Journal of Conflict & Security Law © Oxford University Press 2022.

https://doi.org/10.1093/jcsl/krac035
Advance Access Publication Date: 18 November 2022

The Utility of Weapons Reviews in Addressing Concerns Raised by Autonomous Weapon Systems

Damian Copeland,* Rain Liivoja† and Lauren Sanders‡

Abstract

The obligation to legally review weapons, means and methods of warfare has been identified by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems as one of its Guiding Principles. Despite calls to share practical measures and processes to undertake this review, national practice remains opaque and fragmented. This article describes the traditional weapons review process and explains why this process may need to be modified to adequately evaluate autonomous weapon systems (AWS). It uses three case studies of fictional AWS in various stages of development and acquisition to demonstrate how existing review processes can be adapted for the review of AWS. This article shows the utility of these reviews for ensuring compliance of AWS with existing legal requirements, thereby also demonstrating the suitability of existing law to regulate the use of this novel technology in warfare.

1. Introduction

An international debate about the regulation and use of autonomous weapon systems (AWS) has been ongoing for almost a decade without even reaching

* School of Law, University of Queensland, Brisbane, Australia. E-mail: damian.copeland@uq.edu.au.
† School of Law, University of Queensland, Brisbane, Australia; Lieber Institute for Law and Land Warfare, United States Military Academy, West Point, USA. E-mail: r.liivoja@uq.edu.au.
‡ School of Law, University of Queensland. E-mail: l.sanders@uq.edu.au.
The authors would like to thank Rosie Cavdarski for her helpful research assistance. The views and opinions expressed in the article are those of the authors, and do not reflect the views of any government, any government department or agency, or any other organisation or institution.

consensus on definitions.[1] But it would be unfair to say that there has been no progress whatsoever within the Group of Government Experts (GGE) in the Area of Lethal Autonomous Weapons Systems, convened under the auspices of the Convention on Certain Conventional Weapons (CCW).[2] Earlier proposals to create new treaty law to simply ban this technology outright have been replaced by more nuanced initiatives. At the same time, many States have abandoned their initial resistance to any new normative instrument being developed. In recent GGE discussions, which draw upon 11 Guiding Principles that have previously been adopted by consensus,[3] two broad options have begun to crystallise. One involves a legally binding instrument that would include some combination of prohibitions and restrictions on the use of particular categories of AWS, potentially by articulating standards regarding human control.[4] The other option would be a compendium of relevant principles, best practices and/or interpretations of existing law, without creating new legal obligations.[5]

[1] For the purposes of this article, AWS are weapon systems that select and engage targets without real-time human intervention. This definition aligns with those used by the US Department of Defense and the International Committee of the Red Cross (ICRC). See US Department of Defense, Directive 3000.09: Autonomy in Weapon Systems (21 November 2012, Incorporating Change 1, 8 May 2017) (defining an AWS as '[a] weapon system that, once activated, can select and engage targets without further intervention by a human operator'); International Committee of the Red Cross, 'International Committee of the Red Cross (ICRC) Position on Autonomous Weapon Systems: ICRC Position and Background Paper' (2020) 102(915) IRRC 1335, 1339 ('The ICRC understands AWS to be weapons that select and apply force to targets without human intervention.').

[2] 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 (adopted 10 October 1980, entered into force 2 December 1983) 1342 UNTS 137 (CCW).

[3] Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, Annex III to the Final Report of the Meeting 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 (13 December 2019) UN Doc CCW/MSP/2019/9 (Guiding Principles).

[4] Argentina, Costa Rica, Ecuador, Guatemala, Kazakhstan, Nigeria, Panama, Peru, the Philippines, Sierra Leone, State of Palestine and Uruguay, 'Written Commentary Calling for a Legally-Binding Instrument on Autonomous Weapons Systems' (GGE LAWS, March 2022) <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/G12_March2022.pdf>; Argentina, Costa Rica, Guatemala, Kazakhstan, Nigeria, Panama, Philippines, Sierra Leone, State of Palestine and Uruguay, 'Proposal: Roadmap Towards New Protocol on Autonomous Weapons Systems' (GGE LAWS, March 2022) <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/G13_March2022.pdf>.

[5] Australia, Canada, Japan, the Republic of Korea, UK and USA, 'Principles and Good Practices on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems' (GGE LAWS, 7 March 2022) <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/USgroup_March2022.pdf>; see also United Kingdom, 'Proposal for a GGE Document on the Application of International Humanitarian Law to Emerging Technologies in the Area of Lethal

The prospects of further progress under the auspices of the CCW are, however, uncertain. The GGE works on the basis of consensus, allowing a single State to stall or disrupt discussions. The Russian Federation did so extensively in 2022.[6] It is particularly troubling that these attempts at derailing the GGE process coincide with increasing rumour and reports of the use of uncrewed aerial vehicles with at least some autonomous functionality in Libya, Nagorno-Karabakh, Gaza and Ukraine.[7] Some of these do not appear to be capable of being used in compliance with basic international humanitarian law (IHL) principles, such as accurately distinguishing between legitimate military targets and protected objects and persons.[8]

Despite the uncertainty about the future direction of the GGE debate, the Guiding Principles need operationalisation regardless of the chosen regulatory pathway. The review of AWS for legal compliance, referred to here as 'weapons review', is essential to identifying whether the system may lawfully be employed.[9] Significantly, the current proposals for a legally binding instrument and for a statement of good practices both recognise the importance of weapons reviews.[10]

Autonomous Weapon Systems (LAWS)' (GGE LAWS, March 2022) <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/UK_March2022.pdf>.

6 Sanctions imposed against the Russian Federation in response to its unlawful use of force against Ukraine gave Russia the pretext to object to formal GGE meetings in March 2022. Russia claimed that these 'discriminatory' measures were making it impossible for its delegation to fully participate in the GGE meetings. This proposition was rejected by many other States as being false. See 2022 LAWS GGE, First Session, 7–11 March 2022, Daily Transcripts <https://indico.un.org/event/37347/page/0>, especially transcripts of meetings on 8 and 9 March. In July 2022, Russia objected to the presence of civil society representatives in informal sessions during the drafting of the GGE's report, leading to a full day of 'consultations' with a select group of participants behind closed doors.

7 'Final Report of the Panel of Experts on Libya Established Pursuant to Security Council Resolution 1973 (2011)' (8 March 2021) UN Doc S/2021/229; H Nasu, 'The Alpagu Autonomous Attack Drone: What the Future holds for the Humanity in the Battlefield?' (*Lawfire*, 1 July 2021) <https://sites.duke.edu/lawfire/2021/07/01/hitoshi-nasu-on-the-alpagu-autonomous-attack-drone-what-the-future-holds-for-the-humanity-in-the-battlefield/>; H Nasu, 'The Kargu-2 Autonomous Attack Drone: Legal & Ethical Dimensions' (*Articles of War*, 10 June 2021) <https://lieber.westpoint.edu/kargu-2-autonomous-attack-drone-legal-ethical/>; C Conboy, 'Autonomous Weapon Systems Used in Ukraine' (*Stop Killer Robots*, 22 March 2022 <www.stopkillerrobots.org/news/autonomous-weapon-systems-used-in-ukraine/>; H Nasu, 'Hunter 2-S Swarming Attack Drones: Legal & Ethical Dimensions' (*Articles of War*, 31 March 2022) <https://lieber.westpoint.edu/hunter-2-s-swarming-attack-drones-legal-ethical-dimensions/>.

8 See Nasu, 'The Alpagu Autonomous Attack Drone' (n 7) for a discussion on the limitations of use of autonomous systems to detect and recognise targets without human intervention.

9 Guiding Principles (n 3) [(e)].

10 'Roadmap towards a New Protocol' (n 4) [24]; 'Principles and Good Practices' (n 4) [21]–[24].

This article seeks to explain why weapons reviews will be so central to the responsible use of AWS, how such reviews should be carried out and how they will benefit the reviewing State.

After briefly tracing the legal basis for weapons reviews, and how it will apply to AWS, Section 2 outlines views on the utility of weapons reviews for AWS. Section 3 describes the commonalities in current State practice for the weapons review process generally (the traditional process) and the need to adjust this process in light of the inclusion of autonomous functionalities in new weapons, methods or means of warfare.

To make the discussion more practical, Section 4 offers a selection of case studies to help show the general utility of weapons reviews for AWS at different stages of their development, acquisition and use. The first case study focuses on the State-sponsored 'study and development' of a piece of dual-use technology with military application. The second case study considers the reviewing State's 'acquisition or adoption' of a mature AWS from an allied State. The final case study evaluates the enhancing of an existing weapon system with artificial intelligence (AI).

Without taking a position on whether additional regulation of AWS is necessary or desirable, this article seeks to demonstrate the usefulness of the weapons review process in the context of weapon systems that have autonomous functionality. The article will show how the extant legal framework can be leveraged and the review obligation adapted to ascertain whether a system with autonomous capability can be used lawfully. It further seeks to add to the existing literature by providing examples of how the Guiding Principle affirming the weapons review obligation can be translated into practice.

