SGIA2341: Research Project – Lethal Autonomous Weapons Systems

Z0161906

‘Is the hard governance of the development of Lethal Autonomous Weapons Systems bankrupt?’

Word Count: 4952/5000.

Background and Introduction

Scholars widely regard the likely advent of Lethal Autonomous Weapons Systems (LAWS) as the 'third revolution in warfare,' following after gunpowder and nuclear weapons (FLI, 2022). Accordingly, should LAWS be realised, their effect on warfare and international security paradigms would be profound (Gill, 2017, p. 174; Altmann & Sauer, 2017, pp. 120-122; Sparrow, 2016, p. 110). Consequently, the regulation of LAWS development (and eventual deployment) constitutes much of the academic literature. Before continuing, outlining what actually constitutes LAWS is a definitional prior of the paper. The broad consensus is that a LAWS is: 'a weapon system that, once activated, can select and engage targets without further intervention by a human operator' (Schmitt, 2013, 5).

The current literature on LAWS regulation is rather dichotomous. Essentially, LAWS adversaries espouse LAWS' inhumanity, illegality and threat to international security paradigms (Rosert & Sauer 2021; SKR, 2021; FLI 2022). Comparatively, there is a school of thought that regard LAWS as offering a means of saving countless military personnel when applied correctly (Anderson, 2016, p. 17). Prominent scholars, NGOs and tech moguls voicing these opposing opinions are reintroduced and expanded upon in the forthcoming section.

In line with the United Nations (UN) Convention of Certain Conventional Weapons (CCW), 'human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines,' deeming the deployment of LAWS banned under International Humanitarian Law (IHL) (GGE, 2019, *Principle 1 & 2*). Furthermore, 'the development […] of a new weapon […] determination must be made whether its employment, in some or all circumstances, be prohibited by international law' (Ibid., *Principle 5*). By extension, the development of LAWS is de jure prohibited. In conjunction with this UN ruling, many NGOs and technology titans fervently support the hard governance of 'slaughterbots,' that is governance that operates through 'rules that arise from treaties, directives and regulations' (SKR, 2021; Maggetti, 2015, p. 252).

Despite this, some scholars – principally, Moon & Lin, Horowitz, Maas, Gill, Anderson & Waxman – argue that the hard governance of LAWS has limited efficacy. Instead, they regard a more norms-based form of soft governance as the most appropriate means of regulating LAWS.

To the layman, these dichotomous positions would cause much puzzlement. It is the objective of this paper to contribute to the debate over LAWS regulation from a practical standpoint and assess the extent to which hard governance against the development of LAWS is, in fact, bankrupt. In doing so, I shall echo and expand upon the position of the above scholars. Importantly, this paper shall not consider the ethicality, nor the most appropriate means of regulating LAWS.

In a word, I shall analyse how the hard governance of LAWS will be limited by futility of regulating the constituent components of autonomous technology. I shall discuss issues of regulating dual-use civilian technologies, such as Artificial Intelligence (AI) and open-source software, exacerbated by the nebulous, digital platform upon which they (mostly) exist.

In consequence of these factors, this paper concludes that the hard governance of LAWS is unlikely to prevent their development and eventual deployment. However, I argue that this

likely reality is comparable to conventional criminal activity, in the sense that the law is as good as those who follow it. In saying this, I mean to stress that the implications of their illegal development and deployment could be more damaging than anything in the history of warfare, markedly so. Therefore, I offer brief support of scholars including Horowitz, Maas and Gill in suggesting that greater transparency at all levels of military operation be encouraged to assist in the containment of the destabilising forces that LAWS threaten to unleash.

Research Context and Literature Review

International resistance to the development and deployment of LAWS is spear-headed by Campaign to Stop Killer Robots (SKR), International Commission of the Red Cross (ICRC) and International Panel on the Regulation of Autonomous Weapons (iPRAW). The SKR voice the international coalition supporting hard governance of LAWS within the UN CCW, whilst iPRAW supports the debate within the UN CCW with frequent publications of 'independent, interdisciplinary' research (Rosert & Sauer, 2019, pp. 370-371; iPRAW, *About*). Their arguments in support of hard governance be sub-categorised under legality, ethicality and security.

