

Penn State Journal of Law & International Affairs

Volume 9 | Issue 2

May 2021

Decreasing Unintentional War: Governance Considerations For Regulating Lethal Autonomous Weapons Systems

Aiden Warren

Alek Hillas

Follow this and additional works at: <https://elibrary.law.psu.edu/jlia>



Part of the [International and Area Studies Commons](#), [International Law Commons](#), [International Trade Law Commons](#), and the [Law and Politics Commons](#)

ISSN: 2168-7951

Recommended Citation

Aiden Warren and Alek Hillas, *Decreasing Unintentional War: Governance Considerations For Regulating Lethal Autonomous Weapons Systems*, 9 PENN. ST. J.L. & INT'L AFF. 69 (2021).

Available at: <https://elibrary.law.psu.edu/jlia/vol9/iss2/6>

The Penn State Journal of Law & International Affairs is a joint publication of Penn State's School of Law and School of International Affairs.

**DECREASING UNINTENTIONAL WAR:
GOVERNANCE CONSIDERATIONS FOR
REGULATING LETHAL AUTONOMOUS
WEAPONS SYSTEMS**

By Aiden Warren & Alek Hillas⁺*

ABSTRACT

While the introduction of futuristic technologies will establish new options and precedents for state responses to security scenarios, there are important lessons to be drawn from prior crises. Beginning with a case study of newly sworn-in leaders during their first major foreign policy 'test,' this article examines changes to such security scenarios with reference to the development of policies on Lethal Autonomous Weapons Systems (LAWS). Leadership perceptions and the politics of human versus machine error, or sharing accountability of fault by states, is considered in the context of a detailed thought experiment. The second section of the article identifies enhanced collaborative rules for decreasing the probability of unintentional war at the level of military officers. The final section considers possible avenues to implement restrictions on LAWS in conventional warfare through various arms control models. Through the article's focus on existing architectures of global governance, readers will be presented with an analysis of the challenges that may confront future political leaders and technical experts in the field of emerging technologies.

* Aiden Warren is an Associate Professor of International Relations at the Royal Melbourne Institute of Technology (RMIT) University in Melbourne, Australia. His teaching and research interests are in the areas of international security, U.S. national security and foreign policy, U.S. politics (ideas, institutions, contemporary and historical), international relations (especially great power politics), issues associated with Weapons of Mass Destruction (WMD) proliferation, non-proliferation and arms control, and emerging technologies.

⁺ Alek Hillas is a researcher in the School of Global, Urban and Social Studies at RMIT University, where he graduated with a first-class honors degree in International Studies. His research interests are in global security and emerging technologies, including artificial intelligence and lethal robotics, and Australian foreign policy.

TABLE OF CONTENTS

INTRODUCTION.....	69
I. LEADERSHIP AND SECURITY SCENARIO REACTIONS	71
A. Case Study.....	74
B. Contextualizing the case study.....	76
<i>Table 1</i>	77
C. Drawing applications from the case study.....	80
II. DECREASING THE PROBABILITY OF UNINTENTIONAL WAR VIA ENHANCED COLLABORATIVE RULES.....	82
III. IMPLEMENTING RESTRICTIONS ON LAWS IN CONVENTIONAL WARFARE.....	89
CONCLUSION	96

INTRODUCTION

The development of complex and emerging technologies in settings of conflict presents distinctive challenges to global security. As this article will discuss, particular concern often relates to their management, and the principled and legally ‘permissible’ arguments that encircle them. Although the apprehensions pertaining to drones have been deliberated widely, it is the trajectory of Lethal Autonomous Weapons Systems (LAWS) which has now deepened the debate in this area among those working as military strategists for uninhabited systems, specialists of international humanitarian law, and ethicists working in artificial intelligence (AI). While the international community has not come to an agreement on a classification for LAWS, the most important disparity between existing weapons systems and LAWS is that humans would not directly command the LAWS to execute or kill a specific target. In this regard, the machine would come to its *own* decision independently within the parameters of its algorithm and mission specifications.

The activities LAWS would be permitted to carry out legitimately is a key policy consideration. In articulating the emerging scope of these “machines with the power and discretion to take lives without human involvement,” António Guterres has told the United Nations Security Council (UNSC) that LAWS “must be prohibited by

international law,”¹ and has attempted to provide leadership direction, most recently during an address to the Web Summit in December 2020 where he reiterated his call for a ban on LAWS.² As noted in the Secretary-General’s Roadmap for Digital Cooperation released in 2020, Guterres’ call was also matched by a recommendation from a High-level Panel on Digital Cooperation that he convened, co-chaired by Melinda Gates and Jack Ma. Here, the main emphasis was that, “life and death decisions should not be delegated to machines,”³ necessitating that such “priority actions deserve immediate attention,” so as to spur human rights and human agency.⁴ Guterres also announced his intention to seek a second five-year term from 2022.⁵ Should his bid be successful, Guterres will remain as the Secretary-General until the end of 2026, where he will most likely continue to push for state action on LAWS. Governance considerations for LAWS, thus, will most likely continue to expand into a range of areas,

¹ Press Release, U.N. Security Council, COVID-19 Pandemic Amplifying, Exploiting World’s Fragilities, Secretary-General Tells Security Council Debate on Protecting Civilians in Armed Conflict, U.N. Press Release SC/14196 (May 27, 2020), <https://www.un.org/press/en/2020/sc14196.doc.htm> [http://web.archive.org/web/20201020051526/https://www.un.org/press/en/2020/sc14196.doc.htm].

² U.N. Secretary-General, Secretary-General’s message to Web Summit (Dec. 4, 2020), <https://www.un.org/sg/en/content/sg/statement/2020-12-04/secretary-generals-message-web-summit> [http://web.archive.org/web/20210122223917/https://www.un.org/sg/en/content/sg/statement/2020-12-04/secretary-generals-message-web-summit].

³ U.N. SECRETARY-GENERAL, REPORT OF THE SECRETARY-GENERAL: ROADMAP FOR DIGITAL COOPERATION 17 (June 2020), https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf [http://web.archive.org/web/20210110171903/https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf].

⁴ U.N. Secretary-General, *The Age of Digital Interdependence: Report of the UN Secretary-General’s High-level Panel on Digital Cooperation*, at 4-5 (2019), <https://www.un.org/en/pdfs/DigitalCooperation-report-for-web.pdf> [http://web.archive.org/web/20210110171902/https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf].

⁵ *Guterres to seek second five-year term as UN Secretary-General*, UN NEWS (Jan. 11, 2021), <https://news.un.org/en/story/2021/01/1081832> [http://web.archive.org/web/20210112120831/https://news.un.org/en/story/2021/01/1081832].

including efforts to ‘decrease’ unintentional war, the subject of this article.

While no one can forecast with confidence the actions of LAWS, specialists from a multiplicity of professions across the public and private sectors must contemplate policies to decrease unintentional armed conflict involving LAWS. Such policies would be especially applicable for political leaders, who are more likely to be involved as the decisionmakers in response to an act that could be interpreted as either an accident or a provocation. Politicians may be responsible in deciding whether a ‘skirmish’ would veer into the direction of a diplomatic stand-off, or instead continue towards becoming a military activity. In endeavoring to provide much needed conjunctive material to these multifaceted discussions, this article will begin with a review of political responses during the critical junctures of potential hostilities, providing a case study on crisis response for new leaders. The article also identifies enhanced collaboration rules for decreasing the probability of unintentional wars at the officer level. Finally, it considers possible avenues to implement restrictions on LAWS in conventional warfare through various arms control models.

