

Vindication for Anthropocentrism: The Humanitarian Case for a Pre-emptive Treaty Ban on Autonomous Weapons

Dylan Edmonds, University of Ottawa, Ottawa, Canada

Abstract

This paper will build on a growing chorus of stakeholder voices in civil society, academia and disarmament forums advocating for a pre-emptive treaty ban on lethal autonomous weapons (LAWs). Fundamentally, this piece's *raison d'être* is to address a disparity in the literature, that being the very little attention dedicated to conceptualizing the most feasible strategy which could lead to a pre-emptive ban. Drawing on International Humanitarian Law's established norms of military necessity, Marten's Law and the emerging norm of meaningful human control, this paper makes a compelling humanitarian case for a treaty ban. Framing theory, a theoretical lens borrowed from Goffman's micromobilization theory, is then invoked to make the ultimate argument that civil society's framing of autonomous weapons as a humanitarian threat is the surest strategy towards a pre-emptive treaty ban.

Keywords

International Humanitarian Law, Civil society, Lethal autonomous weapons, Framing theory

Introduction

As the burgeoning debate over artificial intelligence (AI) gains traction in the public sphere, there has been an endless supply of pundits and scholars offering their take on the actual and projected effects of this emerging phenomenon. Whatever their prognosis, these commentators have focused primarily on AI's potential to further automate and autonomize society in peacetime. On the periphery of these discussions, due perhaps to the popular perception that it is little more than an abstraction, is the ethical debate over the increased interfacing of weapons development and AI research. AI has begun to push the boundaries of weapons' autonomous features, and these autonomous modes are part and parcel of what has been dubbed the Third Military Revolution – after gunpowder and nuclear weapons (Future of Life Institute, 2015). Despite most conservative estimates predicting that fully autonomous Lethal Autonomous Weapons (LAWs), or weapons that if fully realized, can identify and select targets without the involvement of a human operator are still twenty years away, this paper will build on the chorus of civil society, scholarly, citizenry and even military voices pushing for a pre-emptive treaty ban on this nascent technology (The Economist, 2018) (Lin, Abney and Beckey, 2012). Since the United Nations' (UN) Convention on Certain Conventional Weapons' (CCW)¹ now five year-long study of the issue has exposed the geopolitical fault lines preventing any progress towards a

¹ For the very first time, UN member states debated the rise of autonomous weapons systems at the United Nations Human Rights Council (HRC) on May 30th 2013. Deliberations were prompted by an alarming report on the growing relevance of autonomous weapons issued by the UN Special Rapporteur on extrajudicial, summary or arbitrary executions which concluded: "If left too long to its own devices, the matter [of prohibiting autonomous weapons] will, quite literally, be taken out of human hands" (Heyns, 2013, p.21). As awareness of autonomous weapons grew, the Convention on Conventional Weapons (CCW) decided to take up the issue in November of 2013. The CCW examination of autonomous weapons has drawn repeated criticism for its lack of progress brought on by cancelled meetings due to funding aberrations, myopic goal setting, and most recently, unambitious governmental expert meetings.

ban or constructive dialogue, there has been a deepening sentiment that International Law is unable to keep pace with technical innovations.

Fundamentally, this piece's *raison d'être* is addressing a lacuna in the literature, that being the very little attention that has been dedicated to conceptualizing the most feasible strategy which could conceivably lead to a pre-emptive treaty ban. The ancillary argument to this ultimate end is as follows: this paper build a humanitarian case against autonomous weapons by drawing on Professor Ian Kerr's military necessity argument as derived from International Humanitarian Law (IHL)², also known as the Laws of War, and Professor Peter Asaro's claim that these weapons would infringe on the emerging IHL principle of meaningful human control. Once this humanitarian case is adumbrated, this text will assume a theoretical lens known as Framing Theory to argue that civil society is the most suitable vehicle to drive the autonomous weapons' treaty ban movement because of its recent successes in framing disarmament issues as humanitarian concerns, and then crystallizing this humanitarian reframing into treaties with a strong legal footing.

Terminology

Before launching into the humanitarian case for an autonomous weapons treaty ban, certain terminology must be defined. It is crucial to understand what is exactly being said when scholars note that weapons development has progressed from automatic to semi-autonomous, and in the

² Although there are overlaps, International Humanitarian Law is categorically separate from other major fields of International Law, such as International Criminal Law and International Human Rights Law, among others. A definition of International Humanitarian Law, according to the International Committee of the Red Cross, "International humanitarian law is a set of rules which seek, for humanitarian reasons, to limit the effects of armed conflict. It protects persons who are not or are no longer participating in the hostilities and restricts the means and methods of warfare" (ICRC, 2004).

span of two decades at most, we will have the technological prowess to develop fully autonomous LAWs. Automated weapon systems follow a script so to speak, or a set of instructions in software programming whose confines dictate the outcome of all scenarios (The Economist, 2018). Many of today's anti-aircraft defense systems incorporate automatic features; a rather ubiquitous example in the United States (US) military arsenal is the MK 15 Phalanx Close-In Weapons System, which when mounted on navy vessels, follows a programming script to shoot down approaching enemy missiles and aircrafts (Bills, 2014, p.181). On the other end of autonomy's spectrum is a fully autonomous LAWs with general AI. According to the US Defense Science Board (2016), to be a LAW, "[the] system must have the capability to independently compose and select among different courses of action to accomplish goals based on its knowledge and understanding of the world, itself, and the situation" (p.4). These LAWs would be supplied with general AI, which theoretically execute any intellectual task humankind has historically undertaken. Currently, humans have managed to, in only a short time, create partially autonomous weapons with narrow AI. Narrow AI follows algorithms for a prescribed task similar in a way to scripted automation, but narrow AI is a far more dynamic structure enabling machine learning. For now, humans retain control of the lethal faculties of these weapons.

