in the Middle East – namely, Hezbollah, Hamas, the Houthi Movement, Islamic State, and the Kurdish Workers’ Party (Veilleux-Lepage and Archambault 2022, pp. 45–52). Armed unmanned aerial vehicles have also been used in the conflict zones of Ukraine, Myanmar, Mexico, and Ecuador, as well as in central Asia (Haugstved 2023, p. 32).

While most rebel groups from South America to the Middle East have extensively used commercially available and low-cost drones purely for surveillance and recognizance purposes, Islamic State of Iraq and Syria became the first terrorist group to use drones for carrying out attacks, demonstrating a more tech-savvy approach in modern warfare (Almohammad and Speckhard 2017, pp. 45–50). More recently, Houthi rebels in Yemen have used drones to attack Saudi oil facilities, demonstrating a high degree of precision in their strikes (Haugstved 2023, pp. 33–35). In early December 2023, the rebels threatened to disrupt the global oil supply by deploying drone swarms in certain parts of the Red Sea, targeting oil tankers crossing toward the Suez Canal.

What makes the use of drones by NSAs stand out is the fact that, unlike other forms of military power, air power has traditionally been associated with statehood and sovereignty; the use of drones by non-state actors, therefore, represents an incursion by non-state groups into the prerogative of sovereign states – not only militarily, but also conceptually and symbolically (Archambault and Veilleux-Lepage 2020, p. 78).

A worrisome development related to the use of drones by non-state actors is the fact that while high-end technology was once an exclusive privilege of states with robust military budgets, over time, access to advanced technology has been made more accessible for individuals and non-state actors who can today acquire affordable commercial-off-the-shelf technology to be integrated into attack planning.

**Autonomous weapons systems and criminal liability (attribution and causation)**

An indispensable element of the concept of crime according to criminal theory, particularly in the context of continental law (Germany, France, Greece, etc.), is the principle of attribution (nullum crimen, nulla poena sine culpa) as a partial principle of the principle of legality of crimes and penalties (nullum crimen, nulla poena sine lege), which also highlights the anthropocentric-personal character of criminal law as the law of the personal-individual responsibility of the specific offender. According to Kantian philosophy, respect for the greatness and value of a man is the basic content of the ethical theory, and man, as a mental character (intelligible character), is self-determined by deciding freely. The principle of attribution is enshrined in the constitutions of most countries. In contrast, in the institutional framework of the Council of Europe, the principle of attribution is derived from Article 7 of the European Convention on Human Rights. In the institutional framework of the EU, the principle of attribution is derived from Article 48 (presumption of innocence) of the Charter of Fundamental Rights of the European Union in conjunction with Article 1 (human dignity).

However, in the case of violation of the IHL through the use of AWS, it becomes challenging to attribute criminal responsibility to the operator-military who operates an autonomous weapon system since, as mentioned above, a vital element of the operation of these systems is the autonomy of ‘decision-making’, i.e., the ability of these systems to select and strike targets independently of human intervention. Therefore, it is inconceivable to attribute criminal liability to the military operator, who operates an autonomous system but then has no possibility of intervening in it in case of violation of the principles of IHL by the system due to the ‘error’ of the system, since, even if the operator detects in time the possibility of violation of the system, there is nothing he can do to prevent it. On the other hand, in such cases, any criminal liability of the manufacturer of AWS must be examined, which may take the following forms: a) if he acted intentionally, i.e., programmed the weapon system precisely to violate the IHL, e.g., to target civilians, etc., then his liability is apparent, b) if he acted negligently, then it should be judged whether the technical/manufacturing protocols used to build the AWS were safe and appropriate to prevent as far as possible any violations of the IHL, and if they were secure, it should be investigated whether it complied with these specifications.

Given the above, it becomes clear that criminal law does not only need a simple adaptation of its structures to the constantly changing social conditions, but for the first time after about two centuries, it is faced with the question of revising its basic fundamental principles, so that it can respond to its mission, in this case the protection of civilians from violations of IHL. In this context, the scientific debate has opened on whether or not to recognize a quasi-legal personality in robots (Chesterman 2020, pp. 820–822), but the scientific community appears divided; some are completely opposed to considering the capacity for imputation as an element of the person, who develops a moral and emotional relationship with himself and his actions in the present and past (Gless 2017, p. 326), something that robots can never acquire, but others are open to a new approach to criminal theory and a revision of the basic concepts of criminal law, such as culpability (Simmler and Markwalder 2017, pp. 37–47; Dremluiga and Prisekina 2020, p. 260), albeit with conditions (Gleß and Weigend 2015, pp. 576–577).