I. LEADERSHIP AND SECURITY SCENARIO REACTIONS

When political leaders opt to use force or to de-escalate a conflict in response to an incident, traditionally, the notion has been that antagonists could be relied upon to be logical actors whose pursuits can be defined and understood. However, when considering how to respond to a crisis where LAWS are present, it may be conceivable for robots to have coding inaccuracies. Conversely, it could be difficult to ascertain the chain of command and the supervision of LAWS from civilian or military authorities (or even non-state actors). In short, how would politicians respond differently to an incident where it is more difficult to determine a state’s intent, due to the inclusion of a machine ‘actor’? The section below engages with the notion that reactions by leaders in volatile scenarios (particularly newly sworn-in leaders), can be viewed as ‘heightened’ decisions in which the politician is facing a foreign policy ‘test’ or crisis. It is instructive, therefore, to consider how previous presidents have been confronted with ambiguous security incidents within months of

their inauguration, and to consider as a thought experiment whether new variables could be introduced in the future by the possible presence of uninhabited, autonomous devices.

An interesting historical example, which will be considered below, is the April 2001 South China Sea collision which occurred early in George W. Bush's first term. The selection of this incident is not intended to indicate that other presidents have not also had their resolve tested early during their first term. In March 2009, Chinese ships and aircraft harassed U.S. vessels in such a manner the Director of National Intelligence testified to the Senate Armed Services Committee that such actions constituted the "most serious" encounter since the April 2001 incident, with President Obama meeting that month with China's foreign minister in the White House.⁶ While intercepts in May and July 2017, early into the term of President Trump, had by then become more routine,⁷ it is possible President Biden may soon face challenges in response to encounters involving the military. Three significant events occurred in January 2021. A week before the Biden inauguration, the Trump Administration declassified and released the 2018 U.S. Strategic Framework for the Pacific, previously classified as "secret" and "not for foreign nationals."⁸

Initially to be declassified in the 2040s, the Strategic Framework for the Pacific indicates the U.S. would "devise and implement a defense strategy capable of . . . *defending* the first-island-

⁶ MICHAEL GREEN ET AL., COUNTERING COERCION IN MARITIME ASIA: THE THEORY AND PRACTICE OF GRAY ZONE DETERRENCE 61, 65 (2017), https://csis-prod.s3.amazonaws.com/s3fs-public/publication/170505_GreenM_CounteringCoercionAsia_Web.pdf?OnoJXfWb4A5gw_n6G.8azgEd8zRIM4wq [http://web.archive.org/web/20190915093922/https://csis-prod.s3.amazonaws.com/s3fs-public/publication/170505_GreenM_CounteringCoercionAsia_Web.pdf?OnoJXfWb4A5gw_n6G.8azgEd8zRIM4wq].

⁷ Idrees Ali, *Chinese jets intercept U.S. surveillance plane: U.S. officials*, REUTERS (July 24, 2017, 9:46 AM), <https://www.reuters.com/article/us-usa-china-military/chinese-jets-intercept-u-s-surveillance-plane-u-s-officials-idUSKBN1A91QE>.

⁸ Laura Tingle, *Previously secret details of Trump administration's Indo-Pacific strategy revealed*, ABC NEWS (Jan. 12, 2021, 3:52 AM), <https://www.abc.net.au/news/2021-01-12/details-of-trump-administrations-indo-pacific-strategy-revealed/13052216>.

chain nations, including Taiwan.”⁹ Several days after President Biden’s inauguration, China flew eight nuclear-capable bombers and four fighter jets into the Taiwanese air defense identification zone,¹⁰ and passed a new law explicitly allowing its coastguard to fire on foreign vessels,¹¹ described by Manila as a “verbal threat of war to any country that defies the [Chinese] law; which, if unchallenged, is submission to it.”¹² While this may presage a future foreign policy challenge in the region, the technology for LAWS does not yet exist. However, as the sophistication of AI continues to expand at an exponential rate, it is not far-fetched to suggest a future president may encounter a scenario with similarities to the one envisioned in our thought experiment outlined below. Through an initial background of the case study, its re-contextualization involving LAWS, and applications for policy development, the thought experiment intends to draw on past lessons to guide future leadership and security scenario reactions.

⁹ Cabinet Memorandum on the U.S. Strategic Framework for the Indo-Pacific, at 2 (Feb. 15, 2018) (declassified Jan. 5, 2021), <https://assets.documentcloud.org/documents/20449107/us-strategy-document-on-indo-pacific.pdf> [<http://web.archive.org/web/20210124024845/https://assets.documentcloud.org/documents/20449107/us-strategy-document-on-indo-pacific.pdf>]

¹⁰ *China flies nuclear-capable bombers, fighter jets over Taiwanese waters*, ABC NEWS (Jan. 23, 2021, 8:22 PM), <https://www.abc.net.au/news/2021-01-23/china-flies-nuclear-capable-bombers-and-jets-over-taiwan-waters/13086192>; see also *US says support for Taiwan ‘rock-solid’ as Chinese fighter jets enter defence zone for second day*, SBS NEWS (Jan. 25, 2021), <https://www.sbs.com.au/news/us-says-support-for-taiwan-rock-solid-as-chinese-fighter-jets-enter-defence-zone-for-second-day> [<http://web.archive.org/web/20210124222042/https://www.sbs.com.au/news/us-says-support-for-taiwan-rock-solid-as-chinese-fighter-jets-enter-defence-zone-for-second-day>].

¹¹ *China authorises coastguard to fire on foreign vessels in disputed waters of East China and South China Seas*, ABC NEWS (Jan. 23, 2021, 5:37 PM), <https://www.abc.net.au/news/2021-01-23/china-authorises-coast-guard-to-fire-on-foreign-ships-if-needed/13084754>.

¹² *Philippines files protest over China’s new coastguard laws in South China Sea*, ABC NEWS (Jan. 28, 2021), <https://www.abc.net.au/news/2021-01-28/philippines-protests-chinas-threat-of-war-coastguard-law/13096898>.

A. Case Study

As Beijing's response to the April 2001 collision bears some similarities to the crisis of its embassy in Belgrade being struck by U.S. missiles in May 1999, it is worthwhile providing some context on the transition from President Clinton to President Bush as written from the perspective of the Chinese leadership. The leaders of Beijing believed the 1999 incident was an intentional act, whereas Washington made efforts to demonstrate the accident was attributed to human error.¹³ Memoirs of the incidents from China's perspective are available, for example, from Qian Qichen (the then Vice-Premier of China). Recalling the "attack" on the embassy, which he believed could not have been a "mistake," Qian asserted that "the midair collision seemed anything but accidental" because the "U.S. spy planes frequently had been flying close to Chinese territory in the South China Sea area."¹⁴

In addition, Tang Jiaxuan (the then Foreign Minister of China) provides a detailed account, from his anticipation of the Bush presidency to the thinking of Chinese leaders following the collision. "From the second half of 2000," Tang recalls an increase in "reconnaissance flights along the edge of China's waters."¹⁵ Moreover, Bush's rhetoric from a presidential debate had described China as a "strategic competitor" instead of a "strategic partner," and Tang considered that, "it was disturbing to think about the possible adverse impact on China–U.S. relations should the Bush argument become the China policy of the U.S. government."¹⁶ On 1 April, less than three months after Bush's inauguration, the collision occurred between the two nations' military aircraft. The Chinese pilot went missing, and the U.S. pilots performed an emergency landing on Hainan, where they were taken into custody by the Chinese authorities. The U.S. pilots considered themselves to be "like hostages," although this characterization is disputed by Tang, who describes the status of the pilots as "neither tourists nor honored guests but intruders" who had

¹³ SUSAN L. SHIRK, CHINA: FRAGILE SUPERPOWER 218 (2007).

¹⁴ QIAN QICHEN, TEN EPISODES IN CHINA'S DIPLOMACY 157 (2005).

¹⁵ TANG JIAXUAN, HEAVY STORM AND GENTLE BREEZE: A MEMOIR OF CHINA'S DIPLOMACY 328 (2011).

¹⁶ *Id.* at 330.