2. Development of Weapons Review Processes and Views on Applicability to AWS

The weapons review process is not specific to novel technology. Rather, it derives from general and specific international legal obligations that require consideration of the way in which all 'tools of war' are utilised.[11] After discussing the source of this obligation, this Section identifies the arguments made in support of the relevance of weapons reviews to AWS, outlines the sceptics' arguments, and lists the key challenges raised in reliance upon weapons reviews to underwrite the use of AWS in armed conflict.

[11] See generally, ICRC, 'A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measure to Implement Article 36 of Additional Protocol 1 of 1977' (2006) 88(864) IRRC 931 (ICRC Guide).

A. Weapons Review Obligation

The right of belligerents to choose means and methods of warfare is not unlimited.[12] International law restricts that choice in multiple ways. On the one hand, it prohibits generally weapons, methods and means of warfare of a nature to cause certain types of harm.[13] On the other hand, it prohibits and limits the use of specific weapons, means and methods of warfare.[14]

Additional Protocol I expressly requires its States Party to review any weapon, means or method of warfare, ahead of its use, for compliance with that State's international legal obligations. Specifically, Article 36 of Additional Protocol I provides as follows:

> In the study, development, acquisition or adoption of a new weapon, means or method of war, 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.

The aim of this provision is to 'prevent the use of weapons that would violate international law in all circumstances, and to impose restrictions on the use of weapons that would violate international law in some circumstances'.[15]

The customary international law status of this obligation remains unsettled.[16] It is not in doubt, however, that States must ensure that the use of force in armed conflict complies with relevant international law, including any limitations that apply to the choice of weapons, means and methods of warfare. We suggest that understanding how a weapon, means or method of warfare operates, having regard to legal obligations and principles, is a logical and necessary prerequisite for a State fielding it for lawful use in armed

[12] Hague Convention (II) with Respect to the Laws and Customs of War on Land (adopted 29 September 1899, entered into force 4 September 1900) Annex: Regulations concerning the Laws and Customs of War on Land art 22; Hague Convention (IV) respecting the Laws and Customs of War on Land (adopted 18 October 1907, entered into force 26 January 1910) Annex: Regulations concerning the Laws and Customs of War on Land art 22; of the 1907 Hague Regulations Respecting the Laws and Customs of War on Land art 22; Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (opened for signature 8 June 1977, entered into force 7 December 1978) 1125 UNTS 3, art 35(1) (Additional Protocol I or API).

[13] See API (n 12) art 35(2) and (3), and art 51(4); see also material cited in ICRC, *Customary IHL Database*, practice to rules 45, 70 and 71 <https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2>.

[14] See material cited in ICRC, *Customary IHL Database*, practice to rules 46–69, and 72–86, <https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2>.

[15] ICRC Guide (n 11) 4.

[16] See N Jevglevskaja, *International Law and Weapons Review* (CUP 2022) for an outline of current State practice and an assessment as to whether weapons review obligations form part of customary international law.

conflict. Accordingly, States not party to Additional Protocol I, including the USA and Israel, are subject to a more general duty to determine the legality of their weapons, means and methods of warfare as a consequence of their obligation to comply with their international law obligations in good faith (*pacta sunt servanda*).[17]

Additional Protocol I does not define the terms 'weapon', 'means of warfare' or 'method of warfare'. This has allowed States and commentators to take divergent approaches to delimiting these concepts and establishing their relationship with each other.[18] As a general matter, we consider AWS, as they feature in ongoing international debates, as capable of constituting weapons, means of warfare or both. In this article, we address the design and development of AWS for specific use cases, and touch upon methods of warfare in that context. Thus, we leave to one side the broader question as to whether autonomous functionality could amount to a distinct method of warfare.

B. Views on the Weapons Reviews of AWS

The GGE has concluded that weapons reviews 'at the national level' are 'a useful tool to assess whether AWS would be prohibited by any rule of international law applicable to that State in all or some circumstances'.[19] This view aligns with the Guiding Principles, adopted by the GGE and endorsed by the Meeting of High Contracting Parties to the CCW, which relevantly specify:

> In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means or method of warfare, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law.[20]

Although there has been formal consensus on the relevance of weapons reviews to determining the legality of AWS, there remain different views as to the practical utility of these reviews in governing the use and deployment of AWS.

[17] Vienna Convention on the Law of Treaties (opened for signature 23 May 1969, entered into force 27 January 1980) 155 UNTS 331 art 26; see also ICRC Guide (n 11) 933.

[18] See J McClelland, 'The Review of Weapons in Accordance with Article 36 of Additional Protocol I' (2003) 85 IRRC 397, 404; Jevglevskaja (n 16) 88–115.

[19] 'Report of the 2019 Session of the Group of Government Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems' (Geneva 25–29 March 2019 and 20–21 August 2019) (25 September 2019) UN Doc CCW/GGE.1/2019/3.

[20] Guiding Principles (n 3).

Many States emphasise the importance of conducting weapons reviews to assess the compliance of AWS with existing international legal obligations, limitations, restrictions or prohibitions.[21] A number of States also support the discussion and dissemination of relevant State practice. This group notably includes many of the States that do not support the call for a legally binding instrument on AWS and argue that existing international law, in particular IHL, is sufficient to regulate the use of AWS.[22]

Weapons reviews have been discussed as a tool to allay concerns that a particular type of weapon may pose for legal compliance. Reviewing how the weapon might be deployed in a specified operational context can lead to the imposition of limitations or restrictions upon the use of that weapon to further compliance with IHL. A weapons review thus enables identification of those types of weapons that cannot comply with legal obligations, or those that may only meet legal standards in particular use scenarios. For example, depending on views adopted relating to levels of human control and agency necessary to enable compliance with legal obligations, the weapon system's ability to meet the legal criteria of distinction and proportionality could be assessed. While this goes beyond the usual scope of a weapons review, it is a practical necessity to support the proper and lawful fielding of these capabilities in armed conflict by a State in compliance with its international legal obligations.

Many of those supporting specific regulation of AWS—but by no means all of them—remain sceptical about the utility or adequacy of weapons reviews as a pathway to AWS legal compliance.[23] Despite the Article 36 weapons review obligation having existed for four decades, there is limited evidence of national

[21] 'Chairman's Summary of the 2020 Session of the Group of Government Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems' (19 April 2021) UN Doc CCW/GGE.1/2020/WP.7, 30 (Austria), 35 (Cuba), 41 (Finland), 44 (France), 48 (Germany), 50 (Guatemala), 52 (Israel), 56 (Japan), 61 (Netherlands), 64 (Panama), 68 (Poland), 72 (Portugal), 77 (Sweden), 88 (Switzerland).

[22] See, eg, Australia, 'Lethal Autonomous Weapons Systems: National Commentary' (GGE LAWS, 20 August 2020) <https://documents.unoda.org/wp-content/uploads/2020/08/20200820-Australia.pdf>; Israel, 'Considerations on the Operationalization of the Eleven Guiding Principles Adopted by the Group of Government Experts' (GGE LAWS, 31 August 2020) <https://documents.unoda.org/wp-content/uploads/2020/09/20200831-Israel.pdf>; United States, 'Commentaries on the Guiding Principles' (GGE LAWS, 1 September 2020) <https://documents.unoda.org/wp-content/uploads/2020/09/20200901-United-States.pdf>.

[23] See, for example, Human Rights Watch, *Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban* (2016) 33 <www.hrw.org/sites/default/files/report_pdf/arms1216_web.pdf>; N Goussac, 'Safety Net or Tangled Web: Legal Reviews of AI in Weapons and War-fighting' (*Humanitarian Law & Policy*, 18 April 2019) <https://blogs.icrc.org/law-and-policy/2019/04/18/safety-net-tangled-web-legal-reviews-ai-weapons-war-fighting/>. Many have also highlighted the challenges of reviewing AWS. See, eg, Austria, 'Contribution of Austria to the Chair's Request on the Guiding Principles on Emerging Technologies in the Area of LAWS' (GGE LAWS, 1 September 2020) 3–4 <https://documents.unoda.org/wp-content/uploads/2020/09/20200901-Austria.pdf>; Switzerland, 'Commentary on Operationalizing the Guiding Principles at a National Level' (GGE LAWS, 25

weapons review practice for conventional weapons.[24] In addition to its limited take up by States generally, the process for conducting a review is not prescribed by international law and therefore varies significantly between States.[25] The absence of uniform practice of weapons reviews undermines arguments that weapons review will prevent the use of unlawful AWS.

Noting that the weapons review process also documents the normal and expected use of a weapon, means or method of warfare, it provides a reference point—and guidance to the military—as to how the reviewing State considers that a particular AWS may be used lawfully. Sceptics of weapons review processes suggest that the evolving nature of some AWS, for example through machine learning, renders the articulation of the use case practically impossible. If a State reviews the weapon for use in one context, but the machine changes itself in a material way, that initial assessment of legal compliance becomes meaningless.