In addition, using AWS also raises issues related to causality, which must exist between an act and a criminal result. Thus, in the context of the violation of IHL and the infliction of harm on civilians through the possibility of committing criminal acts provided by modern technology through AWS, signifi-
significant problems may arise in establishing the objective causal link between an act and a harmful effect. The use of autonomous weapon systems, especially the most advanced ones, often requires the cooperation of many people: There are drones controlled by the operator but assisted by another monitoring team. These drones are driven from the battlefield or a station, which can also be another drone. There are also drones with more excellent autonomy capability. In this case, this drone has been programmed in advance, and the operator can intervene – to a greater or lesser extent – to correct the errors. Technological development leads to constructing systems programmed with full autonomy (Balazünbüll 2021, pp. 17–22). In the event of a breach of the IHL in such cases, known as cumulative damage cases (‘Summationsschäden’), because the conduct of several parties causes the damage, it becomes difficult to ascertain the percentage of harm caused by each party, such as the instances of using AWS described above, in which the attack on the target requires the action and/or cooperation of several persons, from the manufacturer, the programmer to the supervisor and the end user. As can be deduced from the above, the rapid development of new technologies in the field of the war industry, with the help of artificial intelligence applications, will sooner or later lead to the use of systems with a full degree of autonomy. The modern economic organization has highlighted large corporations as the main pillars of production. Therefore, the war industry, by developing autonomous applications of weapon systems, automatically becomes morally and socially responsible for the risks arising from the use of these technologies. It therefore becomes clear that proving criminal responsibility for the violation of IHL through the use of AWS is primarily a difficult task, because we should check each time the lesser or more significant possibility of user/military intervention during the use of an autonomous weapon system. Moreover, in a few years, applications of artificial intelligence. For this reason, it is also necessary for the states to adopt standard uniform rules (protocols) for the operation of factories and companies involved in the manufacture of automatic weapons systems.

It is necessary for states to adopt standard uniform rules (protocols) for the operation of factories and companies involved in the manufacture of automatic weapons systems.

Conclusion

From what has been set out above, the need to respect international humanitarian law has emerged from the severe risks faced by civilians due to the use of modern weapon systems. For such protection to be adequate, it is necessary to adopt initiatives initially at the international level. Of utmost importance is the negotiation between states to adopt international conventions that will recognize a commonly accepted definition of AWS and will impose strict restrictions on the use of AWS in military operations, in the sense that their use will be delimited by the principles of the IHL, notably those of proportionality and discrimination between civilians and combatants. Prima facie, the most effective way to prevent violations of the IHL in using autonomous weapons systems in warfare is to ban them. Apart from the fact that the adoption of similar initiatives at international level would be extremely difficult for reasons primarily political, which have to do with the global power competition between states (Anderson and Waxman 2013, p. 21), such a prohibition would not be relevant from a practical point of view, since the ban on autonomous weapons systems with specific technical characteristics at the present juncture would not legally cover the rapid development of those systems, while a complete ban on any development of autonomous weapons systems would exclude any possibility of manufacturing in the future weapons systems with ethical and legal specifications by the IHL (Anderson and Waxman 2013, pp. 20–22).

However, as mentioned above, it is a given that the development of modern weapon systems will be rapid shortly due to the control of human intervention will be irrelevant, since the use of fully autonomous systems is certain. Therefore, it would be more effective to identify specific protocols for the manufacture of these autonomous weapon systems, which would be compatible with IHL and which would become legally binding and their violation (e.g., intentional or negligent programming a system that attacks combatants and civilians indiscriminately) will automatically prove a violation of IHL. Similar legally binding protocols of action should also be adopted in the framework of international conventions to adopt arrangements for the use of autonomous weapons systems under humanitar-
IHL violations; the U.S. Agency for International Development announced the delivery of nine autonomous drones to the Office of the Prosecutor General of Ukraine to collect material on possible violations of the IHL by the Armed Forces of the Russian Federation (USAID 2023). In addition, it can be agreed in interstate agreements or agreements between international organizations to adopt minimum ethical values as standards for the construction of AWS (Trusilo 2023 pp. 12–13), such as the development of algorithms capable of superior target discrimination capabilities (combatant – noncombatant status) (Arkin 2010, pp. 340–341), but also the monitoring of states’ compliance with the rules of use of AWS through international organizations or committees (such as the International Humanitarian Fact-Finding Commission; Steinberg and Herzberg 2018, pp. 298–299).

Funding • This work received no external funding.

Competing interests • The authors declare no competing interests.

References


Dr. Vasilios Tzoufis is a lawyer (Athens Bar Association) at the Supreme Court of Greece ("Aresos Pagos") and a postdoctoral researcher at Aristotle University of Thessaloniki (Faculty of Law). He has taught at schools of the Greek army and the Greek security forces.

Dr. Nikolaos Petropoulos is a criminologist with a Ph. D. in criminal justice from the City University of New York. He is also a member of the EENet Steering Committee.