The legal opposition to LAWS pertains to their incompliance with IHL. Campaigners and scholars cite Additional Protocol I, Article 36 (New Weapons) of the Geneva Convention, which states:

> "In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party (Protocol Additional, 1977.)"

Accordingly, their collective position is that the UN CCW doesn't sufficiently control the development of LAWS in line with IHL; despite stating: 'human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines' (CCW, 2019, *Principal 1 & 2*). The CCW finds itself in a paradox because the development of intelligent autonomous technology should 'not be hampered' by such measures (CCW, 2019, *Principal 10*). LAWS adversaries demand that the UN adopts a more explicit and enforceable position, such as 'a pre-emptive ban on the weapons' development, production and use of LAWS, as oppose to implying its noncompliance with IHL through the language of meaningful human control (DOCHERTY, 2018).

Additionally, supporters of LAWS' hard governance level an ethical critique against autonomous weaponry on the premise that it is inhumane to allow a robot to discriminate over the taking of a human life (ASARO, 2012, p. 688). Essentially, critics argue that it is wrong to reduce human life to the cogs in a killing machine (Rosert & Sauer, 2019, p. 372.) It is argued that it violates basic human rights such as German Basic Law Article 1.1 that 'human dignity shall be inviolable' (Basic Law FRG, P. 15).

Further, these advocates of the hard governance of LAWS emphasise the threat they pose to international security due to the risk of proliferation to non-state actors and their capacity to radically transforming warfare as it is currently understood (Altmann & Sauer, 2017, p. 117, pp. 120-122; Gill, 2019, p. 172). Tellingly, President Putin stated that whoever is to become

the leader in the AI sphere 'will become the ruler of the world' (Vincent, 2017). The speed, inconspicuousness and destructiveness of autonomous weaponry are likely to be 'more limited by the laws of physics' than 'any deficiencies in the AI systems that control them' (Russel et al., 2015, p. 416). Due to the hugely advantageous military capabilities LAWS would offer, there is concern in the scholarship that this could lead to a state-level arms race as well as risk LAWS proliferation to non-state actors, unconcerned with the protocol of IHL – if unregulated (iPRAW, 2018, p.2; Warren & Hillas, 2020, p. 823; Anderson, 2016, pp. 33-34).

In consequence, there is broad support in the corporate world and in mainstream public opinion for hard governance of LAWS. For instance, prominent private sector figures in the tech industry have called LAWS as 'among the world's worst ideas,' with the likes of Elon Musk and Alphabet's Mustafa Suleyman signatories of an outright ban of 'killer robots' (CNBC, 2018; Gibbs, 2018). Moreover, surveys have indicated that popular opinion of the US and Belgium oppose the 'trend towards autonomous weapons' (Doherty, 2018).

However, there is a school of thought that questions the validity of hard governance of LAWS development, indicating that norms-based governance would not only be more appropriate, but more effective. Anderson and Waxman argue that it would be short-sighted to restrict the development of autonomous technologies that could greatly assist in civilian sectors, 'helping us achieve the 2030 sustainable development goals,' and when applied in a law-abiding, military context (HOROWITZ, 2018, p. 50; UN Discussion on the Emerging Technologies in the area of LAWS; Anderson, Waxman, 2013, p. 1).

Moreover, Maas questions the UN commitment to 'meaningful human control' in a practical sense – identifying potential 'challenge vectors' that could disrupt the efficacy of regulatory measures (Maas, 2019, p. 137). Maas identifies two forms of possible disruption. 'Direct' disruptions and 'indirect' disruptions that would render the imposition of 'established principles' 'more difficulty' (Ibid., p. 138). For example, direct disruption pertains to developments such as 'new innovation' that would cause 'military capabilities [to] qualitatively shift in destabilising ways' (Ibid., 138-140). 'Indirect' disruption refers to paradigm shifts in the international security landscape, such as the 'erosion of nuclear deterrence,' beholden to new state-of-the-art, Military Artificial Intelligence (MAI) being realised (Ibid., p. 129). These challenge vectors could limit the practicality of upholding the hard governance of LAWS.