A final specification to be had is that of the further breakdown of general/narrow autonomy. This can be broken down in the following three groupings: man *in* the loop, man *on* the loop, and man *off* the loop. Man *in* the loop LAWs are only capable of acting after human authorization, such as the South Korean SGR-A1 robot which populates the Korean Demilitarized Zone (Bills, 2014, p.182). In the event that an enemy approaches the robot, the SGR-A1 can command the person to put their hands up and surrender. Should the enemy not pay

heed, prior to initiating lethal force, the SGR-AI must receive authorization from a human in the loop. Conversely, man *on* the loop systems do not require human authorization prior to administering lethal force but a human is actively monitoring the weapon's activities. The United States X-47B UCAV drone fits this description in the sense that the drone makes its own decisions, e.g. the route it takes, but always with human oversight (Bills, 2014, p.183). Fully-fledged autonomy, as defined by the full removal of "human controllers from lethal decision-making altogether" (Carpenter, 2016, 58), are fully as defined by a man *off* the loop system, does not yet exist and contradicts the official policies of the P5 countries, who purport they will maintain effective control of the weapons they deploy (Wallach, 2017). This paper is advocating for a ban on autonomous weapons that fall within the man *off* the loop category but remains cognizant of the spectrum-like quality of these categorizations, and would therefore advise that man *on* the loop systems, which do not provide humans the actual possibility of asserting their control, also be proscribed. Man *off* the loop systems and man *on* the loop systems could be powered by either narrow or general AI.

Characteristic to most global issues, where a panoply of stakeholders jockey for their own policy position, definitions are the subject of rancorous debate. The United Kingdom (UK) delegation to the CCW has maintained that autonomy must be defined in its most extreme version as "near humanlike capabilities," whereas the Campaign to Stop Killer Robots, an umbrella civil society organization, defines autonomy in a similar respect to the concept of man in the loop (Wallach, 2017, p.29). Thus, it is impossible to give a legal definition of autonomy at this point in time, and even the relatively innocuous definition provided in this paper could garner some controversy when deciding its applicability to various weapons systems.

The Humanitarian Case for a Pre-emptive Ban on Autonomous Weapons

For those scholars who are either non-committal or who have voiced their outright dissent to a treaty banning autonomous weapons, their reasoning is suffused with a notion of inevitability. To quote two of these scholars: “The certain yet gradual development and deployment of these systems, as well as the humanitarian advantages created by the precision of some systems, make some proposed responses – such as prohibitory treaties – unworkable as well as ethically questionable” (Anderson & Waxman, 2012, p.36). This assumption that a certain technological development is *certain* represents a technological determinism of sorts, where societies become powerless, captured audiences resigned to the technological developments they presumably engineer. From the vantage point of today’s technologically-driven society, the gift of hindsight affords us an alcove where we can make these deterministic assumptions. Indeed, autonomous systems are well on their way to development, but as Peter Asaro (2012), a philosopher and historian of science and technology, points out, it is problematic to make the assumption that deployment proceeds thereafter. Even when technologies do take hold in a society for a short while, this does not assure their persistence. For instance, only a few decades after the introduction of guns in Japan by two Portuguese adventurers in 1543, the country was home to the most preeminent gun technology in the world. Soon after this technological diffusion, the Japanese government, bolstered by a nationalist movement, restricted foreign products, including guns, until Japan was nearly gun-less (Diamond, 1999). In short, though there is a high likelihood that autonomous weapon systems will eventually be deployed, this does not equate to inevitability as national weapons reviews, existing IHL norms and a possible pre-emptive ban all pose existing and hypothetical hurdles.

Some scholars, such as Robert Arkin, are not only convinced of the inevitability of autonomous weapons' deployment, but consider it to be probable that these weapons will respect the foundational principles of IHL, namely discrimination and proportionality. Discrimination requires the capacity during military operations to distinguish non-combatants from combatants, while proportionality requires that the damage caused through the use of force in a military operation be commensurate with the "strategic importance of the mission" (Kerr & Szilagyi, 2016, p.343). Arkin offers a unique perspective to this debate as he heads a lab on the frontier of AI research, which seeks to test autonomous weapon systems in war-like scenarios. In a 2009 book he stated that if autonomous systems were properly inculcated with an ability to adhere to IHL, they could outperform human soldiers with respect to humaneness (Arkin, 2009). More recently, Arkin (2011) made the claim that due to this superior adherence to IHL, "if the technology is developed properly it could potentially lower non-combatant casualties" (p.10). On the face of it, these are indictments against any argument for a pre-emptive treaty ban. Certainly, autonomous weapons would face glaring challenges in superseding humankind's ability to respect these principles. Discrimination is made even more complicated in modern, urban warfare where non-state actors are increasingly entangled in protean, vicious armed conflicts. As for the principle of proportionality, IHL is equally nuanced. The International Committee of the Red Cross' guidelines for determining whether a civilian is a legitimate target are harrowingly complex and include three requirements that must be met: threshold of harm, direct causation and belligerent nexus (Asaro, 2012).