C. Challenges for Weapons Review of AWS

While we aim to show that a weapons review is not only legally necessary but practically useful to assess the potential use of an AWS to meet legal obligations, we note that the sceptics have correctly identified two challenges. The first has to do with the limited accessibility of State practice and the resultant uncertainty around how exactly States conduct weapons reviews. The second challenge is that some novel features of technology that facilitate autonomous functionality may render existing weapons review processes inadequate.

(i) Sharing of Inconsistent State Practice

State practice in conducting reviews is far from consistent, is conducted in private and without public explanation of its results.[26] The internal nature of weapons reviews raises the concern of a divergence of State interpretation and application of legal norms to AWS. In a submission to the GGE, Austria suggests that States applying different legal standards to weapons reviews would lead to inconsistency in results and ultimately uncertainty.[27] This risk was identified already during the

August 2020) <https://documents.unoda.org/wp-content/uploads/2020/08/20200825-Switzerland.pdf>.

[24] For the (fairly short) list of States known to undertake weapons reviews, see N Jevglevskaja and R Liivoja, 'Weapons Review' (*PREMT*: *Program on the Regulation of Emerging Military Technologies*, 11 March 2022) <www.premt.net/weapons-review/>.

[25] Jevglevskaja (n 16) 122–62.

[26] WH Parks, 'Convention Weapons and Weapons Reviews' (2005) 8 *YBIHL* 55, 135.

[27] Austria (n 23).

Diplomatic Conference responsible for drafting the 1977 Additional Protocols but it was agreed that a determination of legality by a State was not intended to create a new subjective standard.[28] A potential for inconsistent weapons review determinations must be balanced against the reality that few States, for reasons of national security and commercial proprietary interests, would be prepared to publicise the results of their weapons reviews. There is no transparency obligation associated with Article 36.[29]

Noting the limited State practice and internal nature of weapons review, numerous States and non-government organisations have called for sharing of weapons review procedures and best practices.[30] Some States have called for a 'harmonization of weapons review processes and an elaboration of international agreed norms and standards'.[31] To aid operationalising the weapons review obligation, France recommends that States 'exchange information on their national legal review procedures as well as to jointly define best practices'.[32] Similarly, the Netherlands believes that 'sharing information regarding the modus operandi and underlying principles of national Article 36 API procedures would be of added value'.[33] This might result in a compendium of good national practice on weapons reviews and contribute to transparency and information exchange. Japan goes further by recommending 'the introduction of an implementation mechanism of weapons review into the annual report of the CCW as a confidence building measure'.[34]

[28] Jean de Preux, 'Protocol I – Article 36 – New Weapons' in Y Sandoz, C Swinarski and B Zimmerman (eds), *Commentary on the Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949* (ICRC and Martinus Nijhoff 1987) [1469].

[29] Parks (n 26) 135.

[30] See, eg, 'Legal Review of New Weapons: Scope of the Obligation and Best Practices' (*Humanitarian Law & Policy*, 6 October 2016) <https://blogs.icrc.org/law-and-policy/2016/10/06/legal-review-new-weapons/>.

[31] Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (23 October 2018) UN Doc CCW/GGE.1/2018/3, [44].

[32] France, 'Operationalization of the 11 Guiding Principles at National Level' (GGE LAWS, 6 August 2020) <https://documents.unoda.org/wp-content/uploads/2020/07/20200610-France.pdf>.

[33] Netherlands, 'National Commentary by the Kingdom of the Netherlands Regarding the National Interpretation and Implementation of the Guiding Principles Affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System' (GGE LAWS, 2020) <https://documents.unoda.org/wp-content/uploads/2020/09/NL-Comments-LAWS-guiding-principles-formatted.pdf>.

[34] Japan, 'Commentary on the Operationalization of the Guiding Principles Affirmed by the Group of Government Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems at National Level' (GGE LAWS, 28 August 2020) <https://documents.unoda.org/wp-content/uploads/2020/09/20200828-Japan.pdf>.

(ii) Challenges to Review Processes Posed by Autonomous Technology

The answer to the practical question of when novel technologies associated with weapons, means or methods of warfare are subject to review is unclear. This problem is compounded by the increasing ubiquity of autonomous capabilities in augmenting existing weapon systems, or in assisting in the decision-making process to procure lethal effects, previously the subject of human decision-making. We concur with Klaudia Klonowska in supporting an expansive reading of the review obligation when it comes to AI systems that are not central to the use of force, but result in the 'co-production of hostilities'.[35] Further, the requirement to re-review conventional systems augmented by AI or other novel technologies would be triggered by a material change to its previous functionality as a consequence of the insertion of the new technology.

The temporal scope of the weapons review obligation is also problematic when reviewing AWS. Article 36 requires a weapons review during the 'study, development, acquisition or adoption' which occur prior to the use of a weapon in armed conflict. Except in the case of a material modification, weapons reviews do not require reconsideration during a weapon's in-service life, which may be decades long. This represents a risk for AWS that are capable of adaptation, for example through machine learning. While not articulated in the Guiding Principles,[36] the need for weapons review to occur throughout an AWS life cycle has been recognised by a number of States, including Austria, Belgium, Brazil, Chile, Ireland, Germany, Luxembourg, Mexico and New Zealand. They argue that 'a regular evaluation process ... should be applied across the life cycle of a weapon system'.[37] Similarly, Australia regards weapons review as a component of a broader 'system of control' that applies throughout the life cycle of an AWS, from its development through to decommission.[38] Overcoming this challenge of determining appropriate timing requires adaption of the review process through an iterative approach.

Finally, given the nature of autonomy subsuming activities that were previously the preserve of a human, assessment of the standards of 'decision-making' undertaken by the machine, vice the human, must be undertaken. It is difficult to translate into code acceptable levels of machine performance having regard to legal standards: comparison of standards relating to human error versus acceptable standards of machine error are not readily capable of testing and

[35] K Klonowska, 'Article 36: Review of AI Decision-Support Systems and Other Emerging Technologies of Warfare' (2020) 23 *YBIHL* 123.

[36] Ibid.

[37] Austria, Belgium, Brazil, Chile, Ireland, Germany, Luxembourg, Mexico and New Zealand, 'Joint "Commentary" on Guiding Principles A, B, C, D' 2020 (GGE LAWS, 1 September 2020) 2–3 <https://documents.unoda.org/wp-content/uploads/2020/09/GGE20200901-Austria-Belgium-Brazil-Chile-Ireland-Germany-Luxembourg-Mexico-and-New-Zealand.pdf>.

[38] Australia, 'Australia's System of Control and Applications for Autonomous Weapon Systems' (26 March 2019) UN Doc CCW/GGE.1/2019/WP.2/Rev.1.

evaluation. For example, the exercise of due diligence in targeting operations, as is required by Article 57 of Additional Protocol I, is difficult to translate into code. Thus, those actions governed by IHL that have been delegated to the machine must be assessed in terms of meeting the requisite legal standards for use. This standard may, as a minimum, be measured against the same action had a human undertaken it. For example, when a decision on target acquisition is made, the standard of due diligence is required of humans; this standard represents the minimum that would be required of a machine undertaking this function on behalf of a human.

Debate about the requirement for, and efficacy of, humans 'in the loop' or 'on the loop' also persist,[39] but the requirement for 'meaningful human control' as mooted in the CCW with respect to AWS—however that may be eventually defined—means that the review must assess whether or not there is an appropriate mechanism for human control of the system.[40]

3. The Traditional Weapons Review Process and Required Amendments for AWS

The inconsistent State practice in weapons review—insofar as it is publicly known—and the absence of any specified process to undertake reviews, other than the articulation of a general obligation to ensure legal compliance, means that processes to properly complete a weapons review can be derived having regard to the construction of Article 36 and the scope of the review obligation. The review of a weapon for lawfulness is, at its most basic, an assessment of compliance with a State's specific and general legal obligations. Taking this requirement and analysing it against current State practice can identify the generally required steps necessary to conduct a weapons review.[41]

A. What is the Weapons Review Obligation?

The Article 36 obligation is intended to ensure that States, before employing a weapon in armed conflict, make determinations about whether a weapon is

[39] K Leins and A Kaspersen, '7 Myths of Using the Term "Human on the Loop"' (*Carnegie Artificial Intelligence and Quality Initiative*, 9 November 2021) <www.carnegieaie.org/blog/7-myths-of-using-the-term-human-on-the-loop/>; J Barnett, 'AI Needs Humans "On the Loop" Not "In the Loop" for Nuke Detection, General Says' (*Fedscoop*, 14 February 2020) <www.fedscoop.com/ai-should-have-human-on-the-loop-not-in-the-loop-when-it-comes-to-nuke-detection-general-says/>.

[40] See further T McFarland, 'Minimum Levels of Human Intervention in Autonomous Attacks' (2022) 27 *JCSL* (advance access) <https://doi.org/10.1093/jcsl/krac021>.

