Artificial Intelligence and the Obligation to Respect and to Ensure Respect for International Humanitarian Law

Hitoshi Nasu

Exeter Centre for International Law

Exeter Centre for International Law

The Exeter Centre for International Law builds on a long and distinguished tradition of international legal scholarship at Exeter Law School. The Centre's mission is to provide an intellectual environment for the study and development of international law and to stimulate discussion and collaboration in response to the most pressing challenges facing the international community. As part of this mission, the Centre publishes the present Working Paper Series.

Artificial Intelligence and the Obligation to Respect and to Ensure Respect for International Humanitarian Law

Hitoshi Nasu[*]

Abstract

Artificial Intelligence is a discipline of science and engineering of building intelligent machine capable of acting with an appropriate forethought to achieve a task in a complex environment. Its technological application in the context of armed conflict has been anticipated, which has caused significant scholarly and public debate primarily concerning the legality of lethal autonomous weapons systems. However, artificial intelligence technology has the potential for far wider applications in military use, such as target recognition, deception, communication, and the research and development of new weapons. Apart from the legality of specific lethal use, consideration therefore must extend to the general legal obligation to respect and to ensure respect for international humanitarian law so as to address a wide range of challenges and opportunities that artificial intelligence uniquely presents. This paper specifically considers practical measures that States can employ to implement their obligation to respect and ensure respect for international humanitarian law as modern science and technology advances in the area of artificial intelligence.

1. Introduction

Artificial Intelligence (AI) is a discipline of science and engineering of building intelligent machine, which is one that can achieve a specific task in a complex and uncertain environment.[1]

---

[*] Professor of International Law, University of Exeter.

[1] In 1950, Alan Turing proposed the 'Imitation Game' as a method of inquiring whether a computer thinks like a human being: A M Turing, 'Computing Machinery and Intelligence' (1950) 236 *Mind* 433, 433-435. The Turing Test has been both influential and highly criticised, however, there is no universally agreed definition of artificial intelligence to date.

The technology is increasingly used in a wide range of tasks such as image recognition, speech recognition, natural language processing, health care and drug discovery. Its application for military purposes is no exception to this trend, with many technologically developed nations heavily investing in the development and military applications of AI. The United States (US), in its 2008 National Defense Strategy, commits itself to broad investment in military application of autonomy, AI, and machine learning, including rapid application of commercial breakthroughs, to gain competitive military advantages. [2] The People's Republic of China released an Artificial Intelligence Development Plan on 20 July 2017, outlining the national policy to take the lead in AI technology by 2030 including its application for the enhancement of national defence and the protection of national security.[3] The Russian Military Industrial Committee, a body responsible for Russia's military industrial policy, has reportedly set a target of integrating AI and robotic technologies into 30 per cent of military equipment by 2025.[4]

Its technological application in the context of armed conflict has been anticipated, which has caused significant scholarly and public debate concerning the legality of lethal autonomous weapons systems.[5] However, AI technology has the potential for

---

[2] US National Defense Strategy Summary: Sharpening the American Military's Competitive Edge (2018) 7, available at https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

[3] People's Republic of China, New Generation Artificial Intelligence Development Plan (2017) Section II(4), available at http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm (English translation by Graham Webster, Rogier Creemers, Paul Triolo, and Elsa Kania at https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/).

[4] Institute of Land Warfare, 'Integrating Army Robotics and Autonomous Systems to Fight and Win' (Association of the United States Army, ILW Spotlight 17-2, 24 July 2017) 1, available at https://www.ausa.org/publications/integrating-army-robotics-and-autonomous-systems. Cf Alina Polyakova, 'Weapons of the Weak: Russia and AI-Driven Asymmetric Warfare' (Brookings Institution, 15 November 2018), available at https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/.

[5] Most notably, questions related to emerging technologies in the area of lethal autonomous weapons systems were discussed at the 2014-2016 Informal Meetings of the States Parties to the Convention on Certain Conventional

far wider applications in military use, such as target recognition, deception, communication, and the research and development of new weapons. Apart from the legality of specific lethal use, consideration therefore must extend to the general legal obligation, under Common Article 1 of the 1949 Geneva Conventions and 1977 Additional Protocol I as reflective of customary international law,[6] to respect and ensure respect for international humanitarian law so as to address a wide range of challenges and opportunities that AI uniquely presents.

