INTERNATIONAL HUMANITARIAN LAW AND THE CHALLENGES OF CONTEMPORARY ARMED CONFLICTS

RECOMMITTING TO PROTECTION IN ARMED CONFLICT ON THE 70TH ANNIVERSARY OF THE GENEVA CONVENTIONS

REPORT

ICRC

TABLE OF CONTENTS

P. Krzysiek/ICRC

CHAPTER 2

CONTEMPORARY AND FUTURE CHALLENGES IN THE CONDUCT OF HOSTILITIES

2. NEW TECHNOLOGIES OF WARFARE

New technologies are changing human interaction profoundly – including in times of armed conflict. Many States are investing heavily in the development of means and methods of warfare that rely on digital technology. Cyber tools, increasingly autonomous weapon systems, and artificial intelligence are being used in contemporary armed conflicts. The ICRC closely follows the development of new means and methods of warfare and their use by militaries; it also engages all relevant stakeholders on the applicability of IHL to the use of these new means and methods of warfare.

Technological advances can have positive consequences for the protection of civilians in armed conflict: weapons can be used with more precision, military decisions can be better informed, and military aims can be achieved without the use of kinetic force or physical destruction. At the same time, new means of warfare and the way they are employed can pose new risks to combatants and civilians, and can challenge the interpretation and implementation of IHL. The ICRC's assessment of the foreseeable humanitarian impact of new technologies of warfare, and the challenges they may pose to existing IHL rules, focuses on interrelated legal, military, technical, ethical, and humanitarian considerations.

IHL is applicable to the development and use of new weaponry and new technological developments in warfare – whether they involve (a) cyber technology; (b) autonomous weapon systems; (c) artificial intelligence and machine learning; or (d) outer space. States that develop or acquire such weapons or means of warfare are responsible for ensuring that they can be used in compliance with IHL (e).

A) CYBER OPERATIONS, THEIR POTENTIAL HUMAN COST, AND THE PROTECTION AFFORDED BY IHL

The use of cyber operations during armed conflicts is a reality. While only a few States have publicly acknowledged using such operations, an increasing number of States are developing military cyber capabilities, and the use of such capabilities is likely to increase.

The ICRC understands "cyber warfare" to mean operations against a computer, a computer system or network, or another connected device, through a data stream, when used as means or methods of warfare in the context of an armed conflict. Cyber warfare raises questions about precisely how certain provisions of IHL apply to these operations, and whether IHL is adequate or whether, building on existing law, it might require further development.

The use of cyber operations may offer alternatives that other means or methods of warfare do not, but it also carries risks. On the one hand, cyber operations may enable militaries to achieve their objectives without harming civilians or causing permanent physical damage to civilian infrastructure. On the other hand, recent cyber operations – which have been primarily conducted outside the context of armed conflict – show that sophisticated actors have developed the capability to disrupt the provision of essential services to the civilian population.

Understanding cyber operations and their potential human cost

To develop a realistic assessment of cyber capabilities and their potential human cost in light of their technical characteristics, in November 2018 the ICRC invited experts from all parts of the world to share their knowledge about the technical possibilities, expected use, and potential effects of cyber operations.[19]

Cyber operations can pose a particular threat for certain elements of civilian infrastructure. One area of concern for the ICRC, given its mandate, is the health-care sector. In this regard, research shows that the health-care sector appears to be particularly vulnerable to direct cyber attacks and incidental harm from such attacks directed elsewhere. Its vulnerability is a consequence of increased digitization and interconnectivity in health care. For example, medical devices in hospitals are connected to the hospital network, and biomedical devices

[19] See ICRC, *The Potential Human Cost of Cyber Operations*, 2019; available at https://www.icrc.org/en/download/file/96008/the-potential-human-cost-of-cyber-operations.pdf.

such as pacemakers and insulin pumps are sometimes remotely connected through the internet. This growth of connectivity increases the sector's digital dependence and "attack surface" and leaves it exposed, especially when these developments are not matched by a corresponding improvement in cyber security.