[41] L Sanders and D Copeland, 'Developing an Approach to the Legal Review of Autonomous Weapon Systems' (*ILA Reporter*, 27 November 2020) <https://ilareporter.org.au/2020/11/developing-an-approach-to-the-legal-review-of-autonomous-weapon-systems-lauren-sanders-and-damian-copeland/>.

prohibited or restricted as a matter of international law. The determination is based upon the normal or expected use of the weapon under some or all circumstances.[42] Thus, the review deals with the weapon as it is presented to the reviewer and, in its traditional form, contemplates weaponry as static capabilities, where testing and adoption methodologies relate to fixed means and methods of warfare with clearly anticipated and defined use cases. National weapons review determinations are not binding on other States, and are not intended to create a separate legal standard, but are rather intended to 'ensure that means or methods of warfare will not be adopted without the issue of legality being explored with care'.[43]

A weapons review will focus primarily on the legality of a weapon *per se* rather than the legality of its use in particular circumstances.[44] It is generally accepted that determining the lawful use of a lawful weapon depends on the context and the responsibility for making that determination on the basis of IHL rules rests primarily with military commanders, weapon operators and legal advisors made available to commanders, at the appropriate level.[45]

B. A Traditional Weapons Review Process

While weapons review methodologies may vary between States, the process adopted to enable the practical assessment of the legality of a weapon commonly entails four distinct steps.[46] The first, preliminary step is to confirm the applicability of Article 36 by assessing whether there is 'new weapon, means or method of warfare'.[47] As already noted, Additional Protocol I does not define these concepts.[48] States are therefore free to develop their own interpretations through domestic law and internal policies.[49] The normal or expected use of the weapon, anticipated at the time of the review, limits the scope of the analysis.

The second step involves determining whether the type of weapon is specifically prohibited or restricted by treaty or customary law binding the reviewing State.[50] This focuses on the type of weapon or the mechanism for causing harm.

42 de Preux (n 28) [1469].
43 Ibid.
44 WH Boothby, *Weapons and the Laws of Armed Conflict* (OUP 2009) 62.
45 J Farrant and C Ford, 'Autonomous Weapons and Weapons Reviews: The UK Second International Weapons Review Forum' (2017) 93 *Int L Stud* 389, 409; ICRC Guide (n 11) 15; API (n 12) art 82.
46 See generally ICRC Guide (n 11).
47 See ICRC Guide (n 11) 936–38; Netherlands and Switzerland, 'Weapons Review Mechanisms' (7 November 2017) UN Doc CCW/GGE.1/2017/WP.5, [3(c)].
48 See n 18 and accompanying text above.
49 See, eg, Danish Ministry of Defence, *Military Manual on International Law Relevant to Danish Armed Forces in Internationl Operations* (2016) 379–81.
50 ICRC Guide (n 11) 938–42; Netherlands and Switzerland (n 47) [15]; see also 'Chairman's Summary Report' (n 21) 51 (Israel).

For example, if the weapon employs asphyxiating gas to cause injury or death to combatants, it is prohibited by the Geneva Gas Protocol, the Chemical Weapons Convention and customary international law.[51]

The third step considers whether there are any general international law prohibitions applicable to the weapon.[52] This includes consideration of whether the normal or expected use causes unnecessary suffering or superfluous injury to enemy combatants,[53] is inherently indiscriminate in that it cannot be directed at a specific military objective[54] or causes widespread, long-term and severe environmental damage.[55]

For some States, the final weapons review step is quasi-legal, in that it includes consideration of national policy.[56] Drawing on the Martens clause,[57] this step considers domestic policy reflecting the 'principles of humanity' and 'the dictates of public conscience'. While not determinative of a weapon's legality, States may make policy decisions on relevant issues, such as a political decision not to use a particular type of weapon (for example, depleted uranium munitions[58]) or requirements for human control over AWS.[59]

In this article, we call this the 'traditional weapons review process', in that it reflects the currently prevailing interpretation of the obligations under Article 36.[60]

[51] Geneva Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare (adopted 17 June 1925, entered into force 8 February 1928); Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons on their Destruction (adopted 13 January 1993, entered into force 29 April 1997) 1974 UNTS 45.

[52] See ICRC Guide (n 11) 942–45.

[53] API (n 12) art 35(2).

[54] ibid art 51(4)(a).

[55] ibid art 35(3).

[56] See, eg, Australian Department of Defence, 'Defence Article 36 Reviews of New Weapons Guide' (May 2020) <www.defence.gov.au/sites/default/files/foi/187_2021_Document.pdf>; see also SJ White, 'The Legal Review of Weapons, Means and Methods of Warfare' (Presentation, Department of Defence, 2020) <https://documents.unoda.org/wp-content/uploads/2020/10/Weapons-reviews-2020-Australia.pdf>; US Department of Army, Regulation 27-53, Legal Services: Review of Legality of Weapons under International Law (rev 1, 23 September 2019) [1] (Purpose).

[57] The Martens clause first appeared in the preamble to the 1899 Hague Convention II (n 12) and was subsequently reaffirmed in 1907 Hague Convention IV (n 12) and API (n 12) art 1(2); see also 'Chairman's Summary Report' (n 21) 72 (Portugal).

[58] For example, the Australian Government decided not to continue to use depleted uranium munitions. See Commonwealth of Australia, Parliamentary Debates, Senate, 15 October 2003, 16585 <https://parlinfo.aph.gov.au/parlInfo/search/display/display.w3p;query=Id%3A%22chamber%2Fhansards%2F2003-10-15%2F0195%22>.

[59] US Department of Defense (n 1) 2 ('Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force.').

[60] See, eg, the approach described in the ICRC Guide (n 11).

C. Domestic Implementation

The practice of weapons reviews is not universal amongst States Party to Additional Protocol I. Around 20 States are known to have formal legal review processes for traditional, non-AI-enhanced, weapons.[61] At the time of writing, no State has publicly outlined its approach to the legal review of AWS beyond well-intentioned public statements of the requirement to do so.[62] Article 36 implicitly requires States Party to Additional Protocol I to 'establish internal procedures for the purpose of elucidating the issue of legality'.[63] A reviewing State is not required, as a matter of law, to adopt a particular methodology.[64] States generally implement their internal procedures through domestic legislation, regulations or policy directives within the respective defence organisations.

National weapons review policies commonly include the following elements: assigning responsibility for the conduct of reviews to lawyers within the national defence organisation; defining the scope of the weapons review obligation by defining the Article 36 phrase 'new weapon, means or method of warfare'; aligning the weapons review obligation to the military weapon procurement process and identifying the weapon criteria addressed in the steps described above.[65]

Except for the US Department of Defense Directive 3000.09, 'Autonomy in Weapon Systems',[66] national policies are concerned with the weapons review of traditional weapons. This either suggests that States believe that existing processes are sufficient to address the challenges of deploying AWS in conflict or that they continue to evaluate their weapons review frameworks, which are not shared as part of an international community of weapons review practice as proposed by the GGE's Guiding Principles.

D. Amendments of Weapons Review Process Required for AWS

The existing review approach is suitable for traditional weapons where the operators—who are subject to international law accountability mechanisms, such as war crimes prosecution for the misuse of weapons they control—are responsible for ensuring that the method of employment of the weapon complies with the law. However, this approach is too narrow for AWS. They require

[61] See Jevglevskaja and Liivoja (n 24).

[62] For example, Australia's weapons review guide does not detail how such review would consider autonomous technologies, only referencing the option for creating 'multi-disciplinary review committee[s]' to review 'highly complex or novel weapon systems'. Australian Department of Defence (n 56) [8(h)].

[63] de Preux (n 28) [1470].

[64] Boothby (n 44) 343.

[65] For a list of national weapons review policies and guidelines, see Jevglevskaja and Liivoja (n 24).

[66] US Department of Defense (n 1).

legal review throughout their life cycles, given the impact of machine learning in some instances, but more generally as a result of the changing nature of algorithmic performance based on programming and varying data inputs.[67]

The algorithms that drive the AWS must generate results that comply with a State's legal obligations. Thus, a thorough legal review should be part of the entire design and procurement process of an AWS, both informing the AWS development and assessing its legality during use. This process may require intervention in State-sponsored research and development activities, and will require the review to occur in an iterative process, as compared to being a final review issued immediately prior to the incorporation of a product or capability into military use.

Thus, a broader, iterative, multidisciplinary and ongoing approach must be factored into extant review processes. This may extend throughout the weapon's life cycle to assess its ability to operate in multiple environments that require the AWS to interpret data which differs from that upon which its performance was initially reviewed. It will also address advances in technology which change the AWS operation.

Small operational changes could be assessed by an operational or field legal review that builds upon the reviews conducted before introduction into service. Such legal reviews must also be flagged with the original reviewing authority to identify whether to trigger a re-review for that particular capability. More significant changes—for example, the introduction of machine learning or improvements in the AI's contextual reasoning capability—may require specific operational limitations to be placed upon them to ensure ongoing legal compliance. Specifically, an assessment about whether or not the system is capable of changing its normal or expected use will determine whether a re-review of the original assessment for legal compliance is required.