Echoing this, Maas, Gill and Horowitz employ the Collingridge Dilemma to communicate how hard governance of LAWS may be hamstrung. As a developing technology that has yet to be fully realised, the regulation of LAWS faces a two-fold dilemma. Collingridge identifies how controlling a developing technology faces a double bind problem of not knowing how certain measures will impact a technology until it is suitably developed, at which point imposing regulatory measures are challenging because the technology is more entrenched (Collingridge, 1981).

Importantly, Horowitz recognises a 'direct disruption' – to use Maas' term – of regulating the technologies LAWS is composed of that would likely hamstring any hard governance of the development of LAWS. Horowitz considers artificial intelligence as an 'enabling technology,' more closely 'resembling the combustion engine or electricity than a specific weapon,' that will continue to develop commercially and militarily (Horowitz, 2018, p. 37). It is with this practical line of inquiry into the efficacy of regulating the constituent components of LAWS

that this paper means to continue. It follows that this question under review is:

‘Is the hard governance of the development of Lethal Autonomous Weapons Systems bankrupt?’

The foci of this paper shall be in relation to the role of sub-national actors and their centrality to the practical ability to regulate LAWS development. Maleficent appropriation of dual-use hardware and software is likely to present the hard governance of LAWS with its greatest challenge, yet the rhetoric of ban advocates surrounds the international arena and how international law can be implemented to thwart the development of LAWS (FLI, 2022, SKR, 2021).

Through a qualitative methodology I shall, firstly, discuss how dual-use technologies shall invite complications for governing the development of LAWS. I separate these technologies into hardware and software, both of which are manifold applications in the civilian sector, complicating possibilities of export controls. I expand by explaining how AI is likely to evolve vis-à-vis commercial incentives in civilian and military sectors, which will only increase the likelihood that autonomous weaponry capabilities are realised. In doing so, I underline how there is a dialectical relationship between software programmers and state-level and governmental actors over the ability to securitise LAWS. Here, I shall introduce the case studies of Project Maven, Dragonfly and additional corporate-government contracts to evidence the dilemma this will pose to the hard governance of LAWS. Secondly, I shall discuss how digital platforms – the loci of the development of Machine Learning Algorithms (MLAs) behind autonomous weapon technology – are markedly nebulous. I shall illustrate how open-source software and digital anonymity renders the web in almost extralegal realm, where attempts at regulating it have proved largely futile. On this basis, I infer that the hard governance of developing LAWS will be undermined significantly.

Theoretical framework: Securitisation Theory

This paper shall emphasise the practical limitations to the hard governance of LAWS, focussing on the nature of the technology that they are beholden to. Here, the role of commercially driven, dual-use AI and anarchic digital platforms would be seen to undermine state-level actors and international initiatives. To expand, the role of corporations, software communities and the individual programmer become securitising actors.

It follows that this paper shall situate its theoretical framework in Securitisation Theory. Securitisation Theory, a post-Cold War approach to explaining international security, ‘widened the scope of security to include other referent objects beyond the state,’ emphasising the role of language in the quantification of a threat (Eroukhmanoff, p. 105; Buzan, Wæver & de Wilde 1998, p. 26). Despite the rhetoric of LAWS adversaries, employing evocative, sci-fi diction to buttress their arguments, software communities continue to exercise their political agnosticism and counter-establishment culture (Pizza et al. p. 145; Warren & Hillas, 2020, 837). This denial of a self-determining state agenda, despite how ‘slaughterbots’ are narrativized by the KRC and ICRC, communicates the dialectical power relation between individual action and state-actors. In this sense, the role of ‘the audience’ in validating a ‘securitising move’ to the extent all actors agree to pursue every means necessary to thwart it, as oppose to a security threat simply existing ‘out there,’ reflects a central tenet of Securitisation Theory (Eroukhmanoff, p. 104, 106).