Critical civilian infrastructure – including electrical, water, and sanitation facilities – is another area in which cyber attacks can cause significant harm to the civilian population. This infrastructure is often operated by industrial control systems. A cyber attack against an industrial control system requires specific expertise and sophistication, as well as specifically designed cyber tools. While attacks against industrial control systems have been less frequent than other types of cyber operations, their frequency is reportedly increasing, and the severity of the threat has evolved more rapidly than anticipated only a few years ago.

Beyond the vulnerability of specific sectors, there are at least three technical characteristics of cyber operations that are cause for concern.

First, cyber operations carry a risk of overreaction and escalation, simply due to the fact that it may be extremely difficult – if not impossible – for the target of a cyber attack to detect whether the attacker's aim is to spy or to cause physical damage. As the aim of a cyber operation might be identified only after the target system has been harmed, there is a risk that the target will imagine the worst-case scenario and react much more strongly than it would have done if it had known that the attacker's true intent was limited to espionage, for example.

Second, cyber tools and methods can proliferate in a unique manner, one that is difficult to control. Today, sophisticated cyber attacks are carried out only by the most advanced and best-resourced actors. But once a cyber tool has been used, stolen or leaked, or becomes available in some other way, actors other than those who developed it may be able to find it, reverse-engineer it, and repurpose it for their own – possibly malicious – ends.

Third, while it is not impossible to determine who created or launched a particular cyber attack, attributing an attack tends to be difficult. Identifying actors who violate IHL in cyberspace and holding them responsible is likely to remain challenging. The perception that it will be easier to deny responsibility for such attacks may also weaken the taboo against their use – and may make actors less scrupulous about violating international law by using them.

While cyber operations have exposed the vulnerability of essential services, they have not, fortunately, caused major human harm so far. However, much is unknown in terms of technological evolution, the capabilities and the tools developed by the most sophisticated actors, and the extent to which the increased use of cyber operations during armed conflicts might be different from the trends observed so far.

The limits that IHL sets for cyber warfare

The ICRC welcomes the fact that an increasing number of States and international organizations are acknowledging that IHL applies to cyber operations during armed conflicts. It urges all States to recognize the protection that IHL offers against the potential human cost of cyber operations. For example, belligerents must respect and protect medical facilities and personnel at all times, which means that cyber attacks against the health-care sector during armed conflict would – in most cases – violate IHL. Likewise, IHL specifically prohibits attacking, destroying, removing or rendering useless objects indispensable to the survival of the civilian population.

More generally, IHL prohibits directing cyber attacks against civilian infrastructure, as well as indiscriminate and disproportionate cyber attacks. For instance, even if the infrastructure or parts of it become military objectives (such as a discrete part of a power grid), IHL requires that only those parts be attacked, and that there be no excessive damage to the remaining civilian parts of the grid or to other civilian infrastructure relying on the electricity provided by the grid. IHL also requires parties to conflict to take all feasible precautions to avoid or at least minimize incidental harm to civilians and civilian objects when carrying out a cyber attack.

Notwithstanding the interconnectivity that characterizes cyberspace, the principles of distinction, proportionality and precautions can and must be respected. A careful examination of the way cyber tools operate shows that they are not necessarily indiscriminate. While some of the cyber tools that we know of were designed to self-propagate and indiscriminately affect widely used computer systems, they did not do these things by chance: the ability to self-propagate usually needs to be specifically included in the design of such tools. Furthermore, attacking specific targets may require custom-made cyber tools, which might make it difficult to carry out such attacks on a large scale or indiscriminately.

In fact, many of the cyber attacks that have been observed appear to have been rather discriminate from a technical perspective. This does not mean they were lawful or would have been lawful if carried out in a conflict; on the contrary, in the ICRC's view, a number of the cyber attacks that have been reported in public sources would be prohibited during armed conflict. However, their technical characteristics show that cyber operations can be very precisely designed to have an effect only on specific targets, which makes them capable of being used in compliance with IHL principles and rules.

IHL rules protecting civilian objects can, however, provide the full scope of legal protection only if States recognize that cyber operations that impair the functionality of civilian infrastructure are subject to the rules governing attacks under IHL.[20] Moreover, data have become an essential component of the digital domain and a cornerstone of life in many societies. However, different views exist on whether civilian data should be considered as civilian objects and therefore be protected under IHL principles and rules governing the conduct of hostilities. In the ICRC's view, the conclusion that deleting or tampering with essential civilian data would not be prohibited by IHL in today's ever more data-reliant world seems difficult to reconcile with the object and purpose of this body of law.[21] Put simply, the replacement of paper files and documents with digital files in the form of data should not decrease the protection that IHL affords to them.