Similar processes used in analogous fields—such as in the regulation of autonomous vehicles and ensuring self-driving cars can comply with road rules—may assist.[68] For example, placing limitations on implementing self-learning without prior human approval of software updates or patches may address this requirement.

Furthermore, the review process will require additional consideration of the expected operating environment of the AWS. In the case of traditional reviews, the human operators of the system must evaluate the impacts of the operating environment, such as assessing objects in the battlespace as being civilian or military objects. The weapons review of an AWS must take into account the limitations in a cluttered and complex environment. Accordingly, in addition to determining if there has been a material change to the operation of the system, the review must also consider any material change to the operating environment in which the system is operating. This is particularly important if the data used

[67] This subsection draws on A Tattersall and D Copeland, 'Reviewing Autonomous Cyber Capabilities' in R Liivoja and A Väljataga (eds), *Autonomous Cyber Capabilities Under International Law* (NATO CCDCOE 2021) 240–52.

[68] M Ryan, 'The Future of Transportation: Ethical, Legal, Social and Economic Impacts of Self-driving Vehicles in the Year 2025' (2020) 26(3) *Sci & Eng Ethics* 1185.

to train the AI is focused on a particular legal regime. For example, an AI-enabled system designed to operate in an international armed conflict would require re-review if it were to be deployed in a non-international armed conflict to ensure that any rules coded into the system reflect the changed criteria for assessing targetable status.

Further, assessments of the manner in which these determinations are made by the AWS must also be undertaken. For example, in determining compliance with the law, the assumptions programmed into the AWS must also be legally compliant. For example, a fundamental IHL principle is that in cases of doubt, status is presumed to be civilian.[69] The AWS must also therefore achieve the same kind of result in cases of doubt. In this case, the performance of the AWS in a particular context will form part of the weapons review—which is additional to the content included in a traditional weapons review.

E. Outcomes of a Traditional Weapons Review

A traditional weapons review report will generally reach one of three conclusions regarding the use of the weapon in armed conflict:

i. the normal or expected use is lawful;
ii. the use of the weapon is generally lawful but particular methods or circumstances may render the use unlawful;
iii. the use of the weapon is unlawful in all circumstances.[70]

In reaching these conclusions, the report will provide recommendations and guidance to the reviewing State about what limitations or controls must be put in place once the weapon system is introduced into service.

(i) Use of the Weapon in its Normal or Expected Use is Lawful

A weapons review will determine the legality of a weapon within the context of its normal or expected use. A traditional weapons review focusses on the legality of weapons *per se* and therefore its determination is likely be binary in nature. For example, the review of a particular high-velocity bullet may reveal that it is of a nature to cause superfluous injury or unnecessary suffering in contravention of Article 35(2) of Additional Protocol I if it is used to target enemy combatants. Such use would thus be prohibited. In contrast, the weapons review may determine that the anti-materiel use of the same bullet is lawful as the issue of unnecessary suffering is not relevant.

[69] See API (n 12) arts 50(1) and 52(3).
[70] Sanders and Copeland (n 41).

(ii) The Use of the Weapon is Lawful, but Limitations are Required on its Method of Use

A traditional weapons review commonly recommends that restrictions be placed on the use of a weapon. Such restrictions often ensure the discriminate use of a weapon in a particular circumstance. There are broadly two ways in which a weapon may be indiscriminate. The first relates to the weapon's design: the weapon is unable to be aimed or directed at specific military objective as required by Article 51(4)(a) of Additional Protocol I. Examples of such weapons include long range, unguided missile such as the German V-2 Rockets during World War II or, more recently, the Soviet Scud Missiles used by Iraq against Israel and Saudi Arabia.[71] The second reason relates to the weapon's use: while being technically capable of being directed at a specific military objective, the weapon is used in an indiscriminate way. In such a case, the weapon could be certified as lawful following review, because it is the user of the weapon rather than the weapon itself that must adjust to ensure legal compliance during use. That is, the weapon is capable of being used discriminately, but the operator of the weapon must point it at a particular target. In this case, the limits recommended by the review are based upon the intended use and the onus is on the weapon operator to ensure the weapon is employed lawfully.

Accordingly, in the case of a wide-area affect weapon (for example, an artillery piece or mortar), the review may recommend that the weapon not be used to target military objectives located within the civilian population or close to civilian objects unless the military advantage outweighs the risk of collateral damage.

(iii) The Weapon is Unlawful in all Circumstances

A third outcome of a review may be that the use of the weapon is unlawful in all circumstances, effectively resulting in the State being unable to utilise the weapon without incurring international responsibility for any wrongful damage that ensues. Such a determination would not necessarily preclude the State from possessing that weapon technology, but knowingly equipping itself with such a weapon would be considered to 'violate the spirit and letter of the Protocol'.[72]

If such a determination was the outcome of a weapons review, significant design and development adjustments would need to be made to enable the system to achieve an alternative, legally compliant outcome—being (i) or (ii) above. In the case of a conventional weapon system, these adjustments may be relatively easy to pinpoint insofar as the componentry that results in the non-

[71] See, eg, MN Schmitt, 'War, Technology, and International Humanitarian Law' (HPCR Occasional Paper Series No 4, 2005) 10 <https://hhi.harvard.edu/files/humanitarianinitiative/files/occasionalpaper4.pdf> .

[72] de Preux (n 28) [1471].

compliance is likely to be linked to a specific function. For example, an accurate sensor may replace an original, inaccurate one.

F. Weapons Review Outcomes Applied to AWS

The outcome of the weapons review of an AWS may not fit neatly within one of the three options mentioned above. For example, a weapons review may advise that the weapon component of the AWS is lawful *per se* but recommend restricting certain autonomous functionality that is determined to be, or risks being, unlawful. Alternatively, a weapons review may advise that the use of an AWS would be lawful but that, as a matter of national policy, its use in certain circumstances should be under direct real-time human control. For example, a review may identify that an AWS can lawfully be used to target civilians taking a direct part in hostilities, but national policy precludes AWS from being employed against humans without direct human control. Alternatively, because of the inherent complexity of certain operations, State policy may require AWS to be used to target military objectives only, not lawful human targets. Therefore, a review report may recommend the AWS not be used in asymmetric operations typical of non-international armed conflict where humans are the focus of kinetic targeting.

(i) Use of the AWS in its Normal or Expected Use is Lawful

While a weapons review of an AWS is concerned with the same specific and general IHL prohibitions and restrictions as a traditional weapons review process, its outcome is less likely to be binary in nature. This is because it will also need to consider those IHL rules regulating methods and means of warfare relevant to the AWS function. Accordingly, a weapons review of an AWS which fired the same high-velocity bullet as in Section 3.E(i) above would make the same determinations as above but would also consider the IHL rules concerning distinction, proportionality and precautions. AWS functions that are governed by these rules may be subject to specific limitations as a result of the review. To illustrate, a high velocity anti-material round may be assessed as raising little risk of disproportionate collateral damage and therefore be approved for use in an AWS. However, the weapons review of an AWS with wide area affects may lead to a restriction of its use in combat environments where civilians and civilian objects are present, by only authorising its use under direct human control.

(ii) The Use of the AWS is Lawful, But Limitations are Required on its Method of Use

In contrast to a traditional weapons review, a review of an AWS may make multiple recommendations to address the legality of an AWS's normal or

expected use. Consider the earlier example of a weapon that it is capable of being aimed discriminately but relies upon the weapon's user to aim it at an appropriate target. When the AWS undertakes this targeting function, the review must consider both cases of when a weapon may be indiscriminate—by design or by use. For example, the review may determine that an AWS, designed to identify, select and attack military objectives but incapable of assessing collateral damage, is only lawful for use in an environment devoid of civilians or civilian objects. In other circumstances, the AWS may be restricted to real-time human control. A recommendation of the review may restrict the use of autonomous functionality to operational circumstances where the legality is sufficiently certain.

A weapons review is likely to consider individual functions of an AWS and make recommendations for each function that engages IHL or other legal obligations. An AWS designed to distinguish between civilian objects and military objects may be prevented or restricted from doing so in different operational contexts. For example, a weapons review may reveal that an AWS, designed to attack enemy armoured vehicles with distinctive physical characteristics, experiences false positive recognition results when it encounters certain agricultural equipment. The weapons review may recommend preventing the AWS from operating in rural environments or requiring it to obtain human approval to attack enemy tanks in such environments. If not already programmed to do so, the review may recommend that the AWS be programmed to enable its operator to alter its autonomous functioning.

The review could also recommend restricting the use of an AWS to address policy requirements. For example, national policy may require real-time human control over any AWS attack against enemy combatants. In such a case, a weapons review may identify circumstances where an AWS attack against military objectives would involve an unacceptable risk of harm to nearby civilians—even if such use was potentially lawful.

Similarly, weapons review recommendations to prohibit or restrict specific AWS functionality may be temporary to address technical or operational circumstances. For example, an AWS designed to target specific military objectives may demonstrate unpredictability performing this task in certain operational environments. The weapons review may then lead to limiting autonomous targeting and require direct human decision making, or it may prohibit the weapon entirely from being used in such circumstances, until the reason for the unpredictability is determined and resolved.