Whether the difficulties in meeting these IHL norms are surmountable remain to be seen, but this surely constitutes a glaring weakness of this presumptive argument. Notwithstanding

these assumptions, the debate of whether autonomous weapons could meet or exceed human adherence to IHL is a red herring and, aside from much speculation, does little to advance the cause for or against a treaty ban. Asaro (2012) succinctly summarizes the inherent digression in this debate in saying that:

Given the complexity of these systems, and our inability to foresee how they might act in complex operational environments, unanticipated circumstances, and ambiguous situations, there is a further difficulty – how we can test and verify that a newly designed autonomous weapon system meets the requirements imposed by IHL. (p. 692-693)

The argument for autonomous weapons' ability to faithfully follow IHL's cornerstone values lacks the empirical basis it requires to be sound. Thankfully for the proponents of a treaty ban, this paper will outline a humanitarian appeal that does not fall victim to the same empirical paucity. At best, pursuit of this speculative argumentation broaches, albeit tangentially, the ethical debate that rages at the heart of LAWs – that being whether human decision-making and responsibility for lethal violence should be delegated to non-human entities. With this transition from a *can*: can LAWs adhere to the principles of IHL, to a normative *ought*: ought it be acceptable for LAWs to initiate lethal force without human oversight, the debate over LAWs regains the anthropocentric lens it so desperately requires. The time is now fortuitous to turn to the humanitarian arguments which address this crucial question, ones which reemphasize humankind, in both its combatant and non-combatant forms, as the overriding consideration in any disarmament discussion.

Although we cannot know for sure if LAWs would be superior followers of IHL than humans, we can determine beyond a reasonable doubt that their availability would almost certainly be a force multiplier for the IHL principle of military necessity. Put bluntly by Professor Pagallo (2011), “[i]f we have expendable mechanical mercenaries that we can send into battle

instead of our children, how could we not do so?” (p.303). According to Nils Melzer (2009), the UN Special Rapporteur on Torture and other Cruel, Inhuman or Degrading Treatment or Punishment, military necessity can be defined as “only that degree and kind of force, not otherwise prohibited by the law of armed conflict [IHL], that is required in order to achieve the legitimate purpose of the conflict” (p.79). As a disruptive technology, the existence of LAWs, without adequate legal devices to curtail their usage, would expand the “scope of inclusion for what can prospectively be perceived as ‘necessary’” (Kerr & Szilagyi, 2016, p.359). Even the most cursory glance at the history of military technologies trend towards this disturbing pattern – once a new military technology is included in the arsenal of choice, the perspective on what is deemed militarily necessary to reaching a stated objective shifts even if this technology brings greater casualties (Kerr & Szilagyi, 2016, 363).

An illuminating example of this is the introduction of the submarine. As was custom in International Maritime Law, the vanquisher was obligated to rescue the vanquished by bringing them aboard at the end of a battle (Kerr & Szilagyi, 2016).³ This humanitarian approach to regulating war at sea was impaired by the introduction of submarines whose small cabin spaces could not accommodate enemy crew. To borrow an economics’ term, the submarine’s competitive advantage is in its stealth – a quality that would no longer be maximized if it was forced to surface (Kerr & Szilagyi, 2016). The usage of submarines proliferated in the Second World War, and in order to ensure the submarine’s strategic potential, the notion of rescuing sinking enemy crew was soon forgotten. As submarines became deeply entrenched in the naval strategies of all the great powers, the perspective which was manifested that despite the added

³ This custom was even codified in the 1930 and 1936 London Naval Treaties.

casualties which submarines wrought, these were an indispensable weapon. In other words, they become a military necessity and force multiplier. Tragically, with the sinking of this humanitarian custom of saving overboard crew, countless drowned at the hands of submarines. This is but one example of how the malleable principle of military necessity shifted with the inclusion of a new technology, and as a result, humanitarianism at sea was derogated.

With prospective LAWs raising the humanitarian stakes, recourse can be found in the Martens clause, IHL's explicit reference to its moral origins. First appearing in the 1899 Hague Convention II on the Laws and Customs of War on Land in 1899, it states that:

In cases not covered by this Protocol or by other international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established custom, from the principles of humanity and from dictates of public conscience. (Human Rights Watch, 2012, section 2, para.13)

The International Court of Justice has gone so far as to connect the Martens Clause with the proliferation of weapons technology by asserting that this customary law can be an effective preventative measure against growing lethality (Human Rights Watch, 2012). Central to this paper's first argument is that the fact that Martens Clause renders a powerful acknowledgement of "the principles of humanity" and "the dictates the public conscience" as legitimate sources of new IHL (Human Rights Watch, 2012, section 2, para.13). What emerges then from this is a more visible role for civil society, the traditional champion of humanitarian causes, in spearheading movements aimed at assessing the alignment of weapons to IHL's moral underpinnings.