Finally, parties to armed conflicts must take all feasible precautions to protect civilians and civilian objects under their control against the effects of attacks. This is one of the few IHL obligations that States are required to implement in peacetime.

Affirming that IHL applies to cyber warfare should not be misunderstood as encouragement to militarize cyberspace or as legitimizing cyber warfare. Any use of force by States, whether cyber or kinetic in nature, will always be governed by the UN Charter and relevant rules of customary international law. IHL affords the civilian population an additional layer of protection against the effects of hostilities.

In the coming years, the ICRC will continue to follow the evolution of cyber operations and their potential human cost, in particular during armed conflicts. It will explore avenues to reduce that cost and work towards building consensus on the interpretation of existing IHL rules and, if necessary, on the development of complementary rules that afford effective protection to civilians.

The use of digital technology during armed conflicts for purposes other than as means and methods of warfare

In recent conflicts, certain uses of digital technology other than as means and methods of warfare have led to an increase in activities that adversely affect civilian populations. For example, misinformation and disinformation campaigns, and online propaganda, have fused on social media, leading in some contexts to increased tensions and violence against and between communities. Unprecedented levels of surveillance of the civilian population have caused anxiety and increasing numbers of arrests, in some instances possibly based on disinformation. Disinformation and surveillance are not unique or new to armed conflicts; however, the greater scope and force-multiplying effect provided by digital technology can exacerbate – and add to – the existing vulnerabilities of persons affected by armed conflicts.[22] Developments in artificial intelligence

[20] See ICRC, *IHL Challenges Report 2015*, p. 41.

[21] See ICRC, *IHL Challenges Report 2015*, p. 43.

[22] See ICRC, *Digital Risks in Situations of Armed Conflict*, 2019; available at https://www.icrc.org/sites/default/files/event/file_list/icrc_symposium_on_digital_risks_-_event_report.pdf.

and machine learning are also relevant in this regard.[23] IHL does not necessarily prohibit such activities, but it does prohibit acts or threats of violence the primary purpose of which is to spread terror among the civilian population. Moreover, parties to armed conflict must not encourage violations of IHL. Other bodies of law, including international human rights law, might also be relevant when assessing surveillance and disinformation.

The global digital transformation is changing not only warfare but also the nature of humanitarian action. Digital technologies can be leveraged to support humanitarian programmes, for instance by capturing and using data to inform and adjust responses or by facilitating two-way communication between humanitarian staff and populations affected by conflicts.[24] For example, the ICRC analyses "big data" to anticipate, understand, and respond to humanitarian crises, and uses internet-based tools to interact with beneficiaries as well as with parties to armed conflicts. The ICRC also uses digital tools to restore family links and, if possible, to facilitate communication between detainees and their loved ones; the ICRC does all this also to help parties to implement their IHL obligations. These new possibilities entail new responsibilities: humanitarian organizations need to strengthen their digital literacy and data-protection measures, in accordance with the "do no harm" principle.[25] The ICRC encourages further research, discussion, and concrete steps by all revent actors to enable humanitarian actors to safely adapt their operations to digital changes.

B) AUTONOMOUS WEAPON SYSTEMS

The ICRC understands autonomous weapon systems as: *Any weapon system with autonomy in its critical functions. That is, a weapon system that can select and attack targets without human intervention*. Autonomy in critical functions – already found in some existing weapons to a limited extent, such as air defence systems, active protection systems, and some loitering weapons – is a feature that could be incorporated in any weapon system.

The most important aspect of autonomy in weapon systems – from a humanitarian, legal and ethical perspective – is that the weapon system self-initiates, or triggers, an attack in response to its environment, based on a generalized target profile. To varying degrees, the user of the weapon will know neither the specific target nor the exact timing and location of the attack that will result. Autonomous weapon systems are, therefore, clearly distinguishable from other weapon systems, where the specific timing, location and target are chosen by the user at the point of launch or activation.