This paper specifically considers practical measures that States can employ to implement their obligation to respect and ensure respect for international humanitarian law as modern science and technology advances in the area of AI. To that end, it first surveys unique capabilities that AI could potentially offer to enhance military operations. It then examines weapons review and export control as the means to implement the obligation to respect and ensure respect for international humanitarian law, while identifying challenges that AI poses. It concludes with a brief observation regarding practical challenges to the implementation of the general obligation to respect and to ensure respect for international humanitarian law as AI technology becomes mature and readily available for use in hostilities.

Weapons (CCW). At the Fifth CCW Review Conference an open-ended Group of Governmental Experts was established to discuss relevant questions.

[6] Geneva Convention for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field of August 12, 1949, 75 UNTS 31 (entered into force 21 October 1950); Geneva Convention for the Amelioration of the Condition of Wounded, Sick and Shipwrecked Members of Armed Forces at Sea of August 12, 1949, 75 UNTS 85 (entered into force 21 October 1950); Geneva Convention Relative to the Treatment of Prisoners of War of August 12, 1949, 75 UNTS 135 (entered into force 21 October 1950); Geneva Convention Relative to the Protection of Civilian Persons in Time of War of August 12, 1949, 75 UNTS 287 (entered into force 21 October 1950); Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), adopted 8 June 1977, 1125 UNTS 3 (entered into force 7 December 1978) [hereinafter Additional Protocol I]; Jean-Marie Henckaerts and Louise Doswald-Beck, *Customary International Humanitarian Law* (Cambridge University Press, 2005) Rule 139, as updated by the International Committee of the Red Cross (ICRC) at https://ihl-databases.icrc.org/customary-ihl/eng/docs/home.

2. The Role of Artificial Intelligence in Future Warfare

Many transformative technologies such as nuclear power, information technology and autonomous navigation system were originally developed for military purposes, which then found commercial application, for example, in nuclear energy production, the commercial use of the Internet and drones respectively. Unlike these technologies, AI is largely driven by industrial research as a general-purpose technology, with the potential to deliver diverse capabilities that influence both military and economic power. [7] The shift in advanced technological development from military to commercial settings is evident in the area of autonomous systems development, such as driverless cars. While it emerged from a military research programme at the US Defense Advanced Research Projects Agency in 2004,[8] its progress was rather slow and rudimentary compared to the commercial development that proliferated a decade later.[9]

At the current stage the progress for AI is particularly driven by three enabling technological advances: (1) availability of 'big data' sources; (2) increases in computer processing power; and (3) innovation in machine learning methods.[10] Machine learning is a method of data analysis that uses algorithms to build a computerised programme to improve its predictive performance

[7] Michael C Horowitz, Gregory C Allen, Elsa B Kania, and Paul Scharre, 'Strategic Competition in an Era of Artificial Intelligence' (Center for a New American Security, July 2018) 3-5, available at https://www.cnas.org/publications/reports/strategic-competition-in-an-era-of-artificial-intelligence.

[8] Defense Advanced Research Projects Agency (DARPA), 'Grand Challenge 2004: Final Report' (30 July 2004), available at https://www.esd.whs.mil/Portals/54/Documents/FOID/Reading%20Room/DARPA/15-F-0059_GC_2004_FINAL_RPT_7-30-2004.pdf.

[9] Mary L Cummings, 'Artificial Intelligence and the Future of Warfare' (Chatham House, January 2017) 9-12, available at https://www.chathamhouse.org/publication/artificial-intelligence-and-future-warfare.

[10] National Science and Technology Council, 'Preparing for the Future of Artificial Intelligence' (Office of Science and Technology Policy, 12 October 2016) 6, available at https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

at certain tasks based on observed data. [11] Of particular significance is a method of statistical learning called 'Deep Learning', which uses multi-layered algorithms modelled upon the human brain. Although the Deep Learning draws inspiration from the human brain as the theoretical basis, critical differences exist between the human brain and artificial neural networks.[12] For example, the latter requires, for it to function, an accurate and large volume of datasets, which are processed according to algorithm-based statistical operations to produce probabilistic outputs.[13] The algorithm-based probabilistic reasoning based on highly complicated statistical operations or geometric relationships that humans cannot visualise creates a 'black box' problem – it is difficult for humans to predict the decision or output that AI produces, or understand its decision-making process.[14] AI is not, at the current stage of development, a technology that makes a machine think like a human, but rather one that enables algorithm-based computerised programme to perform certain tasks in a complex and uncertain environment, often with greater accuracy and efficiency than human beings.