Despite its notorious ambiguity, to arrive at any sort of semblance of the public conscious, various actors from public forums such as civil society must actively participate in disarmament conferences. Civil society groups and coalitions, such as The Campaign to Stop

Killer Robots, are well-suited to identify the existing and emerging norms and dictates of public conscious as many of these groups work at the grassroots level, reifying individual and stakeholder perspectives. However, as alluded to above, the Martens Clause is notoriously amorphous. Human Rights Watch, one of the first organizations to make the case for an autonomous weapons ban, was cautious in relying too heavily on the Martens Clause simply because of its subjective quality. For example, Professor Peter Assaro (2016) warns that one must not equate public conscious to public opinion, and cautions against the manipulative ways in which public opinion is typically solicited. Thus, the reader is left with the vague impression that perhaps a recent letter written by Canadian AI researchers and technology scholars entreating Prime Minister Justin Trudeau to declare Canada's support for an autonomous weapon's ban would be more constitutive of the public conscious than adding a question about these weapons on the next census.⁴ Still, with all of Asaro's warnings against the legitimacy of public opinion, we are left no closer to defining public consciousness. Even questions over the exact nature of the principles of humanity persist, as Asaro rejects the argument that these are contained in the UN Declaration of Human Rights, while others concur that in fact they are. These definitional issues aside, the Martens Clause, which is a crucible of the destructive Napoleonic Wars, is testament to the legal weight humanitarian ideals and the need for civil society contribution in not only defending but devising new IHL. It is from this clause that a new IHL norm, one

⁴ Published on November 2nd 2017, this open letter was penned by five leading Canadian researches of AI including Professor Ian Kerr, Canada Research Chair in Ethics, Law and Technology, University of Ottawa whose research into the intersection of autonomous weapons and the law was essential to this paper's manifestation. While the letter does implore the Canadian government to take a firm global stance against autonomous weapons, it does acknowledge the "remarkable investment" that Canada is making to assure its place as a leader in AI research and innovation.

propounded by the Campaign to Stop Killer Robots and several member states of the CCW, is beginning to take shape.

Throughout its five-year existence, the Campaign to Stop Killer Robots has touted the position that LAWs would contravene the incipient IHL norm of meaningful human control, which they claim is required when initiating lethal force. One of the ways they have broadcasted this goal to various audiences is by releasing blog posts on their website championing countries such as Pakistan and the Nordic states for explicitly including meaningful human control in their arguments for a treaty ban at any given CCW meeting on the subject (Stop Killer Robots, 2017). Even those countries that have not voiced their support for a treaty ban have converged as early as May 2014 around the shared belief that meaningful human control must remain in place. This fundamentally humanist principle is a natural corollary to the Martens Clause, and though it remains to be seen if states in the long run will feel compelled to abide by it, the fact that there is at minimum consensus amongst a majority of the world's states that humankind must exert some sort of control over LAWs is a positive development. Though this control has yet to be defined, the connotation that the descriptor *meaningful* affords is again positive for those advocating in favour of a ban.

In essence, the emerging meaningful human control norm, which was even italicized for emphasis in the aforementioned letter written by some of Canada's leaders in AI research, answers in the negative the normative question of whether LAWs should be a final arbitrator in life and death situations. Stakeholders in favour of a ban share in the "widely held belief that the law at least aims to express shared norms, and that many legal norms coincide with moral norms" (Asaro, 2016, p.271). Since IHL is a protean system, there is an opportunity to

promulgate this emerging norm until it is fully inculcated as a custom to which states abide. Put differently, in the new moral space that a disruptive technology like LAWs precipitate, there is potential for the international community's beliefs and actions to coalesce around the behavioral expectation of meaningful human control (Asaro, 2016). If common ground is achieved, a treaty ban would be in easier reach, since one of the proximate causes for new law is "the emergence of new behaviours by states or non-state actors" (Asaro, 2016, 371). Since the Campaign to Stop Killer Robots has been so central in propounding this emerging principle, the coalition will be crucial going forward in any efforts to substantiate meaningful human control as a norm worthy of customary law status.

A caveat is in order: this customary status need not be considered *Jus Cogens*, or a fundamental principle of International Law, to be a powerful influencer on state behaviours. There are, of course, other scenarios which lead to the creation of new international law – inconsistencies in the law or the emergence of new technologies that alter state behaviours in such a profound manner that the assumptions of customary or treaty law are violated – one or both of which might be salient to LAWs. The emergence of new norms which shape actors' behaviour, e.g. meaningful human control, has, in antecedence to the actual existence of LAWs, been demonstrably externalized in the discourses of state and non-state actors alike. This antecedent quality is unique because the other proximate motivations for creating new law (e.g. the actual emergence of technologies which disrupt affairs in real time and inconsistencies in existing law) would only be triggered after LAWs are immediately accessible. Having demonstrated the power that this emerging norm has to hasten the codification of treaty law, the

Campaign to Stop Killer Robots, as the leading purveyor of the norm, will be crucial to any successful strategy for a pre-emptive ban.