The ICRC's primary concern is loss of human control over the use of force as a result of autonomy in the critical functions of weapon systems. Depending on the constraints under which a system operates, the user's uncertainty about the exact timing, location and circumstances of the attack(s) may put civilians at risk from the unpredictable consequences of the attack(s). It also raises legal questions, since combatants must make context specific judgements to comply with IHL. And it raises ethical concerns as well, because human agency in decisions to use force is necessary in order to uphold moral responsibility and human dignity.

[23] See chapter II. 2) c. on artificial intelligence and machine learning.

[24] See *ICRC Strategy 2019-2022*, "Strategic orientation 5: Embracing the digital transformation", pp. 22–23; available at https://shop.icrc.org/icrc/pdf/view/id/2844.

[25] See ICRC and Privacy International, *The Humanitarian Metadata Problem: "Doing No Harm" in the Digital Era*, 2018; available at https://www.icrc.org/en/download/file/85089/the_humanitarian_metadata_problem_-_icrc_and_privacy_international.pdf.

Fuller understanding of the legal,[26] military,[27] ethical,[28] and technical[29] aspects of autonomous weapon systems has enabled the ICRC to refine its views.[30] It continues to espouse a human-centred approach, based on its reading of the law and ethical considerations for humans in armed conflict.[31]

Human control under IHL

The ICRC holds that legal obligations under IHL rules on the conduct of hostilities must be fulfilled by those persons who plan, decide on, and carry out military operations. It is humans, not machines, that comply with and implement these rules, and it is humans who can be held accountable for violations. Whatever the machine, computer program, or weapon system used, individuals and parties to conflicts remain responsible for their effects.

Certain limits on autonomy in weapon systems can be deduced from existing rules on the conduct of hostilities – notably the rules of distinction, proportionality and precautions in attack – which require complex assessments based on the circumstances prevailing at the time of the decision to attack, but also during an attack. Combatants must make these assessments reasonably proximate in time to the attack. Where these assessments form part of planning assumptions, they must have continuing validity until the execution of the attack. Hence, commanders or operators must retain a level of human control over weapon systems sufficient to allow them to make context-specific judgments to apply the law in carrying out attacks.

Human control can take various forms during the development and testing of a weapon system ("development stage"); the taking of the decision to activate the weapon system ("activation stage"); and the operation of the weapon system as it selects and attacks targets ("operation stage"). Human control at the activation and operation stages is the most important factor for ensuring compliance with the rules on the conduct of hostilities. Human control during the development stage provides a means to set and test control measures that will ensure human control in use. However, control measures at the development stage alone – meaning control in design – will not be sufficient.

Importantly, however, existing IHL rules do not provide all the answers. Although States agree on the importance of human control – or "human responsibility"[32] – for legal compliance, opinion varies on what this means in practice. Further, purely legal interpretations do not accommodate the ethical concerns raised by the loss of human control over the use of force in armed conflict.

Towards limits on autonomy in weapon systems

In the ICRC's view, the unique characteristics of autonomous weapon systems, and the associated risks of loss of control over the use of force in armed conflict, mean that internationally agreed limits are needed to ensure compliance with IHL and to protect humanity.

Insofar as the sufficiency of existing law – particularly IHL – is concerned, it is clear, as shown above, that existing IHL rules – in particular distinction, proportionality, and precautions in attack – already provide

[26] Neil Davison, "A legal perspective: Autonomous weapon systems under international humanitarian law", in *UNODA Occasional Papers*, No. 30, November 2017; available at https://www.icrc.org/en/document/autonomous-weapon-systems-under-international-humanitarian-law; ICRC, *Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, 2014: available at https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014.

[27] See ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, 2016; available at https://www.icrc.org/en/publication/4283-autonomous-weapons-systems.

[28] See ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, 2018; available at https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control.

[29] See ICRC, *Autonomy, Artificial Intelligence and Robotics:Technical Aspects of Human Control*, 2019; available at https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control.

[30] See ICRC, *IHL Challenges Report 2011*, pp. 39–40. On definitions in particular, see ICRC, *IHL Challenges Report 2015*, p. 45.