The potential applications of AI are envisaged in such areas as intelligence, surveillance and reconnaissance (ISR), logistics and transport, multi-domain command and control, cyber defence, information manipulation, target recognition, and weapons development.[15] For example, the US Department of Defense

---

[11] See eg, Zoubin Ghahramani, 'Probabilistic Machine Learning and Artificial Intelligence' (2015) 521 *Nature* 452; Tom Mitchell, *Machine Learning* (McGraw-Hill, 1997).

[12] See eg, Jeff Hawkins, 'What Intelligent Machines Need to Learn from the Neocortex', *IEEE Spectrum*, 2 June 2017, available at https://spectrum.ieee.org/computing/software/what-intelligent-machines-need-to-learn-from-the-neocortex.

[13] UK Government Office for Science, 'Artificial Intelligence: Opportunities and Implications for the Future of Decision Making' (2017) 6-7, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf.

[14] Yavar Bathaee, 'The Artificial Intelligence Black Box and the Failure of Intent and Causation' (2018) 31 *Harvard Journal of Law & Technology* 889, 897-906; Davide Castelvecchi, 'Can We Open the Black Box of AI' (2016) 538 *Nature* 20.

[15] For details, see eg, Congressional Research Service, 'Artificial Intelligence and National Security' (US Congress, 30 January 2019) 9-15, available at https://fas.org/sgp/crs/natsec/R45178.pdf; Daniel Fiott and Gustav Lindstrom, 'Artificial Intelligence: What Implications for EU Security and Defence?' (European Union Institute for Security Studies, November 2018) 4-6, available

initiated Project Maven in developing advanced computer algorithms that would enhance the ability to extract objects from massive volumes of satellite and drone surveillance data feeds.[16] China has reportedly been developing and using advanced facial recognition technology to track and control the Uighurs, the ethnic Muslim minority group in the western region of Xinjiang.[17]

Many of these applications are dual-use, meaning that they can be used for both military and civilian purposes. They are designed to enhance existing capabilities with algorithm-based probabilistic reasoning, which enables certain complex tasks to be performed with greater accuracy and efficiency. For example, algorithms developed to search and categorise holiday photos can be repurposed to scour spy satellite imagery.[18] Medical algorithms developed for drug discovery and clinical testing could be proven useful in identifying previously unknown causes of disease,[19] which could be manipulated to develop harmful substances that target a particular group of people with specific biological profiles.

As a technology built on computer algorithms, AI is relatively transparent and accessible in that algorithms can be shared in the form of software or scientific papers, which can be incorporated into an existing system.[20] AI does not necessarily take the form of a tangible object to trade as a product and, therefore, is easily transferrable and amenable to rapid diffusion. It is also scalable in that the system can be easily replicated in great numbers to

---

at https://www.iss.europa.eu/content/artificial-intelligence-–-what-implications-eu-security-and-defence.

[16] Cheryl Pellerin, 'Project Maven to Deploy Computer Algorithms to War Zone by Year's End' (US Department of Defense, 21 July 2017), available at https://dod.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

[17] Paul Mozur, 'One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority', *New York Times*, 14 April 2019, available at https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html.

[18] Tom Simonite, 'For Superpowers, Artificial Intelligence Fuels New Global Arms Race', *The Wired*, 8 September 2017, available at https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race/.

[19] Nic Fleming, 'Computer-Calculated Compounds: Researchers Are Deploying Artificial Intelligence to Discover Drugs' (2018) 557 *Nature* S55, S56.

[20] Congressional Research Service, above n 15, 3

perform the same task.[21] Powerful machine learning techniques such as Deep Learning require large datasets to achieve high performance and computing resources to train machines.[22] However, AI may proliferate when technologically less capable actors can make use of previously trained systems.

It is likely to take decades before science advances to develop general AI with advanced forms of cognitive reasoning capable of application across an undefined range of parameters. However, in the short to medium term, the industry is expected to develop and produce a range of narrow and specific AI systems, whether these are specifically designed for military purposes or capable of military application. With the progressive development of AI, there will be the need to revisit the relevant rules of international humanitarian law in determining how those rules that have been developed to regulate the conduct of States and individuals might extend to the use of AI as it starts assuming the tasks that human beings traditionally performed on the battlefield.