At this point, it would be useful to explore other ways in which civil society has shaped diplomatic circles' decision-making on autonomous weapons. Civil society has only been able to propound emerging norms, such as meaningful human control, because it successfully appropriated the science-fiction tropes which had historically left autonomous weapons off the humanitarian agenda. Prior to 2012, autonomous weapons were subject to "the giggle factor" due to its close resemblance to science fiction themes depicted in popular media, and this meant that the topic was not a contender for advocacy according to humanitarian elites (Carpenter, 2016, 57). In a 2010 interview, a Human Rights Watch Officer surmised the general sentiment of civil society towards autonomous weapons in saying: "I don't think there's much of a taste for being too forward leaning on science fiction if I may be blunt. The emphasis is definitely on existing state practice, not on laboratory weapons that are unproven" (Carpenter, 2016, 59). Subsequently, and in a matter of two years, Human Rights Watch published "Losing Humanity" a substantiated manifesto detailing the first complete argument in favour of an autonomous weapons ban. This report was the crucible of two pivotal years wherein mounting, incontrovertible documentation of drone casualties combined with real-world research into growing autonomy and the effects of growing automation spurred civil society into action (Carpenter, 2016). To overcome the science-fiction stigma surround LAWs, civil society organizations like Article36 or Human Rights Watch advanced very technical and legalistic arguments in favour of a ban. In other words, their arguments were diametrically opposed to the futuristic quality of science-fictions.

However, civil society was not averse to the association of LAWs' with science-fiction, especially if this stigma could be manipulated to generate public attention. By adopting the rhetoric of "killer robots", a terminology with origins in popular science-fiction culture, civil society's inflammatory word choice compounded media attention- it is perhaps of little coincidence that the CCW, pressured by the significant spotlight LAWs were garnering, held their first session on LAWs in 2013 (Carpenter, 2016). Civil society has carefully constructed a rhetoric that is both scientifically rigorous, yet tactfully employs science fiction discourse. Thus, the science fiction label was once a disabler, but when wielded effectively, civil society has used it as a sort of "social lubricant" (Carpenter, 54).

Charli Carpenter (2016), a security analyst who interviewed members of civil society extensively on LAWs, captures the rhetorical power in science fiction tropes in the following quote from a civil society member attending a CCW session: "When I first attended a multilateral meeting with a 'killer robots' bumper sticker on my computer, diplomats would walk past and do a double-take then sit down and say: 'Tell me more, I want to know more about these killer robots.' It's the perfect hook to get a conversation started" (63). Thus, civil society, by selectively abstaining from invoking allusions to LAWs' futuristic elements, positioned itself well to appeal to policy wonks and government bureaucrats through legal and scientific argumentation. However, to generate the necessary media attention for these issues to appear on government policy radars, civil society did not shy away from employing science fiction tropes in its public outreach. Without civil society to undertake this two-pronged task of mobilizing public attention and targeting state diplomats, it is perhaps reasonable to conject that the autonomous weapons debate would have yet to reach the UN.

Civil Society's Humanitarian Framing: The Vehicle to a Treaty Ban

Hitherto, this paper has supported an outright ban on autonomous weapons by outlining the major arguments (i.e. that autonomous weapons are an infringement to the principle of military necessity, the Marten clause, and the emerging norm of meaningful action) which engender the humanitarian ethos in favour of a ban on autonomous weapons. To consolidate this first argument, this paper will now draw upon Framing Theory to propound the claim that the strategy spearheaded by civil society groups, of analyzing disarmament issues through a humanitarian lens, has the greatest likelihood of achieving a treaty ban. Framing Theory is derived from the work of sociologist Erving Goffman on micromobilization, which refers to the multitude of interactive and communicative processes that affect the frame alignment (Gamson, 1985). The *frames* of framing theory are best described as the systems of interpretation which result from these multifaceted, and countless interactions between actors (Borrie, 2014). In the context of weapons disarmament, scholars have cited micromobilization amongst social activists as eventually reshaping the discourse on various, pernicious weapons such as landmines and cluster munitions. The eventual result of this reframing, and the strongest indicator of its transformative ability, are examples of state policy makers reversing their long held beliefs on said weapons.

Framing Theory must not be confused with Thomas Kuhn's more radical concept of a paradigm shift, which he applied to the hard sciences but which has been appropriated by social scientists as well. Examples of a paradigm shift are the scientific revolutions ushered in by the supersession of the Ptolemaic system by Copernicus' Heliocentric model of the universe, or Einstein's publishing of the Theory of Relativity. In both these instances, radical changes in even the definitions of the most basic concepts were witnessed, which prompted entirely new methods

of conducting research. In the frame alignment process that this paper seeks to employ, the framing variant in question pertains to “a domain seen as normative and acceptable is reframed as an injustice that warrants change” (Borrie, 2014, p.636). In short, the change expressed in the domain-specific framing is less expansive and extreme than the paradigm shift, and incorporates a uniquely held normative aspect. As was noted following the reframing of cluster munitions debate in the early 21st century, even after the framing occurred, “states continued to act as states often do, and this was especially visible in the bargaining behavior on draft [Convention on Cluster Munitions] provisions during its negotiation” (Borrie, 2014, p.642). Unlike the paradigm shift, the reframing, which occurred in a transitory time and space, did not result in any real fundamental alteration in state behaviour, and even now, many states have yet to ratify the treaty’s contents.