[31] See ICRC, *Statements to the Group of Governmental Experts on Lethal Autonomous Weapons Systems,* March 2019; available at https://www.unog.ch/80256EE600585943/(httpPages)/5535B644C2AE8F28C1258433002BBF14?OpenDocument.

[32] United Nations, *Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, CCW/GGE.1/2018/3, 23 October 2018.

limits to autonomy in weapon systems. A weapon with autonomy in its critical functions that is unsupervised, unpredictable and unconstrained in time and space would be unlawful, because humans must make the context-specific judgments that take into account complex and not easily quantifiable rules and principles.

However, it is also clear that existing IHL rules do not provide all the answers. What level of human supervision, intervention and ability to deactivate is needed? What is the minimum level of predictability and reliability of the weapon system in its environment of use? What constraints are needed for tasks, targets, operational environments, time of operation, and geographical scope of operation?

Moreover, the limits dictated by ethical concerns may go beyond those found in existing law. Anxieties about the loss of human agency in decisions to use force, diffusion of moral responsibility, and loss of human dignity are most acute with autonomous weapon systems that present risks for human life, and especially with the notion of anti-personnel systems designed to target humans directly. The principles of humanity may demand limits on or prohibitions against particular types of autonomous weapon and/or their use in certain environments.

At a minimum, there remains an urgent need for agreement on the type and degree of human control necessary in practice to ensure both compliance with IHL and ethical acceptability.

C) ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Artificial intelligence (AI) systems are computer programs that carry out tasks – often associated with human intelligence – that require cognition, planning, reasoning or learning. Machine learning systems are AI systems that are "trained" on and "learn" from data, which ultimately define the way they function. Both are complex software tools, or algorithms, that can be applied to many different tasks. However, AI and machine learning systems are distinct from the "simple" algorithms used for tasks that do not require these capacities. The potential implications for armed conflict – and for the ICRC's humanitarian work – are broad.[33] There are at least three overlapping areas that are relevant from a humanitarian perspective.

The first area is the use of AI and machine learning tools to control military hardware, in particular the growing diversity of unmanned robotic systems – in the air, on land, and at sea. AI may enable greater autonomy in robotic platforms, whether armed or unarmed. For the ICRC, autonomous weapon systems are the immediate concern (see above). AI and machine learning software – particularly for "automatic target recognition" – could become a basis for future autonomous weapon systems, amplifying core concerns about loss of human control and unpredictability. However, not all autonomous weapons incorporate AI.[34]

The second area is the application of AI and machine learning to cyber warfare: AI-enabled cyber capabilities could automatically search for vulnerabilities to exploit, or simultaneously defend against cyber attacks while launching counter-attacks, and could therefore increase the speed, number and types of attacks and their consequences. These developments will be relevant to discussions about the potential human cost of cyber warfare. AI and machine learning are also relevant to information operations, in particular the creation and spread of false information (whether intended to deceive or not). AI-enabled systems can generate "fake" information – whether text, audio, photos or video – that is increasingly difficult to distinguish from "real" information and might be used by parties to a conflict to manipulate opinion and influence decisions. These digital risks can pose real dangers for civilians (see above).[35]

The third area, and the one with perhaps the most far-reaching implications, is the use of AI and machine learning systems for decision-making. AI may enable widespread collection and analysis of multiple data sources to identify people or objects, assess "patterns of life" or behaviour, make recommendations for

[33] See ICRC, *Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centred Approach*, 2019; available at https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control.

[34] ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, 2019; available at https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control.

[35] See ICRC, *Digital Risks in Situations of Armed Conflict.*

courses of action, or make predictions about future actions or situations. The possible uses of these "decision-support" or "automated decision-making" systems are extremely broad: they range from decisions about whom – or what – to attack and when, and whom to detain and for how long, to decisions about overall military strategy – even on use of nuclear weapons - as well as specific operations, including attempts to predict, or pre-empt, adversaries.

AI and machine learning-based systems can facilitate faster and broader collection and analysis of available information. This may enable better decisions by humans in conducting military operations in compliance with IHL and minimizing risks for civilians. However, the same algorithmically-generated analyses, or predictions, might also facilitate wrong decisions, violations of IHL and exacerbated risks for civilians. The challenge consists in using all the capacities of AI to improve respect for IHL in situations of armed conflict, while at the same time remaining aware of the significant limitations of the technology, particularly with respect to unpredictability, lack of transparency, and bias. The use of AI in weapon systems must be approached with great caution.