“constraints on them . . . to demand that they cooperate in the [Chinese] investigation.”¹⁷ Within days, three U.S. Navy destroyers were dispatched to the vicinity and the presidents of both countries made public statements. Meanwhile, a Chinese search-and-rescue operation involving one hundred thousand people was underway in the area.¹⁸

Susan Shirk, the Deputy Assistant Secretary of State responsible for China in the Clinton Administration, coordinated Washington’s interagency investigation into how the Chinese embassy in Belgrade had been targeted, and determined that there was “a combination of errors” linked to a failure to follow processes.¹⁹ However, Shirk notes the perception in China maintained that the U.S. had superior technology and management techniques, and therefore, according to this perception, human error could not account for the “mistake.”²⁰ In a comparison of the two incidents, Shirk also notes, the Chinese media’s reporting on the airplane collision had “less enflamed rhetoric than they had used during the Belgrade embassy crisis.”²¹ While the Chinese government had asked for students to stay on campus during the airplane crisis (where young people vented anti-American sentiment on the internet), Beijing’s previous response during the Belgrade embassy crisis had been to organize buses from campuses to outside the U.S. embassy, where the students had thrown Molotov cocktails and bricks.²² In addition, according to a May 2013 report commissioned by the Pentagon’s most senior strategist, the Chinese response during the crisis in April 2001 was sophisticated and exploited the U.S. media.²³ In an analysis of coverage between Xinhua

¹⁷ *Id.* at 350.

¹⁸ *Id.* at 340–41.

¹⁹ SHIRK, *supra* note 13, at 218.

²⁰ *Id.*

²¹ SHIRK, *supra* note 13, at 238.

²² *Id.* at 213–14, 237–39.

²³ Peter Mattis, *Chinese Propaganda and Positioning in the Sino-American Crises: The EP-3 and the Impeccable Case*, in CHINA: THE THREE WARFARES 226–245 (Stefan Halper ed., 2013), http://images.smh.com.au/file/2014/04/11/5343124/China_%20The%20three%20warfares.pdf [http://web.archive.org/web/20201109003450/http://images.smh.com.au/file/2014/04/11/5343124/China_%20The%20three%20warfares.pdf]; see also John Garnaut, *US unsettled by China’s ‘three warfares’ strategy: Pentagon report*, SYDNEY MORNING HERALD (Apr. 11, 2014), <https://>

and the New York Times, international opinion was shaped through a campaign of influence which “juxtaposed Xinhua’s version of events with U.S. statements without subjecting either side to analytic scrutiny, giving Beijing’s propaganda pronouncements equal standing,” and thus it created a “false equivalency,”²⁴ which shifted perceptions among some of the U.S. populace further toward the Chinese view.

In a bilateral dispute of this nature, the number of positions is limited to just two: a position supported by either one party, or by both parties. Initially, the U.S. and China had maintained separate positions, accusing the other of fault. The U.S. later acquiesced to some elements maintained by the Chinese position, to such an extent that it created, “from a diplomatic or conflict resolution standpoint, *constructive ambiguity*.”²⁵

B. Contextualizing the case study

The case study of the incident in April 2001, outlined above, provides a limited frame of reference and serves as an example for illustrative purposes only. As a thought experiment, contemplate if there was a collision between a human pilot and an uninhabited device (instead of colliding with another human or group of humans). A method to conduct this analysis would need to consider the potential variables – the states’ positions, the scenario in dispute, and the difficulty in reaching an outcome. *Table 1* (below) considers a simplified hypothetical scenario if the U.S. EP-3 manned plane were replaced with an autonomous aircraft. A more comprehensive framework could, however, rely upon such an examination across multiple scenarios, developed by specialists to further consider the range of policy options available in advising leaders’ decision-making and crisis responses.

www.smh.com.au/politics/federal/us-unsettled-by-chinas-three-warfares-strategy-pentagon-report-20140410-36g45.html

²⁴ Mattis, *supra* note 23, at 235, 240.

²⁵ Kevin Avruch & Zheng Wang, *Culture, Apology, and International Negotiation: The Case of the Sino-U.S. ‘Spy Plane’ Crisis*, 10 INT’L NEGOT. 337, 345 (2005), <http://dx.doi.org/10.1163/1571806054740958>.

Table 1

Contentions and possible arguments for collision with an autonomous system

<i>Contention</i>	<i>Argument in favor</i>	<i>Argument against</i>
<i>Human at fault</i>	Previous history of unsafe maneuvers from the pilot.	Pilot has skills and experience (e.g., flight hours) to not make a fatal mistake.
<i>Machine at fault</i>	Software issues difficult to rule out due to algorithm 'length,' or computer operation affected by hardware malfunction or other issues.	Algorithm indicates the system adheres to codes of conduct and international law, and collectively it may have more 'hours' than any single human.
<i>Shared fault</i>	Military exchanges did not develop adequate policies to prevent incident. Machine may use components in a supply chain from both countries.	One military did not follow advice received during exchange from the other. A state's quality assurance processes were not adequate.
<i>Accident or causes unable to determined</i>	May be difficult to reach an objective finding if 'black box' (or equivalent type device) is unable to be recovered.	Political matter with history and considerations that go beyond the technical aspects of the most recent incident.

Each row of the above table will be briefly discussed below.

The first contention is where the human is at fault. Accounting for the uncertainties associated with new technology, it is likely that the government responsible for the manned plane would indicate the number of flight hours its pilot had undertaken in order to ‘prove’ the collision was not due to human error. The Chinese pilot, Wang Wei, had recorded 1,152 flying hours, “without incidents of negligence, oversights, or anything to indicate the possibility of an accident,” according to Tang’s account of the actual crisis.²⁶ Yet, this would not rule out pilot error. In December 2000, the U.S. had complained about the PLA flying in close proximity. Moreover, weeks before the collision, Wang Wei harassed the U.S. crew by flying close enough to display something through the surface of his aircraft canopy; the U.S. later released a photo of Wang holding a piece of paper against his window, upon which he had written his email address as a taunt, which allowed the U.S. to describe Wang’s reputation as someone “known to the American EP-3 crews as a *Top Gun*-style hotdogger.”²⁷ Clearly, the presence of an autonomous aircraft would not necessarily eliminate the possibility that a human could still have caused the incident.

The second row of the table considers the possibility where the machine is at fault. In analyzing this hypothetical scenario, the proponents defending the autonomous aircraft could suggest the algorithm and hardware were demonstrably reliable. According to such a proposition, if the sensors were not faulty, and without mistakes in the programming, the LAWS could have conceivably followed international regulation and bilateral codes of conduct, ruling out machine error as the cause of the crash. If the other party submitted a request to verify the evidence in support of this claim, it may also be possible to do so, provided both states held the view that the matter could be resolved through an investigation on the technical aspects and without ‘politicizing’ the broader context, e.g., avoiding deliberation on previous encounters which would be considered outside the scope.

²⁶ TANG, *supra* note 15, at 340.

²⁷ SHIRK, *supra* note 13, at 237; *see also* Mattis, *supra* note 23, at 235.

Relying upon technical considerations as the sole criteria, however, is a risky proposition based on the precedent it creates. In the case of the May 1999 bombing of the embassy in Belgrade, China's perception that U.S. technology was 'superior' did not allow its leaders to acknowledge U.S. officers had made mistakes in targeting the wrong building. Moreover, according to the argument by Chinese officials, the belief that machine error was not the cause, had the effect of exacerbating the crisis by suggesting that the incident 'must' therefore have been due to 'intentional' decisions by people. In the hypothetical experiment of an autonomous aircraft crashing with a human pilot, even if a machine had not been at fault in that specific case, elevating the status of technology above the level of humans creates a risk that future incidents are unable to be explained as machine error.

Another justification for the machine not being at fault could be the policy of a state for their autonomous devices to protect military property. In this scenario, the device would effectively be relying upon an argument for self-defense. This aspect would be particularly important if, instead of a collision occurring, the machine, anticipating the possibility of its own destruction, had fired upon the manned aircraft and in the process, injured the pilot. On the other hand, the loss of an uninhabited device could be an alternative policy outcome by providing a moral justification to prevent the use-of-force against a human in peacetime.