3. Means to Implement the Obligation to Respect and Ensure Respect

The obligation to respect and ensure respect for international humanitarian law in all circumstances is primarily derived from Common Article 1 of the 1949 Geneva Conventions and Additional Protocol I,[23] but is also widely considered as reflective of customary international law derived from 'the general principles of humanitarian law to which the Conventions merely give specific expression'.[24] The scope of its application, as well as the specific nature of the obligation, has been the subject of considerable debate.[25] However, there is no disagreement

---

[21] Miles Brundage et al, 'The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation' (February 2018) 16-17, available at https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf.

[22] Horowitz et al, above n 7, 5-6.

[23] See above n 6.

[24] *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v USA) (Merits)* [1986] ICJ Rep 14, 114 para 220. See also Henckaerts and Doswald-Beck, above n 6, Rule 139.

[25] See generally Robin Geiß, 'The Obligation to Respect and to Ensure Respect for the Conventions' in Andrew Clapham, Paola Gaeta, and Marco Sassòli (eds), *The 1949 Geneva Conventions: A Commentary* (Oxford University Press, 2015) 111-34.

regarding the non-reciprocal and general character of the obligation, which means, at the very least, that the obligation applies irrespective of reciprocal acceptance or implementation of obligations by the other parties to the conflict.

It is also widely accepted that this obligation encompasses both a negative duty of States to refrain from violating international humanitarian law, including the obligation not to encourage, aid or assist the commission of violation,[26] and a positive duty to undertake all measures necessary to comply with their obligations under applicable rules of international humanitarian law in peacetime or in situations of armed conflict.[27] It is an obligation of conduct to be exercised with due diligence to prevent and suppress breaches of international humanitarian law, with the choice of means dependent upon the attendant circumstances.[28] There is no disagreement on the internal compliance aspect of this obligation even with the introduction of new technologies, such as AI. The contentious question rather concerns whether this duty of due diligence applies only to military forces and private individuals within a State's own jurisdiction or extends to individuals under the control or authority of that State, or even to other States engaging in an armed conflict.[29] AI, as the technology that allows certain tasks to be performed as a substitute for human labour, can be characterised as an object to which this obligation applies, as well as the means to facilitate compliance with applicable rules of international humanitarian law.

---

[26] *Nicaragua* Judgment, above n 24, 114 para 220.

[27] See Geiß, above n 25, 117-20 and 130-2.

[28] ICRC, *Commentary on the First Geneva Convention: Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field* (2nd ed, Cambridge University Press, 2016) para 150; Geiß, above n 25, 127-8; Birgit Kessler, 'The Duty to Ensure Respect under Common Article 1 of the Geneva Conventions: Its Implications on International and Non-International Armed Conflicts' (2001) 44 *German Yearbook of International Law* 498, 506-7. For a range of measures available, see Umesh Palwankar, 'Measures Available to States for Fulfilling Their Obligation to Ensure Respect for International Humanitarian Law' (1994) 34 *International Review of the Red Cross* 9.

[29] Compare eg, ICRC, above n 28, paras 155-6; with Frits Kalshoven, 'The Undertaking to Respect and Ensure Respect in All Circumstances: From Tiny Seed to Ripening Fruit' (1999) 2 *Yearbook of International Humanitarian Law* 3.

While the means to implement this general obligation is context-dependent, there are also various specific duties, such as giving orders and instructions to ensure observance of the Geneva Conventions, ensuring that legal advisers are available to military commanders, and disseminating the texts of the Conventions.[30] These specific duties are instrumental to the implementation of various rules of international humanitarian law in good faith.[31] AI may provide additional tools, for example in combat simulation for military training, to facilitate the implementation of these specific duties as part of the general obligation to respect and ensure respect for international humanitarian law. On the other hand, there is a limited range of options to control new technologies, such as AI, to the extent that these would come within the scope of this general. The remainder of this paper focuses on weapons review and export control as the specific means to implement this obligation in relation to the study, development, acquisition, adoption, use and transfer of AI technology.

3.1. Weapons Review

With respect to new technologies in general, an obligation to conduct legal review may arise in order to ensure lawfulness of their employment in warfare under the applicable rules of international law. Article 36 of Additional Protocol I to the Geneva Conventions specifically provides that:

> In the study, development, acquisition or adoption of a new weapons, means or method of warfare, a High Contracting Party is under an obligation 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.