A human-centred approach

AI and machine learning systems could have profound implications for the role of humans in armed conflict. The ICRC is convinced of the necessity of taking a human-centred, and humanity-centred, approach to the use of these technologies in armed conflict.

It will be essential to preserve human control and judgement in using AI and machine learning for tasks, and in decisions, that may have serious consequences for people's lives, and in circumstances where the tasks – or decisions – are governed by specific IHL rules. AI and machine learning systems remain tools that must be used to serve human actors, and augment and improve human decision-making, not to replace them.

Ensuring human control and judgement in AI-enabled tasks and decisions that present risks to human life, liberty, and dignity will be needed for compliance with IHL and to preserve a measure of humanity in armed conflict. In order for humans to meaningfully play their role, these systems may need to be designed and used to inform decision-making at "human speed" rather than accelerate decisions to "machine speed".

The nature of human-AI interaction required will likely depend on the specific application, the associated consequences, and the particular IHL rules and other pertinent law that apply in the circumstances – as well as on ethical considerations.

However, ensuring human control and judgement in the use of AI systems will not be sufficient in itself. In order to build trust in the functioning of a given AI system, it will be important to ensure, including through weapon reviews: predictability and reliability – or safety – in the operation of the system and the consequences of its use; transparency – or explainability – in how the system functions and why it reaches its output; and lack of bias in the design and use of the system.

D) HUMANITARIAN CONSEQUENCES AND CONSTRAINTS UNDER IHL RELATED TO THE POTENTIAL USE OF WEAPONS IN OUTER SPACE

Military use of space objects has been an integral part of warfare for several decades. It includes the use of satellite imagery to support the identification of enemy targets and the use of satellite communication systems for command-and-control, and more recently, for remotely controlled means of warfare. The weaponization of outer space would further increase the likelihood of hostilities in outer space, with potentially significant humanitarian consequences for civilians on earth.

The exact scope of the potential humanitarian consequences of the use of weapons in outer space is uncertain. It is clear, however, that the use of weapons in outer space – be it through kinetic or non-kinetic means (such as electronic, cyber or directed energy attacks), using space – and/or ground-based weapon systems – could directly or incidentally disrupt, damage, destroy or disable civilian or dual-use space objects on which safety-critical civilian activities and essential civilian services depend. This includes the navigation satellite systems (such as BeiDou, Galileo, GLONASS, and GPS) that are increasingly employed in civilian vehicles, shipping, and air traffic controls. Satellites are also critical for the weather services used for disaster prevention and mitigation, and for the satellite phone services on which the delivery of humanitarian assistance and emergency relief is reliant.

The use of weapons in outer space would not occur in a legal vacuum. It is constrained by existing law, notably the Outer Space Treaty,[36] the UN Charter, and IHL rules governing means and methods of warfare.

The applicability of IHL in outer space is confirmed by Article III of the Outer Space Treaty, which states that international law applies to the use of outer space; and IHL forms part of international law. Furthermore, the International Court of Justice has recalled that the established principles and rules of IHL applicable in armed conflict apply "to all forms of warfare and to all kinds of weapons, those of the past, those of the present and those of the future".[37] In terms of treaty law, the four 1949 Geneva Conventions and Protocol I of 8 June 1977 additional to the Geneva Conventions (Additional Protocol I) apply "to all cases of declared war or any other armed conflict which may arise between two or more of the High Contracting Parties".[38] Article 49(3) of Additional Protocol I shows that the Protocol's rules on the conduct of hostilities are meant to apply to all types of warfare that may affect civilians on land. This would include hostilities in outer space.

IHL applies to any military operations conducted as part of an armed conflict, including those occurring in outer space, regardless of whether or not the use of force is lawful under the UN Charter (*jus ad bellum*). IHL does not legitimize the use of force in outer space; nor does it encourage the militarization or weaponization of outer space. The sole aim of IHL is to preserve a measure of humanity in the midst of armed conflict, notably to protect civilians.