When weighing civil society’s framing strategy against other possible strategies, such as the conventional practice of letting states deliberate disarmament issues at the CCW, a short analysis of recent disarmament history elucidates that the former is a far more effective option than the latter. That is not to say that states should not debate autonomous weapons at the CCW – in fact quite the opposite as UN fora are ideal grounds for civil society to disseminate their humanitarian appeals. However, taken alone, allowing states sole proprietorship in the drafting of a treaty is an impoverished and myopic approach. After three years of informal talks, it seemed the CCW was beginning to take substantive steps towards a treaty when it slated a Group of Governmental Experts for two weeks of meetings, one in August and the other in November of 2017. The panel’s first meeting date was cancelled due to a shortfall in member payments, and the most recent meeting in November ended with the unambitious decision to roll the 2017

mandate into 2018 (Stop Killer Robots, 2017). Even for the UN, where progress is measured in glacial terms, the discussion about autonomous weapons has languished in the CCW. States seem hard-pressed to relinquish their hard security mindset. In July 2017, China produced its Next-Generation Artificial-Intelligence Development Plan, which designates AI as the transformative technology ensuring economic and military pre-eminence. In September of the same year, Vladimir Putin told Russian children returning to school that “artificial intelligence is the future, not only for Russia but for all of mankind...whoever becomes the leader in this sphere will become the ruler of the world” (quoted in *The Economist*, 2018, para. 17). It is with these provocative statements in mind that we turn to a brief history of civil society’s success in tempering naked pursuit of national supremacy, by reframing several disarmament issues as ultimately humanitarian ones with substantive results to show.

Injecting humanitarianism in disarmament diplomacy has correlated with the outright proscription or regulation of various weapons systems by way of binding treaties. Humanitarianism’s place in disarmament diplomacy was cemented in the 1997 Ottawa Treaty banning landmines, the 2008 Convention on Cluster Munitions, and more recently in the 2017 Treaty on the Prohibition of Nuclear Weapons. These treaties were spurred by the efforts of civil society coalitions, including the International Campaign to Ban Landmines (ICBL), the Cluster Muniton Coalition (CMC) and the International Campaign to Abolish Nuclear Weapons (ICAN) to redraw landmines, cluster munitions and nuclear weapons respectively as devastating humanitarian threats. Early calls for banning landmines and cluster munitions in the 1970s focused on the unnecessary suffering cluster munitions cause combatants and largely ignored the humanitarian consequences of these weapons on non-combatants (Borrie, 2014). The ICBL

began the process, prior to the drafting of the 1997 Ottawa Treaty banning landmines, of reframing landmines as a pernicious humanitarian threat to civilians.⁵ Buttressing their arguments with data-gathering in the field, the International Campaign to Ban Landmines formed a sort of epistemic community by drawing a clear picture of just how ordinances were afflicting civilians' lives (Borrie & Randin, 2006). To aid their cause, survivors were enlisted as ban advocates, and just as the Campaign to Stop Killer Robots applauds humanitarian-minded states and shames those states that are less so, the ICBL was adept at rewarding and shaming states (Borrie, 2014). These practices culminated in the Ottawa Treaty, whereby the number of allied states (principally small and medium sized progressive states and many from the Non-Aligned Movement) noticeably grew. Though several of the world's superpowers refuse to ascend to the Ottawa Treaty, there is a widespread understanding that the humanitarian norms it has codified have universally proscribed any usage of landmines.

In subsequent years, as the CMC launched its humanitarian appeal against cluster munitions in 2003, those policy makers most susceptible to the ICBL's reframing efforts some ten years previously were the first the CMC lobbied. Rather than duplicating previous work, the CMC borrowed from the ICBL's strategies in constructing its own parallel humanitarian arguments against cluster munitions. In the language of Framing Theory, this is referred to as "frame bridging" (Borrie, 2014, p.638). On the coattails of the Ottawa Treaty, non-state actors had succeeded in reframing cluster munitions which "helped to empower champion states in terms of delineating the issues to be solved and identifying a plausible way forward towards

⁵ The banning of landmines, for example, encompassed a multi-step effort in the form of the Ottawa Treaty and its flawed antecedent, the 1996 CCW Amended Protocol II. Many civil society groups and states were unsatisfied by the Amended Protocol II since it did not categorically ban landmines. The Ottawa Treaty of 1997 realized a total ban on landmines- an unimaginable feat to the state parties who drafted the CCW amendments a year prior.

meaningful action” (Borrie, 2014, p.635). From 2003 to 2007, the CMC built linkages with friendly small and medium state powers such as Norway in an effort to circumvent the CCW’s stalled cluster munitions talks (Hulme and McGoldrick, 220). Eventually, Norway agreed to sponsor the Oslo Process, and in February 2007, 46 states agreed to the Oslo Declaration which established a timeline of just under two years to realize a treaty banning cluster munitions (Hulme and McGoldrick, 220). The CMC was instrumental in including a reference in the declaration to humanitarian ideals through the inclusion of a statement which affirms that these weapons “cause unacceptable harm to civilians” (clusterconvention.org).