This obligation, in relation to new weapons and means of warfare, is also implicit in any weapons law obligations under customary international law that prohibits, for example, weapons of a nature to cause superfluous injury or unnecessary suffering or of an

---

[30] Additional Protocol I, above n 6, arts 80(2), 82 and 83.

[31] See eg, Carlo Focarelli, 'Common Article 1 of the 1949 Geneva Conventions: A Soap Bubble?' (2010) 21 *European Journal of International Law* 125, 151-154.

indiscriminate nature, as well as those that each State may have under various treaties it is party to. [32] Therefore, the core obligation to conduct legal review of weapons as a means to ensure compliance with any applicable rule of weapons law is reflective of customary international law, even though Article 36 of Additional Protocol I, in and of itself, may not satisfy the requirements for such status.[33]

Concerns over automation in conducting warfare were already looming at the background during the drafting of Additional Protocol I, with weapons review under Article 36 envisaged as the solution by placing responsibility within the domestic governance of nations.[34] Thus, any incorporation or integration of AI into the existing weapons system or weapons development programmes to build new capabilities will be subject to the obligation to conduct weapons review. However, the relevance of weapons review to the regulation of the study, development, acquisition or adoption of AI as the means to implement the obligation to respect and to ensure respect is rather limited for the following three main reasons.

First, only a handful of States are known to have systematic approaches to the legal review of new weapons.[35] The research and development of AI is currently pursued in many different countries, including those that have not traditionally been known for weapons manufacturing and therefore may not have systematic weapons review processes in place. Many technology companies engaging in the research and development of AI

---

[32] For differences between the weapons review obligation under customary international law and under Article 36 of Additional Protocol I, see eg, Damian P Copeland, 'Legal Review of New Technology Weapons' in Hitoshi Nasu and Robert McLaughlin (eds), *New Technologies and the Law of Armed Conflict* (TMC Asser, 2014) 43, 47-52.

[33] On the customary international law status of Article 36, see Natalia Jevlevskaja, 'Weapons Review Obligation under Customary International Law' (2018) 94 *International Law Studies* 186.

[34] Yves Sandoz et al, *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (ICRC, 1987) 427-428 paras 1476-1478.

[35] ICRC, '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' (2006) *88 International Review of the Red Cross* 931, 934 fn 8; James D Fry, 'Contextualized Legal Reviews for the Methods and Means of Warfare: Cave Combat and International Humanitarian Law' (2006) 44 *Columbia Journal of Transnational Law* 453, 473-479.

independently from any military funding programmes or procurement are not subject to government regulation requiring weapons review. This means that many computerised programmes and associated innovations in AI, including those designed for hostile purposes or with potential military applications, will be left unchecked.

Second, given the dual-use nature of AI technology, many applications of AI may not qualify as 'weapons' or 'means of warfare' for the purpose of this legal review obligation. These terms are not defined in Article 36 of Additional Protocol I, however, "weapons" are widely considered to mean any instrument of warfare that is 'used, designed, or intended to be used to cause injury to, or death of, persons or damage to, or destruction of, objects'.[36] The 'means of warfare' encompasses both stand-alone weapons (such as guns, bombs and other munition) and weapon systems (including delivery systems such as ballistic missiles and unmanned aerial vehicles).[37] AI is not an independent weapon system of its own, but rather is more likely to form part of a weapon system.[38] The obligation to conduct weapons review arises only when AI is incorporated or integrated into a particular application to serve its designed or intended function. This presents an issue regarding the extent to which States are reasonably expected to exercise due diligence in

---

[36] *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (prepared by the International Group of Experts at the Invitation of the NATO Cooperative Cyber Defence Centre of Excellence, 2017) Rule 103 para 2. See also *Manual on International Law Applicable to Air and Missile Warfare* (Program on Humanitarian Policy and Conflict Research at Harvard University, 2009) Rule 1(ff). For discussion of the meaning of these terms, see eg, Duncan Blake and Joseph S Imburgia, '"Bloodless Weapons"? The Need to Conduct Legal Reviews of Certain Capabilities and the Implications of Defining Them as "Weapons"' (2010) 66 *Air Force Law Review* 157, 168-72; W Hays Parks, 'Conventional Weapons and Weapons Review' (2005) 8 *Yearbook of International Humanitarian Law* 55, 115-9.