The Outer Space Treaty prohibits the placement in orbit around the earth of objects carrying nuclear weapons or other weapons of mass destruction, the instalment of such weapons on celestial bodies, and the stationing of such weapons in outer space in any manner. It also forbids the establishment of military bases, installations and fortifications, the testing of any type of weapon, and the conduct of military manoeuvres on celestial bodies; it also requires that celestial bodies be used exclusively for peaceful purposes. For its part, IHL notably prohibits weapons that are indiscriminate by nature, as well as a number of other specific types of weapon. These prohibitions are not limited to the terrestrial domains.

Even when resorting to weapons that are not prohibited, a belligerent has to respect the IHL rules governing the conduct of hostilities. These include the principle of distinction, the prohibition against indiscriminate and disproportionate attacks, and the obligation to take precautions in attack and against the effects of attack. Furthermore, attacking, destroying, removing or rendering useless objects indispensable to the survival of the civilian population is prohibited. While specific protections, such as the latter, apply to a broad range of military operations, the rules affording general protection to civilian objects apply mostly in relation to attacks. Under IHL, a kinetic operation against a space object would constitute an attack. However, a space object could also be disabled (rendered dysfunctional) without being physically damaged, for example by directed energy/laser weapons or a cyber attack. In the ICRC's view, such non-kinetic operations would constitute attacks under IHL.

IHL forbids targeting civilian objects in outer space. However, civilian satellites or some of their hosted payloads may also be used by the armed forces, meaning they are of a dual-use nature. They may become military objectives, provided that their use for military purposes is such that they fulfil the definition under Article 52(2) of Additional Protocol I. If such a dual-use satellite or its payload is attacked, the expected incidental harm to civilians and civilian objects, directly or through knock-on effects, must be taken into consideration while assessing the legality of the attack under the principles of proportionality and precautions. Furthermore, the consequences for civilians of putting an end to or impairing the civilian use of the targeted

[36] 1967 Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and other Celestial Bodies.

[37] International Court of Justice, *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion, 8 July 1996, para. 86.

[38] Art. 1(3), Additional Protocol I; Art. 2 common to the four 1949 Geneva Conventions.

satellite or payload must also be considered. As noted above, disabling the civilian functions of satellites could disrupt large segments of modern-day societies, especially if they support safety-critical civilian activities and essential civilian services on earth.

Another issue of concern is the risk posed by space debris. Debris can be created by a number of space activities. A kinetic attack on a satellite, for example, risks causing far more debris than other space activities. Debris may continue to travel in the orbits in which it was produced for decades or more. Given the speed at which it travels, debris risks damaging other satellites supporting civilian activities and services. This would have to be considered in – and may limit – the choice of means and methods of warfare in outer space.

The ICRC is concerned by the potentially high human cost of the use of weapons in outer space. It recommends that future multilateral processes acknowledge:

- the potentially significant humanitarian consequences, for civilians on earth, of the use of weapons in outer space
- the protection afforded by the IHL rules that restrict belligerents' choice of means and methods of warfare, including in outer space.[39]

As with the development of any new means or methods of warfare, the weaponization of outer space is not inevitable but a choice. States may decide to set limits in this regard for a range of reasons, including humanitarian ones. The fact that IHL applies does not prevent States from agreeing on additional rules to prohibit or limit specific military activities or weapons in outer space, as they did in the Outer Space Treaty. States may decide that further prohibitions or limitations may be warranted to reduce the risks of the significant civilian harm that could ensue from the use of weapons in outer space.

E) CHALLENGES POSED BY CERTAIN NEW TECHNOLOGIES OF WARFARE TO LEGAL REVIEWS OF NEW WEAPONS

As noted above, the development and use of new technologies of warfare, such as autonomous weapon systems or military cyber capabilities, do not occur in a legal vacuum. As with all weapon systems, they must be capable of use in compliance with IHL, particularly its rules on the conduct of hostilities. The responsibility for ensuring this rests with every State that is developing, acquiring and using these new technologies of warfare. In this respect, legal reviews are as critical now as they were when Article 36 of Additional Protocol I was conceived during the Cold War arms race. To assist States in implementing this obligation, in 2006, the ICRC published *A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol I of 1977*. What follows is drawn from that *Guide* and addresses new questions regarding the challenges to legal reviews posed by new technologies of warfare.