The contention in the third row of the table relates to fault being shared between the parties. Such a scenario would require the national interests of both sides to value the bilateral relationship. In addition, leaders would need to support the further deepening of military-to-military exchanges, including by a joint development of policies intended to mitigate the reoccurrence of such an incident. Even if fault could be attributed to both parties, it is unlikely to be an entirely equal 'share' of the blame. For example, it is possible one side could have called upon the other to make certain adjustments during previous 'close calls,' or that a greater portion of the hardware or software may originate from only one of the countries.

The adoption of joint responsibility as an outcome would, nevertheless, most likely signify greater influence by foreign ministries

in seeking to reset the agenda in the interests of stabilizing the bilateral relationship, compared to the specialized role of the military where the protection of the life of enlisted personnel may be of a greater concern. In the hypothetical scenario, there would be less motivation for such intensive contact between militaries, and less incentive for a state to admit fault with the operation of the autonomous device and apologize. In April 2001, the U.S. led by its ambassador (who was himself a former navy pilot) sought extensive consultation with the Chinese foreign ministry to secure the release of the EP-3 crew. In the absence of prisoners, the injured nation may evaluate different retaliatory measures, in juxtaposition to a human being taken into custody.

Finally, as noted in the fourth row of the table, the contention here would be that the collision was an accident or that the causes were unable to be determined. In this scenario, it is probable that the matter would encompass political judgements that are beyond the scope of the stand-alone incident. Originating either due to a lack of evidence regarding the causes of the incident – or where a finding indicated that the incident occurred despite the proper adherence to process by all parties – in this scenario it may be possible to allow for a political consensus to be reached with greater ease. While the prospect that a future incident may be unable to be avoided due to an absence of ‘lessons learned,’ and could thereby appear to be unsatisfying, this outcome would not necessarily preclude the benefits of remediation efforts that could arise in response to the ‘shared fault’ scenario. Conversely, it may even be convenient (for some actors) to not find evidence during their investigation, which could suggest a ‘human at fault’ or ‘machine at fault’ contention and, thus, a rearrangement of the political circumstances to something less neutral.

C. Drawing applications from the case study

Evidently, the presence of autonomous devices in the future would make such decision-making even more difficult than it is today. Given the intense uncertainty that decisionmakers will potentially confront when evaluating and ultimately reacting to such scenarios – perhaps in time-critical situations – it is essential they recognize all possibilities of human or machine error and the broader milieu before

rushing to any conclusion. As the precipitous speed and intensification of technology progresses in the digital age, political leaders will encounter circumstances requiring determinations involving uncertainty inherent to incidents where LAWS are likely to be involved.

If a dispute involving LAWS escalates, leaders should continue to communicate with their counterparts, when feasible, to prevent miscalculation if there is continued ambiguity. The establishment or use of a ‘hotline’ could provide logistical capabilities. However, its presence may indicate only a pre-determined intention to communicate, rather than facilitate a pathway toward a successful resolution. Some leaders may even be unwilling to pick up the phone as a negotiation tactic. While states undoubtedly could reciprocate against hostile force, human assessment is necessary before scaling up the antagonism as LAWS could not make these judgments objectively. Ultimately, while the Charter of the United Nations does not place restrictions on states’ intrinsic rights to individual or collective self-defense against an armed attack until the UNSC takes measures to restore international peace and security, state responses should nevertheless demonstrate military necessity, proportionality, and clearly distinguish the enemy combatants from non-military targets in accordance with the principles of Just War.²⁸ Furthermore, the Guiding Principles developed by the Group of Governmental Experts (GGE) have “affirmed that international law, in particular the UN Charter and international humanitarian law,” will guide the work towards an outcome on LAWS through the Convention on Certain Conventional Weapons (CCW). The status of the principles, however, is not currently legally binding, and thereby it remains unclear as to whether the international community has an appetite to introduce stricter rules in this area.²⁹

²⁸ Peter M. Asaro, *How Just Could a Robot War Be?*, in CURRENT ISSUES IN COMPUTING AND PHILOSOPHY 55 (Adam Briggie et al. eds., 2008).

²⁹ *Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, at 13, U.N. Doc. CCW/GGE.1/2019/3 (Sept. 25, 2019), <https://undocs.org/en/CCW/GGE.1/2019/3> [<http://web.archive.org/web/20200919072200/https://undocs.org/en/CCW/GGE.1/2019/3>].

As the article will illuminate further, the proliferation of autonomy in AI is already establishing uncertain benchmarks for a largely untested military technology, with some of these consequences becoming foreseeable in the form of leadership reactions to ambiguous security scenarios. Having considered the broader context and variables at the top of the chain-of-command in the section above, the next section focuses on the officer level. It presents a set of measures aimed at decreasing the scope of unintentional war through enhanced collaborative rules, before an incident escalates into a crisis requiring decision-making to move from the military to political levels.

II. DECREASING THE PROBABILITY OF UNINTENTIONAL WAR VIA ENHANCED COLLABORATIVE RULES

Militaries can redefine the security environment they inhabit by discovering areas for collaborative development with each other. The precepts of International Humanitarian Law (IHL) and the Rules of Engagement (ROE) are generally well understood as they apply to various professions. Among other potential areas for collaboration, regulations such as IHL and ROE will need to be clarified from a legislative standpoint and converted into ‘zeroes and ones,’ which would allow a robot to implement the programs in the intended manner of its instructions. It is more likely such a process would face challenges similar to those which already exist today. Digital tools to calculate and reduce collateral damage have been used by the U.S. Department of Defense since at least 2003.³⁰ Yet, while software to assist officers with estimating the collateral damage is based on particular variables, it currently does not provide an answer to all of the requirements to make a decision, such as military necessity.³¹ Of

³⁰ Michael Press, *Of Robots and Rules: Autonomous Weapon Systems in the Law of Armed Conflict*, 48 GEO. J. INT’L L. 1337, 1359 (2017), <https://www.law.georgetown.edu/international-law-journal/wp-content/uploads/sites/21/2018/05/48-4-Of-Robots-and-Rules.pdf> [http://web.archive.org/web/20210126045354/https://www.law.georgetown.edu/international-law-journal/wp-content/uploads/sites/21/2018/05/48-4-Of-Robots-and-Rules.pdf].

³¹ Kenneth Anderson et al., *Adapting the Law of Armed Conflict to Autonomous Weapon Systems*, 90 INT’L L. STUD. 386, 403 (2013), <https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1015&context=ils> [http://web.archive.org

course, even if IHL and ROE could be agreed upon by legal scholars in collaboration with software developers, in general, algorithms tend to have some flaws which become evident during their use. The probability of a mistake being coded into such a program to execute a legal decision by LAWS, will require the highest level of scrutiny.

We therefore recommend applying collaborative principles in this area to foster greater exchanges between states' militaries, including those engaged at the officer level. Developing a shared understanding could promote a higher level of assurance than perhaps could be indicated inhouse or by a contracted audit partner alone. That said, there are seemingly contradictory priorities between being 'completely' transparent and maintaining information that would fast-track an adversary's development for their own technologies. Each state will need to make its own judgment relating to their appropriate management of risks. The forfeiture of confidential information relating to the legal assessments of IHL and ROE compliance by a machine, may otherwise decrease the probability of inadvertent outcomes occurring from disputes involving LAWS. These would include the possibility of a confrontation between two sets of opposing LAWS where each nation state has deployed a model with proprietary legal interpretations.

Specific models may be tailored to a particular state's national interest within the architecture of global governance. While informal dialogue certainly could assist in the early stages of development, a treaty may promote a broader level of consensus. Alternatively, states may proclaim to observe key components as a matter of policy, while some activities remain separated from the agreement in specific and limited circumstances. The CCW has a total of 125 state parties and four signatories; however, this number is also spread across different protocols. For example, only 86 states have signed up to Amended Article I, which extends the application of the CCW and its protocols to non-international armed conflicts, whereas an additional 32 states have also agreed to Protocol I, prohibiting non-detectable fragments.³²

[/web/20210126023920/https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1015&context=ils](https://web/20210126023920/https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=1015&context=ils)].