[37] See eg, Danish Ministry of Defence and Defence Command Denmark, *Military Manual on International Law Relevant to Danish Armed Forces in International Operations* (September 2016) 380.

[38] See eg, New Zealand Defence Force, *DM69 Manual of Armed Forces Law: Vol 4 Law of Armed Conflict* (2nd ed, 2017) §7.4.6 (specifically referring to 'weapons with artificial intelligence'). See also Michael N Schmitt, 'War, Technology and the Law of Armed Conflict' in Anthony M Helm (ed), *The Law of War in the 21st Century: Weaponry and the Use of Force* (US Naval War College, International Law Studies Vol. 82, 2006) 137, 142.

relation to the AI that has been developed outside the military industry but can be converted for military use by other actors

Third, in the course of weapons review, the legality of a new weapon is examined only in light of the broad and general circumstances in which the weapon is intended for use.[39] This means that it does not cover a particular use of a weapon, which is to be assessed against the targeting rules of distinction, proportionality, and precautions in the operational context of a particular attack.[40] Moreover, for the use of AI as a method of warfare in a particular operational context, the legal review of AI itself is of limited value. For example, AI-enabled satellite imagery analysis could be employed for the purpose of target selection, but at least as far as the customary international law obligation is concerned, that particular method of warfare does not attract a positive duty to conduct legal assessment as a means to ensure its compliance with international humanitarian law.[41] The analysis may differ under Article 36 of Additional Protocol I due to the inclusion of methods of warfare in the review obligation.[42]

These limitations qualify the practical significance of weapons review in relation to AI, as a means to implement the obligation to respect and ensure respect for international humanitarian law. Nevertheless, it provides an essential process to ensure that AI-enabled weapon systems are capable of complying with international humanitarian law rules; in other words, their ability to function within the parameters set by legal requirements when the systems are used as intended.[43] This aspect of weapons review

---

[39] William H Boothby, *Weapons and the Law of Armed Conflict* (2nd ed, Oxford University Press, 2016) 346-7.

[40] US Department of Defense, *Law of War Manual* (June 2015, as updated in December 2016) §6.5.9.3.

[41] Cf *Tallinn Manual 2.0*, above n 36, Rule 110 para 5 (noting disagreement among experts as to whether this obligation under customary international law extends to methods of warfare).

[42] See Alan Backstrom and Ian Henderson, 'New Capabilities in Warfare: An Overview of Contemporary Technological Developments and the Associated Legal and Engineering Issues in Article 36 Weapons Reviews' (2012) 94 *International Review of the Red Cross* 483, 490-5; Justin McClelland, 'The Review of Weapons in Accordance with Article 36 of Additional Protocol I' (2003) 85 *International Review of the Red Cross* 397, 405-6.

[43] William H Boothby, 'Highly Automated and Autonomous Technologies' in William H Boothby (ed), *New Technologies and the Law in War and Peace* (Cambridge University Press, 2018) 137, 145-7; Boothby, above n 39, 348-9.

will have to play a critical role in implementing the general obligation to respect and ensure respect for international humanitarian law as AI becomes incorporated and integrated into weapon systems.

3.2. Exports Control

There is no obligation to regulate arms export under customary international law,[44] nor is it explicitly provided in the Geneva Conventions or the Protocols thereto. Nevertheless, there is a range of arms control regimes that can be considered to facilitate the implementation of the obligation to ensure respect for international humanitarian law.

The Arms Trade Treaty prohibits any transfer of conventional arms as defined therein if it is known that the arms or items would be used in, most relevantly, 'the commission of genocide, crimes against humanity, grave breaches of the 1949 Geneva Conventions, attacks directed against civilian objects or civilians protected as such, or other war crimes as defined by international agreements to which it is a Party'.[45] In any event, States Parties are required to assess the potential that the arms or items could be used to commit or facilitate a serious violation of international humanitarian law.[46] Thus, these measures are considered as instrumental to the implementation of the general obligation to ensure respect for international humanitarian law. Indeed, the ICRC's 2016 Commentary to Geneva Convention I makes reference to arms transfer as an example of the negative duty under Common Article 1 of the Geneva Conventions, observing that 'Common Article 1 requires High Contracting Parties to refrain from transferring weapons if there is an expectation, based on facts or knowledge of past patterns, that such weapons would be used to violate the Conventions'.[47]

---

[44] *Nicaragua* judgment, above n 24, 135 para 269 (observing that 'in international law there are no rules, other than such rules as may be accepted by the State concerned, by treaty or otherwise, whereby the level of armaments of a sovereign State can be limited').