Every State party to Additional Protocol I is obliged to determine whether the employment of a new weapon, means or method of warfare that it studies, develops, acquires or adopts would, in some or all circumstances, be prohibited by international law.[40] In the ICRC's view, the requirement to carry out legal review of new weapons also flows from the obligation to ensure respect for IHL under Article 1 common to the Geneva Conventions.[41] Besides these legal requirements, all States also have an interest in assessing the lawfulness of new weapons. Legal reviews are a critical measure to help ensure that a State's armed forces can conduct

[39] See also ICRC, "Humanitarian consequences and constraints under international humanitarian law (IHL) related to the potential use of weapons in outer space", working paper submitted to the Group of Government Experts on Further Practical Measures for the Prevention of an Arms Race in Outer Space, 2019; available at https://undocs.org/GE-PAROS/2019/WP.1.

[40] Sweden and the United States, for example, first established mechanisms for legal review in 1974, three years before the adoption of Additional Protocol I.

[41] This is also the view of some States. See Australia, "The Australian Article 36 review process", working paper *submitted to the Group of Government Experts of the High Contracting Parties 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* (CCW), 2018, para. 3; available at https://unog.ch/80256EDD006B8954/(httpAssets)/46CA9DABE945FDF9C12582FE00380420/$file/2018_GGE+LAWS_August_Working+paper_Australia.pdf; The Netherlands and Switzerland, "Weapons review mechanisms", working paper submitted to the CCW, 2017, para. 17.

hostilities in accordance with that State's international obligations. They also help prevent the costly consequences of approving and procuring a weapon the use of which is likely to be restricted or prohibited.

Weapon systems of all types should be subjected to legal review, including physical systems (hardware) and digital systems (software). This extends to military cyber capabilities intended for use or expected to be used in the conduct of hostilities. It also includes software components that form part of the weapon system (the "means" of warfare) or the way in which the system will be used (the "method" of warfare), such as software that controls a physical system or supports decision-making processes for use of that weapon system. Since a weapon cannot be assessed in isolation from the way in which it will be used, the normal or expected use of the weapon must be considered in the legal review.

Weapons that include a software component that permits the critical functions of selection and attack of targets (the defining characteristics of autonomous weapon systems) to be triggered by the weapon system's environment, rather than by a commander, make it challenging to assess whether the weapon can be used in compliance with IHL rules. A reviewer will need to be satisfied that the proposed weapon's design and method of use will not prevent a commander from exercising the judgement required by IHL. If the reviewer is not satisfied of this, they must not allow the weapon to be used; alternatively, they may need to impose limitations on the weapon's use to ensure the commander's ability to comply with IHL.

Foreseeing the effects of weapon systems through testing may become increasingly difficult, as weapon systems become more complex or are given more freedom of action in their tasks, and therefore become less predictable, such as weapon systems that incorporate machine learning. Unpredictability in the functioning of the system, and the interaction of the system with a dynamic environment, cannot be simulated in advance of use. This challenge will be compounded, in some cases, by the inability of the commander to understand how a weapon system using artificial intelligence – particularly machine learning – reaches its output from a given input, which makes it difficult (if not impossible) to foresee the consequences of its use.

For legal reviews to be effective, States that develop or acquire new weapon technologies need to navigate these complexities. Therefore, legal reviews of weapons, means and methods of warfare, relying on these new technologies may need to be conducted at an earlier stage of weapon development, and at shorter intervals, than for more traditional technologies, and may need to be repeated during development. The unique characteristics of new technologies and the related processes of legal review require new standards of testing and validation. States should also share information about their legal-review mechanisms and, to the extent feasible, about the substantive results of their legal reviews, especially where a weapon's compatibility with IHL may be in question – so that other States will not encounter the same problems and can benefit from reviewing States' conclusions on whether the use of the weapon in question is prohibited or restricted by IHL. When States exchange information about conducting legal reviews of new technologies, it can help build expertise and identify good practices, and also assist States that wish to establish or strengthen their own mechanisms.