(iii) The AWS is Unlawful in All Circumstances

A review finding that the AWS is unlawful *per se* because of its autonomous functionality may lead to considerable redesign, for example in terms of the training and testing of the neural network or the input data and algorithmic performance of an AI-based system. In this regard, the early correction of the

design and programming of the AWS will enable the algorithms (or other computing processes used) to be adjusted to prevent illegal operation. The use of an iterative review process can enable such corrections to be undertaken without having to unravel the complete design and development of the system. The potential for data bias, the incorrect legal presumptions being applied by algorithms, or the inability to reach appropriate conclusions based on changing inputs can be corrected through ongoing test, evaluation and verification processes that are linked to the weapons review process. This departure from the traditional process—the redesign and ongoing development of the system in response to legal review—marks a significant adjustment to account for autonomous functionality.

4. AWS Review Case Studies

This section presents hypothetical case studies that have been selected to demonstrate the general utility of weapons reviews for AWS at differing acquisition markers—the design and development of a new capability, the acquisition of a pre-existing capability and the enhancement of an existing weapons capability. In the Australian context, for example, the One Defence Capability System (ODCS) articulates a system of acquisition gates that mark certain technological readiness levels aligned with the Australian Defence Force's internal acquisition, resourcing and capability use processes.[73] Within the four phases of the ODCS, there are critical points to consider—and shape—the potential lawfulness of the capability prior to its use by the acquiring State.

During the Strategy and Concepts phase, for example, the setting of the capability priorities will assist in scoping the use case for the AWS and delimiting how it may be utilised by Defence. Considering legal limitations at this time through a weapons review (albeit a preliminary one) will assist in preventing a capability being planned for acquisition when its proposed use could never lawful. Equally, at this point in the ODSC, engagement in the weapons review process may result in an adjustment to the proposed use of the capability to fill the identified capability gap.

The Risk Mitigation and Requirement Setting phase is also important for considering legal risk. The preliminary weapons review conducted at this time will support an assessment as to the level of risk posed by the AWS and whether this legal risk remains at a level acceptable to Defence. This assessment may support the adjustment of the capability use requirements, which will impact on the manner in which the capability is designed and developed.

The Acquisition phase is obviously most relevant when it comes to assessing legal compliance of AWS through weapons review. This is the formal point of

[73] Australian Defence Force, 'Defence Capability Manual' (version 1.1, 3 December 2021) <https://defence.gov.au/publications/docs/Defence-Capability-Manual.pdf>.

review under current systems, at which point a final weapons review will determine whether or not the weapon would be capable of lawful use, and if limitations would need to be placed upon the use.

Finally, the In-Service and Disposal phase requires consideration of the use of the AWS. The relevant rules of engagement (ROE), and tactics, techniques and procedures (TTP), will reflect the parameters or limitations identified in the use of the AWS during its weapon review. During this phase, changes to the system based upon its operating environment or adjustments to use over time through either self-generated changes via machine learning, or adjustment by the systems' operators, may trigger reconsideration of review.

These case studies do not comprehensively cover the steps of weapons review processes outlined earlier but are intended to highlight critical issues that differentiate AWS from traditional weapons, and how risks relating to those differences can be mitigated through the use of weapons reviews.

A. Case Study One: AI Target Classifier

(i) The Capability

The Artificial Intelligence Target Classifier (AITC) is a fictitious AI-supported decision tool designed to enhance the use of weapons by tracking, identifying and prioritising targets.[74] The AITC analyses video footage and makes recommendations or gives warnings to military commanders to consider when planning or deciding upon an attack. The object classifier employs a convolutional neural network (CNN)[75] programmed to recognise specific visual objects and classify them based upon indicia relevant to categories of persons and objects on the battlefield. For example, the classifier can identify an individual in a pre-determined geographic area, carrying a recognised weapon, wearing military disruptive pattern uniform and not possessing any distinctive protective emblem as an enemy combatant with a high degree of certainty (95%). The classifier is in its early stages of study and development by a private company with military funding.

(ii) Preliminary Questions

(a) Does a Sponsor State's Article 36 Obligation Apply to the AITC? The traditional weapons review steps focus on determining the legality of a weapon *per*

[74] See, eg, Athena AI Defence target classifier, under development, <https://athenadefence.ai/capabilities>.

[75] A convolutional neural network is a deep learning algorithm which can take an input image, assign importance to aspects of the image and differentiate one image from another. See S Saha, 'A Comprehensive Guide to Convolutional Neural Networks: The ELI5 Way' (*Towards Data Science*, 16 December 2018) <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.

se by examining specific weapon types and certain effects. However, the AITC is likely to be categorised as a means of warfare in that it is an integral component of a weapon rather than a weapon itself. Let us assume that the reviewing State's weapons review policy includes a definition of 'weapon, means or methods of warfare' that includes 'components of a weapon system integral to the use of force'.[76] If so, the State's Article 36 obligation applies to the classifier on the basis that it is designed to inform decisions concerning the use of force. Alternatively, a reviewing State's policy may require all AI systems designed to perform functions that are governed by IHL (or contribute to belligerent acts) to undergo a weapons review to determine its ability to function lawfully.[77]

(b) Is the Study and Development of AITC by a Private Company, Prior to Military Acquisition or Adoption of the Capability, Subject to the Article 36 Review?. Some regard the study of dual-use technology that at some point may become capable of development as a weapon as outside the Article 36 weapons review obligation.[78] Similarly, a government's study of weapons, for example for intelligence purposes without a view to acquisition, is argued to be beyond the Article 36 obligation.[79] An alternative view is that weapons reviews should be undertaken when the State undertakes procurement processes of the new weapon for the purpose of operational deployment, although reviews can be done at all stages with different levels of intensity.[80]

We submit, however, that States sponsoring companies or universities to research military uses of technology should undertake reviews.[81] In this fictitious case, the State's military sponsored the development of the AITC for the specific purpose of enhancing the use of a weapon or weapons. Accordingly, we would argue that this study falls within the weapons review obligation. Conversely, if the reviewing State sponsored the study and development of the same AI technology that underpins the AITC, but without the specific application of the AI having been determined, it is less likely that the Article 36

[76] See, eg, Australian Department of Defence (n 56), defining 'means of warfare' as 'weapons or weapon systems'.

[77] Although there is differing opinion about the extent of the weapons review process to military capabilities that are not traditional weapons, the case studies in this article will not assess this preliminary question in depth. For a detailed argument supporting the inclusion of capability such as the AITC in weapons review processes (irrespective of extant State practice), see Klonowska (n 35).

[78] WH Boothby, *Conflict Law: The Influence of New Weapons Technology, Human Rights and Emerging Actors* (TMC Asser Press 2014) 168.

[79] Parks (n 26) 113.

[80] N Tsagourias and G Biggio, 'The Regulation of Cyber Weapons' in EPJ Myer and T Marauhn (eds), *Research Handbook on International Arms Control Law* (Elgar 2022) 440.

[81] See also M Kotlik, 'Reviewing Legal Weapons Reviews: Is It Possible to Verify Compliance?' (*EJIL:Talk!*, 19 March 2020) <www.ejiltalk.org/reviewing-legal-weapons-reviews-is-it-possible-to-verify-compliance/>.

review obligation would arise until a decision was made to use the AI as part of a weapon system.

A related question is whether Article 36 requires States to review weapons produced by domestic private industry for export, but which are not acquired or adopted by the State itself. While the temporal application of Article 36 is very broad,[82] the obligation is generally understood to apply to weapons at the stages of 'study, development, acquisition or adoption' by a State and not a non-State entity. The increased prominence of private industry in weapon manufacture may give States cause to consider the temporal application of the weapons review obligation in relation to their national weapons review directives and export control obligations. We consider that the obligation to respect and ensure respect for IHL[83] supports a broad reading of the review obligation for exported weapons.

(iii) Review

During the early stages of design and development, the weapons review focuses on the AI functions that will be governed by IHL. Given that the AITC's normal or expected use is to track, identify and prioritise targets, the application of IHL's cardinal rule of distinction is the key part of the analysis.

Noting that those responsible for the design and development of the AI are likely to be civilian employees of a private company, they are unlikely to be intimately familiar with IHL rules and their application during armed conflict. It may be necessary for those undertaking the weapons review to inform them of the distinction rule and how it is applied in the circumstances of the AITC's intended use. This includes explaining the categories of persons and objects that the CNN will be required to distinguish between, and the possibility of a person's or object's IHL status changing as a consequence of a person's actions (e.g., taking a direct part in hostilities or, conversely, being wounded) or an object's use. Notably, these criteria will change depending on the operational environment and nature of armed conflict in which the capability is deployed; however, the training of the CNN requires training in the ability to make such assessments.