[45] Arms Trade Treaty, adopted 2 April 2013, 3012 UNTS (entered into force 24 December 2014) art 6.

[46] Ibid, art 7.

[47] ICRC, above n 28, para 162.

However, there are inherent complexities in the assessment of risks that exported items in question might be used in the commission of a serious violation of international humanitarian law as the basis for decisions to prohibit or suspend arms export to a third country.[48] The difficulty is illustrated in the dispute regarding the legality of the arms export by the United Kingdom (UK) to Saudi Arabia due to allegations of grave breaches of international humanitarian law. In its decision to dismiss the challenge, the UK High Court justified its deferential approach to the Executive's export decision as follows:

> The picture was acknowledged to be far from a black and white. The decision involved balancing a series of complex and competing factors. Such self-evidently finely balanced judgements are paradigm matters for evaluation and decision by the Executive in conformity with the scheme established by Parliament.[49]

By contrast, while acknowledging that export licensing decisions for the sale of arms always require fine judgements, the Select Committee on International Relations in the UK House of Loads observed that 'given the volume and type of arms being exported to the Saudi-led coalition, we believe they are highly likely to be the cause of significant civilian casualties in Yemen, risking the contravention of international humanitarian law'.[50]

In addition, the range of weapons regulated under the Arms Trade Treaty is limited to certain categories of conventional arms only.[51] This means that dual-use technologies, such as AI, are not subject to the arms control obligations under the Treaty unless they form part of the conventional arms. On the other hand, the Wassenaar Arrangement addresses broader coverage of items, as an international forum established in 1995 to promote

---

[48] See Maya Brehm, 'The Arms Trade and States' Duty to Ensure Respect for Humanitarian and Human Rights Law' (2008) 12 *Journal of Conflict & Security Law* 359, 375-9.

[49] *R (Campaign Against Arms Trade) v Secretary of State for International Trade* [2017] EWHC 1726, para 209. The Court of Appeal subsequently deviated from the High Court in its judgment by finding that the necessary precondition for making a rational decision was missing due to the absence of definitive legal assessment of Saudi Arabia's past conduct during the armed conflict: [2019] EWCA Civ 1020, paras 138-145.

[50] UK House of Lords Select Committee on International Relations, 'Yemen: Giving Peace a Chance' (HL Paper 290, 16 February 2019) para 73.

[51] Arms Trade Treaty, above n 45, art 2.

transparency and greater responsibility in transfers of sensitive dual-use goods and technologies, as well as conventional arms.[52]

While legally non-binding, the Wassenaar Arrangement has established a cooperative regime to facilitate information exchange for enhanced transparency, coordination of national control policies, and commitment to national control of all items set forth in the List of Dual-Use Goods and Technologies (including Sensitive List and Very Sensitive List), as well as those on the Munitions List. The Lists are reviewed regularly to reflect technological developments and on-going studies of experience gained in the field of dual-use goods and technologies.[53]

Regulatory reach of those participating States already extends to several types of computer algorithms of certain function which, as of December 2018, include:

- signal processing algorithms used for, inter alia, acoustic-optic devices, [54] digital data recorders, [55] and radar systems, equipment and assembles specifically designed for airborne operation;[56]
- image enhancement algorithms for the processing of externally derived information-bearing images;[57]
- real-time algorithms to analyse component sensor information to predict and pre-emptively mitigate impending degradation and failures of components within an 'active flight control system';[58] and
- real-time algorithms to identify component failures and reconfigure force and moment controls to mitigate degradations and failures of 'active flight controls system'.[59]

It is reasonable to expect that any future improvement of these algorithms with the benefit of machine learning will remain subject to the exports control regime within the States

---

[52] Wassenaar Arrangement on Export Controls for Conventional Arms and Dual-Use Goods and Technologies, signed 19 December 1995, Section I.
[53] Ibid, Section III para 3.
[54] Wassenaar Arrangement: List of Dual-Use Goods and Technologies (as of December 2018) §3-A-1 para c.3.
[55] Ibid, §3-A-2.
[56] Ibid, §6-A-8 paras g and k.
[57] Ibid, §4-A-3.
[58] Ibid, §7-E-4 para b.3.
[59] Ibid, §7-E-4 para b.4.