The CMC movement culminated in a treaty ban, this time in 2008, and though there remains non-signatories, even those states most heavily addicted to planting landmines and cluster munitions have substantially curtailed these now stigmatized practices. For example, in March 2009 the US Congress passed a one-year moratorium on the transfer of cluster munitions with a failure rate of less than 1%, which represents the majority of the American arsenal (Garcia, 2012). The Obama administration had decided that after 2018, the US will only transfer and deploy munitions containing less than 1% unexploded ordnance (Garcia, 2012). The Trump administration has since reversed this policy of only deploying munitions with a lower than 1% failure rate, but retained the restriction that their deployment must first be cleared by the combat commander (Arms Control Association, 2017). In short, even the world’s greatest military power, and a non-signatory of the cluster munitions treaty, has altered its behaviour in the face of the treaty’s normative force.

Even five years ago, it might have been absurd to many to attempt to transpose this humanitarian logic from landmines and cluster munitions treaty making (which are similar in

some notable respects) to weapons that are so intimately linked to great power struggles, such as nuclear bombs. However, just last year, in a direct rebuke to the stratified nuclear world order perpetuated by the Non-Proliferation Treaty, the International Coalition to Abolish Nuclear Weapons (ICAN) successfully drafted the Treaty on the Prohibition of Nuclear Weapons (TPNW). The majority of the TPNW's support comes from the Non-Aligned Movement, whose ambassador to the CCW mentions in the same breath the Movement's support for an autonomous weapons ban and the TPNW on humanitarian grounds (UN General Assembly, 72nd Session). ICAN campaigners successfully indulged in frame bridging as they explicitly linked their call for a nuclear ban treaty with other weapons such as landmines and cluster munitions that are deemed to have unacceptable effects on civilian populations (Borrie, 2014). Those campaigners who were involved in the Oslo processes for cluster munitions persuasively remarked that, much like the reasoning behind the cluster munition ban, nuclear weapons would saturate large surface areas with explosive force, thereby imperilling civilian lives. All this to say that the detractors of a pre-emptive ban, who rightly argue that autonomous weapons are a unique case in point (mainly because they are not a specific weapon but rather a weapon system/modality and that they do not yet exist), are wrong to infer that these differences nullify any prospect of a ban. Humanitarian concerns for the wellbeing of civilians and combatants alike are ubiquitous, no matter the weapon or weapon system in question. Though the TPNW seems to have added to the polarization which defines nuclear armed countries and the non-nuclear majority, the more NPT nuclear weapon states and its allies dismiss humanitarian concerns, the more they cause a feedback loop where these dismissals contribute to civil society's framing of the issue as driven by a few problem states who are unpreoccupied with humanitarian ideals (Borrie, 2014). The same

could be expected with LAWs, where presumably the most advanced powers would control the most advanced autonomous weapons. Thus, there is hope for humanitarian framing to precipitate a treaty even if this treaty was not pre-emptive.

Conclusion

This text wishes to humbly express a warning to the proponents of a humanitarian ban for autonomous weapons. Such movements can fall prey to infighting and can be equally impassive when faced with adversarial states. Going forward, it will be essential that the Campaign to Stop Killer Robots and other stakeholders remain abreast on research on LAWs and receptive to some of the potentially positive effects of their existence. For instance, defensive autonomous systems will surely not be as threatening as those used for offensive combat. Already, these defensive technologies are very advanced, with some examples being the U.S. Aegis Ballistic Missile Defense System and Israel's Iron Dome. Disarmament issues are inextricably linked to states' jealously guarded assertions of sovereignty, and it is more than likely that any pre-emptive treaty wishing to ban autonomous features on defensive systems would be a sticking point. However, the first humanitarian argument raised by this paper, against the existence of offensive autonomous weapons systems that fall within the ambit of man *on* the loop without veritable human control and *off* the loop systems, remains intact. As was outlined, this humanitarian approach is derived from existing IHL in the form of both the military necessity principle and the Martens Clause, and what many scholars now consider to be an emerging IHL norm of meaningful human control. What followed was the upholding of the humanitarian argumentation as the best strategy for achieving a pre-emptive ban. Drawing on framing theory to understand the recent shift towards a humanitarian lens in disarmament fora, this paper conducted a brief

analysis of the consecutive successful attempts of civil society campaigns at building humanitarian arguments and borrowing best practices to galvanize the traditionally state-dominated treaty making process. Guiding this second argument is the old adage posing the rhetorical question of why reinvent the wheel. If humanitarian reframing has been so persistently successful for the making of treaty bans that cover vastly different weapons, would current stakeholders in the age of autonomy not be remised to heed to this accumulated wisdom?