³² United Nations, *High Contracting Parties to the Convention on Certain Conventional Weapons*, [https://www.unog.ch/80256EE600585943/\(httpPages\)/3CE](https://www.unog.ch/80256EE600585943/(httpPages)/3CE)

Therefore, even within a model shaped through the architecture of the United Nations, state parties can join over time and, thus, increase the universality of an arms control regime progressively. As the next section will also indicate with reference to the Missile Technology Control Regime (MTCR), an agreement outside of the United Nations can still influence global norms when its design and legacy are given legitimacy or recognition by peak UN bodies. Evidently, the pursuit of a common goal can be achieved through collaborative efforts on an essentially non-binding basis, without prejudice to the national interest or commitments to an agreement that ‘locks in’ a future administration to potentially irreversible national security risks.

Militaries may also choose to reach an in-principle understanding with non-governmental organizations (NGOs) to further their implementation of IHL and ROE. An advantage is how such an approach could conceivably provide for the external detection of inaccuracies that were ‘missed’ during a preliminary self-assessment, or to provide additional assurance to humanitarian workers on the compliance of LAWS as new (but verifiable) weapons. Noting the current discussions on the framing of LAWS in legal versus ethical terminologies,³³ such talks could also consider the compliance aspects of machine autonomy from a non-theoretical perspective. For example, while an algorithm should be able to visualize and acknowledge the unique status of non-combatants, NGO workers voluntarily could provide military officials with information on their activities to update the databases used to qualify the legality of a strike. As noted by Lieutenant Colonels John Cherry and Christopher Korpela, BONUS sub-munitions are advanced enough to compare an image of a target to their databases and self-destruct in the middle of their trajectory if a medical vehicle is identified instead of a combatant

7CFC0AA4A7548C12571C00039CB0C?OpenDocument [http://web.archive.org/web/20201130052613/https://www.unog.ch/80256EE600585943/(httpPages)/3CE7CFC0AA4A7548C12571C00039CB0C?OpenDocument] (last updated June 17, 2020).

³³ Elvira Rosert & Frank Sauer, *How (Not) to Stop the Killer Robots: A Comparative Analysis of Humanitarian Disarmament Campaign Strategies*, 42 CONTEMP. SEC. POL'Y 4, 22 (2021), <https://doi.org/10.1080/13523260.2020.1771508>.

vehicle.³⁴ As stated, “munitions such as a BONUS round properly employed by a commander and other similar technologies are the current and future of LAWS, not killer robots detached from command and control.”³⁵

In this manner, NGOs which otherwise may have reluctance to provide their assistance in areas of conflict, would have an opportunity to collaboratively develop protections for use by humanitarian workers who may not believe LAWS inherently possess the proper situational awareness and capability for distinction, compared with humans. While such actions alone are a small component of the requirements for ‘meaningful human control,’ they could also assist with weapons review processes before and throughout deployment. Yet, the premeditated distortion of this system “would arguably amount to perfidy,”³⁶ which could be considered under IHL to “betray the confidence of the adversary” through “simulation of protected status,” and depending on the circumstances, may be considered as a war crime.³⁷

Nevertheless, effectively validating compliance with IHL and ROE does not necessarily overcome the dangers arising from LAWS failing to have enough circumstantial awareness. Some actions are symbolic but can be crucial for sending a warning. As indicated in the previous section, the actions by Wang Wei in his previous flights could have multiple interpretations. While the U.S. interpretation had been

³⁴ John Cherry & Christopher Korpela, *Enhanced Distinction: The Need for a More Focused Autonomous Weapons Targeting Discussion at the LAWS GGE*, INT’L COMM. RED CROSS HUM’N L. & POL’Y (Mar. 28, 2019), <https://blogs.icrc.org/law-and-policy/2019/03/28/enhanced-distinction-need-focused-autonomous-weapons-targeting/> [http://web.archive.org/web/20210126053636/https://blogs.icrc.org/law-and-policy/2019/03/28/enhanced-distinction-need-focused-autonomous-weapons-targeting/].

³⁵ *Id.*

³⁶ Maziar Homayounnejad, *Ensuring Lethal Autonomous Weapon Systems Comply with International Humanitarian Law*, TRANSNAT’L L. I. THINK 22 (2017), <http://dx.doi.org/10.2139/ssrn.3073893>.

³⁷ *Rule 65. Perfidy*, INT’L COMM. RED CROSS, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule65 [http://web.archive.org/web/20210125085734/https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule65] (last visited Mar. 3, 2021).

that Wang was ‘taunting’ its crew, in the view of the Chinese, he may have been warning them. Going back even further into history, simulated attacks from the USSR had “enable[d] U.S. commanders to recognize these attacks during crises” and thus provided “an ‘action language’ for signaling their U.S. adversaries and American political leaders.”³⁸

Although the above proposals on decreasing the probability of unintentional war via enhanced collaborative tools primarily relate to LAWS not possessing the requisite recognition in the context of signaling, it is worth noting that leaders may not yet have come to a common understanding on how the deployment of LAWS could, itself, be used as a signal. Further underscoring the continued use of messaging in heightened security crises, one of the policies of the U.S. today is to “deter regional aggression and assure distant allies” through the flights of bomber aircraft which provide “effective signaling for deterrence and assurance.”³⁹ In the forthcoming years, one leader may need to determine how the presence of LAWS could impact another’s perceptions. Would the deployment of LAWS be a similar threshold? Thirty states have so far have called for a ban on fully autonomous weapons,⁴⁰ and they may hold different perceptions than nations considering the potential benefits that may arise from developing such autonomous systems.

Furthermore, people may also need to consider how robots ‘themselves’ may signal an adversary. As both states would have an interest in preventing an inadvertent conflict that leads to unplanned escalation, political and military strategists should consider what

³⁸ Sean M. Lynn-Jones, *A Quiet Success for Arms Control: Preventing Incidents at Sea*, 9 INT’L SEC. 154, 159 (1985), <http://dx.doi.org/10.2307/2538545>.

³⁹ U.S. DEP’T OF DEF., NUCLEAR POSTURE REVIEW 47 (2018), <https://media.defense.gov/2018/Feb/02/2001872886/-1/-1/1/2018-NUCLEAR-POSTURE-REVIEW-FINAL-REPORT.PDF> [<http://web.archive.org/web/20190920051805/https://media.defense.gov/2018/Feb/02/2001872886/-1/-1/1/2018-NUCLEAR-POSTURE-REVIEW-FINAL-REPORT.PDF>].

⁴⁰ MARY WAREHAM, STOPPING KILLER ROBOTS: COUNTRY POSITIONS ON BANNING FULLY AUTONOMOUS WEAPONS AND RETAINING HUMAN CONTROL 4 (2020), https://www.hrw.org/sites/default/files/media_2020/08/arms0820_web_0.pdf [http://web.archive.org/web/20210130055002/https://www.hrw.org/sites/default/files/media_2020/08/arms0820_web_0.pdf].

customs are likely to eventuate from the introduction of this emerging technology. By means of ensuring a ‘failsafe’ is available in the event of ambiguous signaling, it may be worthwhile to have communications infrastructure between the humans who administer or oversee the LAWS. The essential factor is humans preserving the ability to communicate with their opponent. Assuming the political and military leadership supported such a proposal, the method would need to be reached technically. One option may include having centralized databases for communications across certain geographical regions, as agreed by a treaty between two or more countries. Another pathway may allow for one LAWS in proximity to another to pass on messages sent between the respective human counterparts. In yet another scenario, following misinterpretation between a human and the LAWS that could result in the LAWS not understanding a signal of intent, the human could seek to send a standardized ‘message-alert’ through the LAWS to its human commander, who may possess the disposition to avert escalation or mitigate the scale of retaliation. Implementation of the above measures should, thereby, decrease the likelihood of unintentional actions from LAWS or avert unnecessary escalation.