participating in the Wassenaar Arrangement. With respect to other computer algorithms whose intended function is not listed, there is a shared understanding that exporters should be required to inform national regulatory authorities if they are aware that non-listed dual-use items in question are intended, entirely or in part, for a military end-use.[60]

National authorities also retain the right to adopt and implement regulatory measures to restrict exports for other reasons of public policy, such as national security. For example, the export of dual-use technology from the United States is subject to regulatory control under the Export Controls Reform Act 2018, which repealed and replaced the Export Administration Act 1979 with a view to extending regulatory reach to emerging and foundational technologies that are essential to its national security.[61] In the European Union, Council Regulation (EC) No 428/2009 has established a regulatory regime for the control of exports, transfer, brokering and transit of dual-use items, with the adoption of a common list based on the Wassenaar Arrangement's List of Dual-Use Goods and Technologies.[62] In Australia, the Wassenaar Arrangement's List is incorporated into the Defence and Strategic Goods List, which provides the basis upon which the Minister may grant or revoke a permit for a specified supply on grounds of the security, defence or international relations of Australia.[63]

However, the national requirement for regulatory control is more likely to be driven by national security interests, rather than as an implementation of the obligation to ensure respect for international humanitarian law. This means that while national security sensitive algorithms may well be subject to regulatory control, other AI-enabled dual-use technologies that do not cause an immediate national security concern could be transferred even if there is a foreseeable risk that these might be used to facilitate breaches of international humanitarian law, such as autonomous vehicles used in the delivery of explosives for detonation

---

[60] Statement of Understanding on Control of Non-Listed Dual-Use Items, adopted at the 2003 Plenary Meeting of the Wassenaar Arrangement, available at https://www.wassenaar.org/app/uploads/2015/07/Non-listed_Dual_Use_Items.pdf.

[61] National Defense Authorization Act for Fiscal Year 2019, Public Law No: 115-232 §1758.

[62] Council Regulation (EC) No 428/2009, OJ L 134 (29 May 2009) 1.

[63] Defence Trade Controls Act 2012 (Cth), as amended by the Defence Trade Controls Amendment Bill 2015, ss 11(4) and (9).

targeting civilians in a densely populated area. It is particularly challenging, in this respect, to regulate the transfer of computer algorithms due to the inherent complexities in export control of intangible transfer, which necessarily involves a great deal of uncertainty as to the range of end-use options and the likelihood of intended use.

4. Concluding Observations

The implementation of the general obligation to respect and ensure respect for international humanitarian law poses practical challenges with respect to the study, development, acquisition, adoption, use or transfer of AI technology. Weapons review, as a means of implementing that obligation, is limited in scope to the AI that is incorporated or integrated into a particular application to be used, designed or intended to be used to cause injury or death to a person, or destruction or damage to an object. Export control, on the other hand, may extend to dual-use technologies, such as AI, yet the regulatory control tends to be based on national security concerns, rather than compliance with international humanitarian law. This appears to leave a critical gap in the implementation of the positive duty to ensure that the use of AI does not involve any breach of international humanitarian law, especially when it is transferred to third parties involved in an armed conflict.

The gap exists where AI is studied, developed, acquired or adopted without being designed or intended to be used to cause injury or death to a person, or destruction or damage to an object, but is capable of application that facilitates a serious violation of international humanitarian law. This problem is relevant to many dual-use technologies, but may exacerbate with the rise of AI due to the potentially wider variety of unintended applications for hostile purposes. The data-driven probabilistic reasoning unique to machine learning as one of the critical techniques that drive the current progress for AI means that the same computer algorithm can be employed in a variety of ways by feeding a different set of data depending on the purpose of application.

Although there is no easy solution to this issue, the Wassenaar Arrangement's Participating States recognise the importance of post-export monitoring with reference, for example, to 'the imposition of a requirement on industry, academia, and

individuals to keep records, for an appropriate period of time, that clearly identify all controlled technology transferred, the dates between which it was transferred, and the identity of the end-user of all intangible transfers of technology for which licenses have been issued that may be inspected by, or otherwise provided to, export control authorities upon request'.[64] In cases where it is feasible, post-export monitoring is a practical measure that can be usefully employed to mitigate these concerns in implementing the general obligation to ensure respect for international humanitarian law.

---

[64] Best Practices for Implementing Intangible Transfers of Technology Controls, adopted by the 2006 Plenary Meeting of Wassenaar Arrangement.