Accordingly, an important element of the early weapon analysis will be to understand what aspects of an image the CNN is being trained to recognise as indicative of an IHL category. In the case of certain objects, for example an enemy tank, there will be less of a risk that a CNN will mis-identify civilian objects with features similar to a certain model of enemy tank. However, classifying persons as combatants will be more difficult. A CNN trained to recognise features such as the possession of a weapon, wearing a military uniform and

[82] ICRC Guide (n 11) 23.

[83] See generally, E Massingham and A McConnachie (eds), *Ensuring Respect for International Humanitarian Law* (Routledge 2020).

proximity to other combatants may result in a higher risk of false-positive identifications, noting that such visually apparent criteria are not determinative of targetable status. For example, although a combatant may meet the primary criteria to be targetable, they may also be *hors de combat* because of secondary considerations such as having effectively surrendered or being injured. Membership of armed forces cannot be determined just by assessing visual indicators of the person.[84] While not only a challenge for AI but also for humans required to assess the targetability of adversaries, the ability of an AI system to incorporate these additional, complex, context-based assessment criteria is more complex than adding this to the list of considerations for a human decision-maker.

Depending on the expectations of the AITC capability, initial positive identification will require assigning differing importance to different outputs, or adjusting filters to appropriately capture differing spatial and temporal dependencies to reflect these multiple considerations—all informed by the IHL performance requirements of the AITC.

(iv) Conclusion

Without input in relation to the specific operational use of the AITC, this capability would not be capable of IHL compliance. The outcome of the weapons review would be a conditional clearance pending further information about the ability of the AITC to meet the IHL requirements of distinction in its specific operational use scenario.

(v) Reflections

While it is difficult, and perhaps premature, to make final determinations during the study and development of an emerging weapon capability, a weapons review plays an important role in the study and development of the AITC. First, it informs and educates those responsible for the design, programming and training of the AI classifier of the reviewing State's IHL obligations. This ensures that the programming, data selection for training and testing and development

[84] See, eg, the list of indicia of functions that may support membership of organised armed groups as listed in the Explanatory Memorandum to the *Criminal Code Amendment Act (War Crimes) 2016* (Commonwealth of Australia) 9, [11] <https://parlinfo.aph.gov.au/parlInfo/download/legislation/ems/r5736_ems_73251af8-f937-4534-8dd8-ba98099a2399/upload_pdf/496771.pdf;fileType=application%2Fpdf>: ' Indicia of such functions may include: carrying arms openly; exercising command of the organised armed group or elements of it; giving or taking orders or acting on instructions from the organised armed group; direct involvement in achieving the military aims or objectives of the organised armed group; and other activities indicative of membership in an organised armed group which could include intelligence gathering, maintaining communications or providing engineering or logistics support.'

integrates IHL obligations into the system's operation. Secondly, it provides early indications of system capabilities and limitations. In this way, limitations in the AI functionality that impact on IHL compliance can be further developed or restricted early in the system development. This provides the reviewing State with an early indication of their ability to use the AI classifier in compliance with IHL.

B. Case Study Two: Acquisition of an Autonomous Aerial Platform

(i) The Capability

This fictitious scenario deals with the reviewing State purchasing an AWS Loiterer from an allied State. The Loiterer is an aerial platform designed to, once launched, fly along a pre-designated flight path for up to 6 hours and target military objectives employing specific communications systems.[85] It is equipped with sensitive radio frequency receivers that can detect and locate radio signals from radio equipment used in military headquarters and command vehicles. Operators can program the Loiterer to fly up to 3000 m in altitude and search within a designated geographic area for military objectives emitting the target radio frequencies.

Once a radio signal is identified, the Loiterer flies a specific path to triangulate the source of the radio signal and pinpoint its location. The platform carries two 50 gramme high explosive glide drones, which can be launched up to 2 km from the intended target. Once the target location is pinpointed, the platform passes the target information to one or both glide drones which are released and glide, undetected, to strike their target. Once the Loiterer launches the glide drones, the targeting process is fully autonomous. The Loiterer is programmed to return to a specified landing point on demand or when its fuel supply is low.

The normal or expected use of the Loiterer is the kinetic targeting of military objectives, including military headquarters. Testing has demonstrated that the glide drones strike their designed targets with 98% accuracy. While the 50 gramme payload has a lethal weapons effect radius of 10 m, it can cause collateral damage within 50 m of its point of impact. A human controller can monitor the Loiterer mission and, where necessary, cancel or suspend an attack. If there is a low risk of collateral damage, or if communications between the platform and its controller are unavailable, the Loiterer can operate autonomously once launched.

(ii) Preliminary Questions

The preliminary weapon review question is whether the Loiterer is a new weapon, means or method of warfare. It has been in service with other States

[85] See, eg, the Boeing Airpower Teaming System, the 'Loyal Wingman' or 'Ghost Bat', <www.boeing.com/defense/airpower-teaming-system/>.

and is therefore not new from a technology perspective. But it is new to the reviewing State and its particular international legal obligations, and therefore subject to weapons review.[86]

The Loiterer's normal or expected use is to attack specified military objectives, identified through radio signal emission, within a pre-designated geographic area with high explosive glide drones. Once launched, it is capable of searching, selecting and attacking its targets without human input. It does, however, have a human on the loop and in-built multiple fail safes in the event of communications loss.

(iii) Review

To give effect to the requirement of informing the government and military decision-makers of the legality of fielding the weapon, a reviewing State's weapon policy or directive may require an interim weapons review report. This report provides an assessment of legal risk in adopting a technology such as this, where its use is subject to an error rate, and thus an acceptance of risk in certain critical legal functions. In this case, the interim legal review may find that the normal or expected use of the Loiterer raises several legal risks that can then be assessed for acceptability, capability adjustment or determining limitations on the capability's use.

(a) *Distinction*. There is a risk that the Loiterer may unlawfully target civilian objects which use communications equipment with similar characteristics to the intended targets. The platform identifies military objectives by detecting specific radio frequencies. The platform's on-board system interprets the type of radio signal and detects signal modulation techniques typical of military communications. This identification of the specific radio emission is accurate (98% probability) but it presumes the absence of civilian radio communications within its search area using the same radio frequencies or modulation techniques. Civilian medical and emergency services in several countries are known to possess radio communications capable of emitting signals within the frequency range targeted by the Loiterer. In addition, the platform is unable to interpret the emitted radio signals, and thus cannot identify when combatants located within the target area expressing an intention to surrender; or identify if the objects are protected medical facilities. The interim review might therefore recommend, prior to approving acquisition that:

1. the platform's operating instructions require direct human approval prior to attacking military objectives identified by radio emissions, to ensure the

[86] ICRC Guide (n 11) 24.

objective is accurately assessed as a military objective and not otherwise protected; and
2. the operating system is designed to enable GPS coordinates of all military objects to be pre-programmed into the platform's operating instructions to enable no-strike locations to be identified and cross-checked against proposed targets; or alternatively,
3. the use of the platform is limited only to operating environments absent of civilians and civilian objects.

(b) *Collateral Damage*. There is a risk that the Loiterer will launch an attack that results in a loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which is excessive in relation to the concrete and direct military advantage gained. While the platform is intended to target military headquarters and other military objectives, it is unable to detect the presence of other persons or objects within the vicinity of a target. As a consequence, civilians or civilian objects located within the 50 m collateral damage radius may be injured and objects damaged. The review may therefore recommend that:

1. the platform's operating instructions require direct human approval prior to attacking military objective identified by radio emissions; or
2. alternatively, the use of the platform is limited only to operating environments absent of civilians and civilian objects.

(c) *Precautions in Attack*. The Loiterer's operating system is not programmed to take into account precautions in attack. In particular, it is not able to identify civilians and civilian objects and therefore is not able to take constant care to spare the civilian population, civilians and civilian objects. Moreover, the use of a single data source to determine objects emitting a specific radio signal to determine whether objects are military or civilian may not satisfy the requirement to 'do everything feasible' to verify the nature of objective to be attacked nor whether it is subject to special protection. To address these shortfalls, the review may recommend that:

1. the platform utilises a secondary data source, either integrated into the platform or via external, networked sensors, to distinguish objects as military objectives; or alternatively
2. the use of the platform is limited only to operating environments devoid of civilians and civilian objects.

(iv) Conclusion

The use of the Loiterer, in its proposed normal or expected use, presents legal risks and is unlikely to comply with IHL rules of distinction and proportionality. These identified legal risks may be addressed by modifying the platform's use to require direct human control over targeting decisions to ensure compliance with the rules of distinction, proportionality and precautions. The risks may be further reduced by ensuring the use of the platform is restricted to operating areas which lack civilians and civilian objects.

(v) Reflections

This case study differs from the first case study in a number of important ways. First, the Loiterer is a mature system that was developed by a commercial entity and was in use by another State. This limits the acquiring State's ability to modify the design or system programming. It does not, however, exempt or preclude the State from undertaking its own assessment to independently determine the limitations of the capability and determine whether it fulfils the military's needs. This study may include comparing several similar commercially available weapon capabilities to provide the government with several options to consider.