The creation of the requisite infrastructure for such transmissions would be complicated in certain types of conditions, such as underwater. As Krepinevich notes, “It seems plausible that, given the cost disparity” between autonomous weapons compared to submarines, “maritime competitors could arm and deploy them as undersea ‘kamikaze’ devices or delivery systems” for anti-submarine warfare.⁴¹ However, the underwater links needed for the observation of such devices would be difficult to construct, not to mention who wears the cost. Unlike land or aerial systems, maritime operations will have greater difficulties with relaying data signals.⁴² Nevertheless, innovations in the area of underwater communications indicate some possibilities for considering the development of “hydrospatial

⁴¹ ANDREW F. KREPINEVICH, MARITIME COMPETITION IN A MATURE PRECISION-STRIKE REGIME 82 (2014), <https://csbaonline.org/uploads/documents/MMPSR-Web.pdf>.

⁴² ANTHONY FINN & STEVE SCHEDING, DEVELOPMENTS AND CHALLENGES FOR AUTONOMOUS UNMANNED VEHICLES: A COMPENDIUM (2010), <http://dx.doi.org/10.1007/978-3-642-10704-7>.

infrastructure”⁴³ with shared applications. Enabling such an autonomous device to remain surreptitious inside of a delineated area may be possible while maintaining infrastructure for an enhanced set of collaborative rules. While not a military example, the voyage of filmmaker and explorer James Cameron to the Mariana Trench, the deepest point in the ocean, demonstrated how a civilian was able to remain in constant contact with the surface via an underwater telephone, respond to a voice message from his partner, and send a tweet from that location.⁴⁴ Communication with underwater LAWS would be significantly less cost-effective if each nation pursued separate programs that were, nevertheless, intended to permit communication with each other’s officers. Most concerning would be a scenario that involves states having no control over their underwater deployed autonomous systems, ruling out the suggestions we have articulated in the above pertaining to decreasing the probability of unintentional war.

In summary, a common understanding of how robots would be programmed by nation states to interpret and respond to their environment will increase the capacity to approximate more precisely, the extent to which LAWS would react proportionately to hostilities. And further, the capacity to undertake the requisite actions in fulfilment of political and military objectives. Such efforts would spur reciprocal understanding and lessen some of the capriciousness of robot-to-robot encounters. The implementation of common guidelines and legal understandings through its most open approach would involve collaboration across the public, private, and non-governmental sectors, and may even use the principles of an open-

⁴³ Mathias Jonas, *How to Use the Term Hydrospatial?*, HYDRO INT’L (Jan. 26, 2021), <https://www.hydro-international.com/content/article/how-to-use-the-term-hydrospatial> [http://web.archive.org/web/20210127101012/https://www.hydro-international.com/content/article/how-to-use-the-term-hydrospatial] (explaining that the term “Hydrospatial infrastructure” could be inspired by the “geospatial infrastructure”).

⁴⁴ Paul Roberts et al., *Communications to the Deepest Point on Earth: Underwater Communication Solution for James Cameron’s Deepsea Challenge*, HYDRO INT’L (Mar. 26, 2013), <https://www.hydro-international.com/content/article/communications-to-the-deepest-point-on-earth> [http://web.archive.org/web/20201021190726/https://www.hydro-international.com/content/article/communications-to-the-deepest-point-on-earth].

source approach, such as observed with the creation and maintenance of internet standards.⁴⁵ A more targeted option, however, would involve military exchanges and dialogue with relevant NGOs. Overall, these essentially non-binding initiatives could lessen the scope of miscalculation, or otherwise provide a means to simplify the different probabilities of how LAWS could, respectively, administer IHL in their own way.

III. IMPLEMENTING RESTRICTIONS ON LAWS IN CONVENTIONAL WARFARE

One benefit of LAWS would be their capacity for traversing regions or conditions beyond human tolerance. If the power source were adequate to fuel the operating system and hardware, and human-assisted maintenance were not required during a particular mission, LAWS could withstand heat and freezing temperatures in desert environments for an extended period, the pressures of deep-sea ocean diving, navigate across high altitudes, or protect satellites through use as a robotic spaceplane across various orbits. For some operations, LAWS may only need to contain munitions specific to the mission parameters; for instance, in an urban environment where close combat is expected, small arms would be more appropriate than artillery.⁴⁶ The dilution of weapons that a machine far more powerful than a person may carry, could assist states to comply with IHL-related elements pertaining to the proportionate use-of-force. Consequently, classifications of LAWS would have a distinctive set of risk management guidelines once consensus is reached on separate categories and a method is devised to safeguard arms sales from surpassing such thresholds. As indicated above, separate articles or protocols could exist under the framework of an overarching arms control agreement to allow for governance and regulations on different types of autonomous weapons. These variations may promote action at first where it is easier to reach an agreement – for example, sparsely populated areas where the risk to civilians is much lower – before more

⁴⁵ Thomas Richard Davies, *Governing Communications*, in *GOVERNING THE WORLD? CASES IN GLOBAL GOVERNANCE* 114–27 (Sophie Harman & David Williams eds., 2013).

⁴⁶ Press, *supra* note 30, at 1360.

complex tasks can be considered and the principle of distinction is more difficult for LAWS to observe.

Upon joining the CCW, states parties agree to “the principle of international law that the right of the parties to an armed conflict to choose methods or means of warfare is not unlimited,” and thus prohibit or restrict “the use of weapons which may be deemed to be excessively injurious or to have indiscriminate effects.”⁴⁷ For example, the protocols of the CCW restrict usage of incendiary weapons, landmines, non-detectable fragments, and blinding lasers.⁴⁸ The CCW promotes disarmament by banning or limiting the use of weapons which may be indiscriminate or cause disproportionate damage or casualties, and it can be revised with additional protocols to “respond to new developments in weapons technology and challenges in armed conflicts.”⁴⁹ From 2014 to 2016, through meetings held under the CCW, the United Nations assembled yearly informal Meetings of Experts on LAWS.⁵⁰ Its original decree was “to discuss the questions related to emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention,” and dialogue encompassing technical issues, ethics and sociology, international law, and military aspects.⁵¹

Since 2017, discussions have continued through the GGE on emerging technologies in the area of LAWS.⁵² The CCW seemed the most plausible setting for a protocol on LAWS, although the Campaign to Stop Killer Robots, a coalition of NGOs, has called on

⁴⁷ UNITED NATIONS, CONVENTION ON CERTAIN CONVENTIONAL WEAPONS 4 (2014), <https://unoda-web.s3-accelerate.amazonaws.com/wp-content/uploads/assets/publications/more/ccw/ccw-booklet.pdf> [<http://web.archive.org/web/20190920084430/https://unoda-web.s3-accelerate.amazonaws.com/wp-content/uploads/assets/publications/more/ccw/ccw-booklet.pdf>].

⁴⁸ *Id.* at 6.

⁴⁹ *Id.* at 11.

⁵⁰ Rosert & Sauer, *supra* note 33, at 20.

⁵¹ Chairperson of the Meeting of Experts, Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), ¶ 1, U.N. Doc. CCW/MSP/2014/3 (Nov. 13, 2014), <https://undocs.org/ccw/msp/2014/3> [<http://web.archive.org/web/20201213013451/https://undocs.org/ccw/msp/2014/3>].

⁵² Rosert & Sauer, *supra* note 33, at 20.

states to launch formal negotiations no later than the CCW's Review Conference in December 2021, warning that "failure to do so would fatally undermine public trust and confidence in the CCW framework's ability to solve this challenge."⁵³ Although decision-making through the CCW theoretically needs to achieve only a simple majority vote (rather than consensus) to create a legally binding instrument, Rosert and Sauer's analysis of the Campaign's strategy finds it has missed "the most straightforward argument," which they suggest should be "not a legal but an ethical one . . . less susceptible to consequentialist counterpositions (which argue that the illegality of LAWS will be remedied by technological progress)."⁵⁴ Based on the possibility of failure to reach an agreement by the end of 2021, two other options – outside of the framework of the CCW – are considered below.