References

- Anderson, K. & Waxman, M. (2012). Law and ethics for robot soldiers. (Report). *Policy Review*, 176, 35.
- Arms Control Association. (2017). *Cluster Munitions at a Glance*. Retrieved from <https://www.armscontrol.org/factsheets/clusterataglance>
- Arkin, R. (2009). *Governing Lethal Behavior in Autonomous Robots*.
- Arkin, R. (2011). Governing Lethal Behavior in Robots [T&S Interview]. *Technology and Society Magazine, IEEE*, 30(4), 7-11.
- Arkin, R., Ulam, P. & Wagner, A. (2012). Moral Decision Making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust, and Deception. *Proceedings of the IEEE*, 100(3), 571-589.
- Asaro, P. (2016). Jus nascendi, robotic weapons and the Martens Clause. In Calo, M., Fromkin, M. & Kerr, I. (2016). *Robot law*. Cheltenham: Edward Elgar Publishing.
- . (2012). On Banning Autonomous Lethal Systems: Human Rights, Automation and the Dehumanizing of Lethal Decision-Making. *International Review of the Red Cross*, 94(886), 687-709. Retrieved from <https://www.icrc.org/eng/resources/documents/article/review-2012/irrc-886-asaro.htm>
- Bills, G. (2014). LAWS unto themselves: Controlling the development and use of lethal autonomous weapons systems. *George Washington Law Review*, 83(1), 176-208.
- Borrie, J., Randin, V. M. (2006). *Disarmament as Humanitarian Action: From Perspective to Practice*. Geneva, Switzerland: United Nations Institute for Disarmament Research.
- Borrie, J. (2014). Humanitarian reframing of nuclear weapons and the logic of a ban. *International Affairs*, 90(3), 625-646.
- Carpenter, C. (2016). *Rethinking the Political / -Science- / Fiction Nexus: Global Policy Making and the Campaign to Stop Killer Robots*. *Perspectives on Politics*, 14(1), 53-69. doi: 10.1017/S1537592715003229
- Clusterconvention.org. (2008). *Oslo Conference on Cluster Munitions 22-23rd February 2007 Declaration*. Retrieved from <http://www.clusterconvention.org/files/2012/11/Oslo-Declaration-final-23-February-2007.pdf>.

- Defense Science Board. (2016). *Report of the Defense Science Board Summer Study on Autonomy*. US Department of Defense. Retrieved from <https://www.hsdl.org/?view&did=794641>
- Gamson, W. (1985). Goffman's legacy to political sociology. *Theory and Society*, 14(5), 605-622.
- Garcia, D. (2012). "Banning Evil: Cluster Munitions and the Successful Formation of a Global Prohibition Regime." *Cornell International Affairs Review*, 5(2). Retrieved from <http://www.inquiriesjournal.com/a?id=1244>
- Human Rights Watch & Harvard Law School's International Human Rights Clinic. (2012). *Losing Humanity: The Case Against Killer Robots*. Bonnie Docherty. Retrieved from <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>
- International Committee of the Red Cross. (2015). *International Humanitarian Law: Answers to Your Questions*. Retrieved from <https://www.icrc.org/en/publication/0703-international-humanitarian-law-answers-your-questions>
- Kerr, I. & Szilagyi, K. (2016). Asleep at the Switch? How Killer robots become a force multiplier of military necessity. In Calo, M., Froomkin, M., Kerr, I. (2016). *Robot law*. Cheltenham: Edward Elgar Pub.
- Lin, Patrick, Abney, Keith, and Beckey, George. (2012). *Robot Ethics*. Cambridge: MIT University Press.
- Melzer, N (2009). Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law, *International Committee of the Red Cross*. Retrieved from <https://www.icrc.org/en/publication/0990-interpretive-guidance-notion-direct-participation-hostilities-under-international>
- The Economist. (2017, January 25). Special Report: Getting to grips with military robotics. *The Economist*. Retrieved from <https://www.economist.com/news/special-report/21735478-autonomous-robots-and-swarms-will-change-nature-warfare-getting-grips>
- Pagallo, U. (2011). Robots of Just War: A Legal Perspective. *Philosophy & Technology*, 24(3), 307-323.
- Sartre, P. & Hosotte, O. (2012). New publications in humanitarian action and the law. *International Review of the Red Cross*, 94(887), 1173-1187.
- The Future of Life Institute. (2015). *Autonomous Weapons: An Open Letter From AI & Robotics Researchers*. Retrieved from <https://futureoflife.org/open-letter-autonomous-weapons/>

United Nations General Assembly, 72nd Session. First Committee. *Statement by H.E. Ms. Ina H. Krisnamurthi*. New York, 2017.

Wallach, W. (2017). Toward a ban on lethal autonomous weapons: Surmounting the obstacles. *Communications of the ACM*, 60(5), 28-34.

Dylan Edmonds graduated from the University of Ottawa in April 2017 with a Specialization in Political Studies. Next Fall, he will be pursuing joint civil and common law degrees at the McGill Faculty of Law where he hopes to continue studying and building a career in International Humanitarian Law, Military Law and Weapons Review policy.