Secondly, the normal use of this capability is pre-determined by the manufacturer. In this regard, the acquiring State's acquisition process will assess whether this normal or expected use meets the State's requirements. In addition to the question of whether the platform is capable of use in compliance with IHL, a weapons review will consider national policy requirements, including engineering and programming standards of system performance, degree of human control, lines of accountability and acceptable targets.

C. Case Study Three: Uncrewed Armed Reconnaissance Vehicle

(i) The Capability

The fictional Uncrewed Armed Reconnaissance Vehicle (UARV) is a modified armed vehicle designed to conduct autonomous long-range reconnaissance and surveillance operations with a conventional heavy calibre machine gun mounted on its weapon station.[87] The machine gun is intended primarily for defensive use, although the UARV may be authorised to attack pre-designated targets.

The heavy machine gun was introduced into service in 2009 and a weapons review completed prior to this. The machine gun fires 0.50 calibre rounds to a

[87] See, eg, the Mission Master <www.rheinmetall-defence.com/en/rheinmetall_defence/systems_and_products/unbemannte_fahrzeuge/mission_master/index.php>.

distance of 4000 m from a remotely operated weapon station with an accuracy of 3–5 m at the maximum effective range.

The modified armed vehicle has not previously been considered by a weapons review as it was an unarmed reconnaissance vehicle that did not fall within the reviewing State's definition of a 'new weapon, means or method of warfare'. In addition to mounting the machine gun to this platform, modifications include an AI operating system that enables its mission profile to be programmed (for example, reconnaissance within defined geographic bounds). The AI determines how it will undertake its mission based on pre-programmed tactics, techniques and procedures, environmental conditions and operational information such as known enemy locations. The armed vehicle has video, infra-red and thermal cameras to provide its operating system with environmental data. The system can classify hundreds of objects, including all known enemy tracked and wheeled vehicles, and can identify known combatants' camouflage uniforms. Military commanders can monitor the location, imagery and actions of the UARV and provide additional directions, including suspending or restricting its operations, via secure satellite communications.

(ii) Preliminary Questions

The preliminary question is whether the new use of the heavy machine-gun will make the weapon 'new' for the purpose of Article 36 and require an additional determination of its legality through a further weapons review. This further review will consider the existing weapons review and likely adopt a number of its findings as they relate to the weapon design and use, which is not changed by its use with the UARV. The focus of the updated weapons review will be the autonomous functionality of the weapon and decisions concerning its use.

The question of whether the UARV itself is subject to a weapons review is answered by reference to the reviewing State's weapons review policy defining 'weapon, means and methods of warfare'. In this case, the armed vehicle platform is neither designed nor intended to be used as a weapon, for example by colliding with military objectives. However, the UARV operating system controlling the operation of the machine gun is likely to fall within the definition of means of warfare as its proposed functions—specifically target identification—are governed by IHL.

(iii) Review

Accordingly, the weapons review analysis centres on those aspects of the AI system responsible for operating the heavy machine gun including its target recognition, sighting and fire control systems. Other aspects of the AI system's functionality, for example vehicle driving speed, route selection and navigation, are only relevant to the weapons review to the extent that they affect the lawful

use of the weapon. For example, if the UARV is unstable when travelling at speed, affecting the accuracy of the machine gun, the weapons review may restrict the use of the machine gun to certain speeds.

While the UARV's intended use is surveillance and reconnaissance, the weapons review's focus is on the normal or expected use of the heavy machine gun. The AI system is programmed to engage the heavy machine gun only as a defensive measure in conjunction with other defensive actions, including evasive manoeuvres and deploying a smoke screen. Defensive use of the machine gun may not be intended to kill an enemy combatant or destroy a military objective, but to force an attacking enemy to undertake a defensive manoeuvre to terrain, affording protection from the machine gun fire. The purpose of defensive fire is to regain a tactical advantage and enable the UARV to conduct a safe withdrawal from danger and avoid being decisively engaged.

(iv) Conclusion

The new use of the existing capability, both on a new platform and with the introduction of autonomous functionality, will trigger the reviewing State's weapons review obligation. The review will be less about the weapon itself, but will focus upon the ability of the weapon system to sense and interpret its environment, identify lawful targets and engage targets in compliance with IHL. In reviewing only those components of the updated system that are intended to cause harm (to objects or people), the result of the review will then focus on whether those defensive methods comply with IHL. A different legal analysis is required if the system is used only in self-defence of persons outside of an armed conflict.[88]

(v) Reflections

Unlike the previous case studies, the focus of this review was on the method of use of existing military capabilities when combined with AI technology; and when combined with systems that would otherwise not have triggered a State's review obligations. In this way, reliance on previous assessments could be leveraged. They would require updating based upon the manner in which the previous capability was designed to be used by human operators. They would need to consider the methods by which that capability was to be limited in use by the associated AI technology. This review also requires consideration of the source

[88] See, eg, differing views by States on the right to use force in self-defence in armed conflict as deriving from combatant status as compared to a common law right to protect threats to life: CG Cooper, *NATO Rules of Engagement* (Brill Nijhoff 2020) 320–88; J Cherry and M Rizzotti, 'Understanding Self-Defense and the Law of Armed Conflict' (*Articles of War*, 9 March 2021) <https://lieber.westpoint.edu/understanding-self-defense-law-armed-conflict/>.

of authority for the use of force because some States refer to other legal frameworks in the use of weapons for the purposes of self-defence.

5. Conclusion

Weapons reviews of AWS have clear utility in assessing the compliance of these novel systems with a State's international legal obligations. The ability to identify how the system in intended to be used, as well as limiting the contexts in which an AWS may be fielded, is a critical outcome of such a process. This is particularly important given where machines assume functions that were previously undertaken by human operators.

It is true that behind the use of any machine is a human responsible for fielding the system—while possibly through more convoluted causal chains than assumed in a traditional weapons review. But when machines subsume more tasks, it becomes imperative to confirm that the human operator can rely on those functionalities. The additional considerations include, among other things, determining the standards for use, the expansion of weapons review to capabilities that fall within the States' definition of means of warfare, the tasks authorised to be performed by machines and the limits applied to a machine with respect to the use of force compared to a human conducting the same activity.

Arguably, a weapons review process only derives from Article 36 of Additional Protocol I and thus only applies to States party to that protocol, but several States that are not party to the Protocol are known to conduct weapons reviews.[89] The nature of AWS capability means, as we have argued, that the weapons review obligation needs to be interpreted broadly. This article has demonstrated the ways in which weapons reviews are necessary to enable the lawful use of AWS in compliance with States' general IHL obligations.

The integration of an iterative review process during the design and development of AWS will result in a system that is more capable of IHL compliance insofar as it ensures the necessary IHL standards are built into those functions of the AWS that rely upon computing decisions where humans would previously have undertaken them.

This iterative process may complement an existing Test and Evaluation, Validation and Verification processes associated with the design and implementation of new technologies. It can offer synergies to militaries and States looking to streamline their governance processes, for example through the concurrent certification of the AWS from both a legal perspective and for the purposes of existing acquisition processes.

Further, the output of the weapons review may be a recommendation about limitations on the autonomous functionality of a weapon system, mitigating the risk associated with accountability or IHL compliance. This output also provides an opportunity to address one of the primary concerns of the use of AWS:

[89] See, eg, 'Chairman's Summary Report' (n 21) 52 (Israel).

the extent of human control mandated in the use of the system. By assessing when an AWS can be lawfully utilised, and also articulating the limitations on use of the system's autonomous functionality, the concerns of a lack of human control and accountability of AWS can be addressed through policy constraints.

The case studies of this article are fictitious but they refer to existing technology. Although highly autonomous capabilities are being designed and fielded, it is likely that AWS will not reach certification or approval through review if they are not sophisticated enough to meet the reviewing State's IHL obligations. Accordingly, while a capability may be designed to be highly autonomous, it is likely to be reined in—so as not to authorise use in a highly autonomous mode—as a consequence of a legal review. For this reason, iterative reviews of AWS will be required where the system develops, whether through self-learning or general advancements in technology.

There remain challenges in articulating compliance of an AWS with legal standards in terms of translating the subjective and objective legal tests required to enable lawful use of lethal force in armed conflict—which is inherently a complex, unpredictable and congested environment. However, such an exercise is perhaps the only way to ensure the lawful use of these systems in future conflicts. Sharing the lessons learned from conducting these reviews will enhance the principles of the CCW; as well as States' understanding of the limitations of these systems and broader compliance requirements when commissioning novel technologies incorporating autonomy.

Funding

The authors' work received funding from the Australian Government s Next Generation Technologies Fund through Trusted Autonomous Systems, a Defence Cooperative Research Centre.

Conflict of Interest

Damian Copeland is a legal officer in the Australian Defence Force. Rain Liivoja has served on governmental delegations to the Group of Governmental Experts on Lethal Autonomous Weapon Systems; he is currently a Visiting Legal Fellow to the Australian Department of Foreign Affairs and Trade. Lauren Sanders is a reserve legal officer in the Australian Defence Force; she also advises the defence industry on issues relating to weapons reviews.