The MTCR has "set an effective international standard" and "international benchmark" for national export control. Its guidelines and technical annex allow for "adherence to a common export policy" and can assist with the obligation on all states to implement UNSC Resolution 1540,⁵⁵ including the adoption of the MTCR standards by the United Nations.⁵⁶ The MTCR is "unique in its focus on the means of delivery" of Weapons of Mass Destruction (WMD) "rather than the WMD themselves," promoting best practices among its membership and, significantly, through outreach activities to non-members.⁵⁷ Since

⁵³ *Convention on Conventional Weapons Policy Brief: Nov. 2020*, CAMPAIGN TO STOP KILLER ROBOTS 2, <https://www.stopkillerrobots.org/wp-content/uploads/2020/05/CCW-Policy-Brief-November-2020.pdf> [<http://web.archive.org/web/20201101021258/https://www.stopkillerrobots.org/wp-content/uploads/2020/05/CCW-Policy-Brief-November-2020.pdf>] (last updated Nov. 2020).

⁵⁴ Rosert & Sauer, *supra* note 33, at 22–23.

⁵⁵ MISSILE TECHNOLOGY CONTROL REGIME, WHO ARE WE AND WHAT DO WE DO? (2020), <https://mtrc.info/wordpress/wp-content/uploads/2020/09/MTCR-Pamphlet.pdf> [<http://web.archive.org/web/20201026231034/https://mtrc.info/wordpress/wp-content/uploads/2020/09/MTCR-Pamphlet.pdf>].

⁵⁶ Jeffrey Taylor, *Public Statement from the Plenary Meeting of the Missile Technology Control Regime, Auckland, Oct. 11, 2019*, MISSILE TECHNOLOGY CONTROL REGIME (Oct. 18, 2019), <https://mtrc.info/public-statement-from-the-plenary-meeting-of-the-missile-technology-control-regime-auckland-11-october-2019/> [<http://web.archive.org/web/20201126231639/https://mtrc.info/public-statement-from-the-plenary-meeting-of-the-missile-technology-control-regime-auckland-11-october-2019/>].

⁵⁷ Taylor, *supra* note 56.

its establishment more than thirty years ago, participation has increased to 35 countries,⁵⁸ including France, Germany, India, Russia, Turkey, South Korea, the United Kingdom, and the U.S.⁵⁹ For a comparison using the same number, the MTCR membership accounts for a large portion of the top 35 arms exporting countries in the world (71%), and most of its value (92%).⁶⁰ Outreach activities are also conducted with non-members, including China and Israel (which rank as the fourth and ninth largest arms suppliers globally).⁶¹

Keeping stride with technical innovations and advancements via frequent updates, the MTCR harmonizes rules without disrupting legitimate trade, and the incoming Chair has indicated the privatization of space programs is an emerging area with “substantial technological overlap between the technology used to transport satellites into space and that used to deliver WMD.”⁶² The MTCR presently defines a threshold on specific weapons, including rockets, unmanned aerial vehicle systems, and associated technologies adept to transporting WMDs or any payload of at least 500 kg (1102 lbs.) up to a distance of at least 300 km (186 miles).⁶³ LAWS, as uninhabited devices, could arguably be perceived to meet these criterion, implying their proliferation already may be constrained under the MTCR. The model, consequently, demonstrates its capacity to tackle the challenges

⁵⁸ MISSILE TECHNOLOGY CONTROL REGIME N.Z., MISSILE TECHNOLOGY CONTROL REGIME NEWSLETTER 2, 4–5, (Sept. 3, 2020), <https://mtrc.info/wordpress/wp-content/uploads/2020/09/MTCR-newsletter-final-.pdf> [<http://web.archive.org/web/20200927020050/https://mtrc.info/wordpress/wp-content/uploads/2020/09/MTCR-newsletter-final-.pdf>].

⁵⁹ Taylor, *supra* note 56.

⁶⁰ SIPRI *Arms Transfers Database*, STOCKHOLM INT’L PEACE RES. INST., <https://www.sipri.org/databases/armstransfers> [<http://web.archive.org/web/20210131000300/https://www.sipri.org/databases/armstransfers>] (last visited Jan. 31, 2021) (evaluating a comparison of the Stockholm International Peace Research Institute’s arms transfers database trend-indicator value of arms exports from the top 35 largest exporters of 2019).

⁶¹ *Id.*

⁶² MISSILE TECHNOLOGY CONTROL REGIME NEWSLETTER, *supra* note 59.

⁶³ *Equipment, Software and Technology Annex*, MISSILE TECHNOLOGY CONTROL REGIME 16 (2019), https://mtrc.info/wordpress/wp-content/uploads/2019/10/MTCR-TEM-Technical_Annex_2019-10-11-1.pdf [http://web.archive.org/web/20201203224358/http://mtrc.info/wordpress/wp-content/uploads/2019/10/MTCR-TEM-Technical_Annex_2019-10-11-1.pdf].

emanating from the acquisition risks of LAWS with larger payloads. Conceivably, this form of informal model of partnership by states, with national responsibility (and control) over the implementation of policy, may be suitable for the regular updates of technical facets, creating best practice, and encouraging state participation on a voluntary basis.

Despite the lack of official verification requirements compared to some other arms control treaties, each state may pass laws or adopt a policy to address breaches, and in essence, give ‘teeth’ to the accord. For example, the President of the United States may impose sanctions to deny contracts or licenses to foreign persons engaging in the trade or facilitation of MTCR equipment or technology, or conspiring to do so.⁶⁴ As an interesting historical parallel, President George H. W. Bush sought China’s commitment to arms proliferation prevention, with Beijing promising to abide by the MTCR Guidelines and Annex in return for the U.S. lifting a ban on the export of satellites to China, a negotiation which then-Foreign Minister Qian considered as marking “the beginning of the lifting of sanctions that had been imposed on China by the United States and other Western countries for two years or more [since 1989].”⁶⁵ Furthermore, the regime promotes accountability because members are unable to undermine one another. Where an MTCR participant adheres to their obligation to not export a banned item, they will need to be communicated with before a comparable item can be sold by another member, with authorizing conditions set out in the Guidelines.⁶⁶ Such a mechanism has significant benefits in promoting rules-based trading and as noted above, is the ‘norm’ amongst the biggest arms suppliers in the world.

⁶⁴ 22 U.S.C. § 2797b (2020); *see also* MISSILE TECHNOLOGY CONTROL REGIME, CTR. ARMS CONT. & NON-PROLIFERATION (2017), <https://armscontrolcenter.org/wp-content/uploads/2017/03/Missile-Technology-Control-Regime-MTCR-Factsheet.pdf> [<http://web.archive.org/web/20200423225135/https://armscontrolcenter.org/wp-content/uploads/2017/03/Missile-Technology-Control-Regime-MTCR-Factsheet.pdf>] (last updated Mar. 3, 2021).

⁶⁵ QIAN, *supra* note 14, at 146–49.

⁶⁶ *Guidelines for Sensitive Missile-Relevant Transfers*, MISSILE TECHNOLOGY CONTROL REGIME, <https://mtcr.info/guidelines-for-sensitive-missile-relevant-transfers/> [<http://web.archive.org/web/20210323003750/https://mtcr.info/guidelines-for-sensitive-missile-relevant-transfers/>].

Whereas the MTCR focuses on WMD proliferation and associated delivery systems, the Wassenaar Arrangement (WA) on Export Controls for Conventional Arms and Dual-Use Goods and Technologies may better suit the nature of LAWS through a nonstrategic weapons ‘prism.’ The WA encompasses the same membership as the MTCR (excluding two MTCR members – Brazil and Iceland), while Mexico and eight additional European countries (Croatia, Estonia, Latvia, Lithuania, Malta, Romania, Slovakia and Slovenia) are members of the WA solely.⁶⁷ Considering some technologies – many of which are now available to civilians – can have their origins traced to inventions for military purposes, applying a framework on dual-use technologies can be deemed a viable approach for managing such a transition over time; although it is essential to qualify that it can also have inadvertent consequences which must be factored into the calculation. Chertoff recommends applying the WA as “a ready platform for the near-term creation of a new export control on LAWS and critical LAWS components” to “reduce the risk of transfer to malicious non-state actors,” but notes the WA had once “accidentally criminali[z]ed many of the necessary tools for stopping malware” until an amendment was later created to address this error.⁶⁸ In comparing the MTCR and WA as optional models for arms control sitting outside of the CCW framework, a distinction is the obligation of members to inform the WA secretariat once they finalize an export which hitherto had been rejected by a different participant. This is in stark contrast to the requirements under the MTCR which calls for consultation between participants prior to the supply of listed armaments.⁶⁹

⁶⁷ *About Us*, WASSENAAR ARRANGEMENT, <https://www.wassenaar.org/about-us/> [<http://web.archive.org/web/20210203042123/https://www.wassenaar.org/about-us/>]; see also Taylor, *supra* note 57.

⁶⁸ Philip Chertoff, *Perils of Lethal Autonomous Weapons Systems Proliferation: Preventing Non-State Acquisition*, 2 GENEVA CTR. SEC. POL’Y 1, 7 (2018), <https://dam.gcsp.ch/files/2y10RR5E5mmEpZE4rnkLPZwUleGsxaWXTH3aoibziMaV0JJrWCxFyxXGS> [<http://web.archive.org/web/20190915220303/https://dam.gcsp.ch/files/2y10RR5E5mmEpZE4rnkLPZwUleGsxaWXTH3aoibziMaV0JJrWCxFyxXGS>].

⁶⁹ *The Wassenaar Arrangement at a Glance*, ARMS CONT. ASS’N (Dec. 2017), <https://www.armscontrol.org/factsheets/wassenaar> [<http://web.archive.org/web/20210126193346/https://www.armscontrol.org/factsheets/wassenaar>].

Given the improbability of constraints on LAWS being negotiated before the CCW's Review Conference is held in December 2021, discussions on the regulation of autonomous systems could advance into a new direction, perhaps incorporating the arms control model of the MTCR or the WA. A key risk is that where robots are dual-use and states are unwilling to place prohibitions on their civilian industries. For example, a statement by Japan in 2020 warned of a "chilling effect" that would "hinder[] technological development and innovation in the civil sectors," should restrictive rules be developed.⁷⁰ Furthermore, ensuring adequate capacity to detect and trace a person in urgent need of being rescued, such as during or following a disaster, is analogous to providing humanitarian assistance that lawfully cannot be denied during wartime. The export lists of the above arms control agreements are further mitigation instruments to lessen the misappropriation of LAWS, provided absolute prohibition is not adopted as a CCW additional protocol.

Certainly, granted the amount of consensus it usually takes to complete arms control agreements, an additional option is for the states adept in researching the highly complex technologies necessary to develop LAWS, to execute a tailored approach that meets current needs, and would seek to expand the membership and rules over time. They could incorporate a sunset clause to incentivize the remainder of the international community to come to the negotiation table, conceivably through the United Nations. An imminent expiry can galvanize action as leaders respond with a greater sense of urgency, such as during President Biden's first phone call with President Putin in January 2021, where both leaders indicated their "willingness to extend New START for five years, agreeing to have their teams work urgently to complete the extension by February 5,"⁷¹ the date the

⁷⁰ Japan, Commentary on the operationalization of the Guiding Principles affirmed by the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems at national level 4 in GGE ON LAWS (Sept. 21, 2020), <https://documents.unoda.org/wp-content/uploads/2020/09/20200828-Japan.pdf> [<http://web.archive.org/web/20200916100055/https://documents.unoda.org/wp-content/uploads/2020/09/20200828-Japan.pdf>].

⁷¹ Press Release, White House, Readout of President Joseph R. Biden, Jr. Call with President Vladimir Putin of Russia (Jan. 26, 2021),

Strategic Arms Reduction Treaty was set to expire. Containing any experiences discovered from an initial agreement on LAWS, a mechanism later developed would also be a useful way to ‘update’ the agreement before its preparation for signature and represents incremental steps toward an ongoing and marked challenge in decreasing, as the title of this article denotes, unintentional war.

CONCLUSION

Developing advanced technology that gives an advantage in the battlefield is an expensive element of military strategy, yet in many ways, it can be more costly to not do so at all. In the case of LAWS, these devices may increase the capabilities of armed forces by countering a greater quantum of threats presented by opponents. Nonetheless, for as long as autonomous systems will remain an untested or low-maturity technology, they would harbor the risk of unintentional attacks, possibly inciting greater volatility. Therefore, the emergence of LAWS has clearly engendered an ongoing political challenge. In the above contemplation, we have suggested an approach for decreasing the dangers emanating from the potential misuse of LAWS. The anticipated result of considerations is the lessening of the probability of unintentional war.

We began with discussion of leadership and security scenario reactions, focusing on the response of newly sworn-in leaders to a potential military crisis. As indicated in the case study, U.S. presidents have faced such incidents with reference to negotiating ‘provocations’ with Chinese leaders early into their first term since the inauguration of President George W. Bush in 2001. While President Biden will certainly not face a decision where it is unclear if an uninhabited device had caused an incident, the exponential growth in AI suggests a Commander-in-Chief may one day face a relatively ambiguous confrontation, perhaps involving LAWS. A thought experiment

<https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/26/readout-of-president-joseph-r-biden-jr-call-with-president-vladimir-putin-of-russia/>
[<http://web.archive.org/web/20210130070210/https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/26/readout-of-president-joseph-r-biden-jr-call-with-president-vladimir-putin-of-russia/>].

modelled on a previous collision elaborated on some of the envisaged changes to diplomacy that may result. One outcome is clear: If the use of LAWS is based on the premise that robots could possess enhanced skills compared to humans, political leaders may find it difficult to provide a justification for machine failure, particularly if attempting to explain the actions committed by such a device as unintentional.

The second section considered enhanced collaboration rules to decrease the probability of unintentional war. Placing an emphasis on non-binding efforts to identify some common interests among parties, including adversaries, proposals were provided in the form of promoting mutual understanding of algorithmic interpretations of the regulations governing conduct in war, as well as protections for humanitarian workers. Furthermore, other areas were contemplated, such as the funding challenges (and potential collaboration opportunities) associated with hydrosatial infrastructure projects that would be required for the use of LAWS in underwater domains.

The final section considered arms control frameworks under which a mechanism on LAWS could be developed. Campaigners are calling for negotiation on LAWS to begin ahead of the CCW's Review Conference in December 2021.⁷² While this timeframe may produce a sense of urgency that provides for more substantive discussions during the next GGE, if this target is not met, it may dent the credibility of the CCW (in the eyes of some participating NGOs) as a forum for responding to emerging technologies. The MTCR, which meets non-proliferation standards recognized by the UNSC and has outreach activities that include even those large arms exporters who are not its members, was proposed as a model designed to promote responsible adherence by participants. The WA, which has broader membership, may also be considered as an arms control mechanism due to its focus on dual-use technologies, which certainly would apply to the fields of AI and robotics.

In examining how these governance options and considerations for regulating LAWS may help states to prepare for an already uncertain future, the principal motivation outlined above has

⁷² CAMPAIGN TO STOP KILLER ROBOTS, *supra* note 54, at 2.

been for states to find a means to collaborate in order to decrease the probability of an unintentional war. The postwar rules-based order of global governance, bolstered by the concert architecture of five designated powers presiding over the veto of the UNSC, gives “authority and responsibility” for “action as it deems necessary in order to maintain or restore international peace and security.”⁷³ While imperfect, this relationship has in many instances generated incentives to pursue increased measures and, as the title of this article denotes, must now contribute to decreasing unintentional war through incremental steps and the greater universalization of arms control mechanisms.

⁷³ U.N Charter art. 51, ¶ 1.