Furthermore, Securitisation Theory means to 'sectoralise' security, identifying economic, societal, military, political and environmental sectors (Ibid., p. 104). Consequently, 'by talking about referent objects we can ask: Security for whom? Security from what? And security by whom?' (Ibid., p. 104). This is important because it reinforces the idea that security is not simply a given and agreed upon by all actors at the state and non-state level. Again, the role of software communities and coders in influencing the trajectory of military applications of AI – as shall be expanded on below – is a technocratic, as oppose to democratic, process. Due to the anarchic nature of world wide web, hackers, spoofers, leakers and open-source software programmers are in a position to judge whether or not LAWS constitute a sufficient security threat as to warrant them abandon their principles of anti-establishment, total liberalism and political agnosticism, irrespective of the verbosity of LAWS adversaries on the international stage.

In addition, the practical scope of analysis of the paper, stepping away from ethical imperatives to regulate autonomous weaponry, resonates with the Securitisation Theory. Eroukhmanoff regards how Securitisation Theory is 'more concerned with 'how' rather than 'why' question,' viewing securitisation as a process operating at multiple levels of analysis.

Methodology

This paper shall be employing a qualitative methodology for two reasons. Firstly, when considering the theoretical framework of this paper – the idea of security as an articulation and securitisation as a relational concept; a conceptual plane where actors at multiple levels of analysis compete over the quantification of a threat – and the speculative inquiry into the prospective ability to control the development of LAWS, I am making inference 'based on bits and pieces of […] observations that address different aspects of a problem,' such is qualitative analysis (Gerring, 2017, p. 18). Secondly, I am employing case studies – such as Project Maven, Dragonfly, etc. – as examples of how the hard governance of developing autonomous weaponry could be undermined. This form of inferential analysis does not pertain to quantitative data.

Discussion

The development of a LAWS would require the necessary hardware, software, data sets – to train the software – and sufficient technical know-how to combine the three into a functioning autonomous weapon. The hardware relates to 'physical platforms and delivery systems, like aircraft vehicles,' but also 'sensors, actuators and processors' (iPRAW, 2020, p. 3). Software is 'a component of a weapon system that orchestrates the functions of a physical system,' such as programs and algorithms (Ibid., p. 3). Under the umbrella of software, 'Artificial Intelligence' exists as a 'catch-all term used in a technical sense to refer to a set of computational techniques' (iPRAW, 2017, p. 9). In its technical sense, AI 'refers to a highly diverse set of computational techniques' such as machine learning or Deep Learning, all of which are advanced applications of 'mathematical logic, advanced statistics, and other computational techniques' (Ibid., p. 9). Lastly, data is the 'basis for information about the world that system engineers can use to train data-driven systems with computational methods' (e.g., AI) (iPRAW, 2020, p. 3). Combined together correctly vis-à-vis apt technical expertise, hardware and AI 'are mainly applied to find patterns, classify and categorise inputs, and produce sufficiently optimised courses of action in computationally efficient ways, according to specified goals (Ibid., pp. 9-10).

Now, this could be applied to the mundane, such as the case of AlphaZero, an algorithm developed by Google DeepMind that used self-play reinforcement learning to train 'itself' to play chess and defeat the world champion in four hours (Warren & Hillas, 2020, p. 834). It could also be employed to foster an unmanned urban 'search and destroy mission' or carry out ethnic genocide (Russell, 2015, p. 416).

This 'complex and polymorphic nature' to autonomous technology 'represents a special challenge' to the efficacy of regulating LAWS through hard governance (Sauer & Rosert, 2020, p. 16). The regulation of LAWS is less straightforward than more conventional arms-control agreements due to the dual-use nature of its constituent components. Processors and sensors are commonplace in contemporary society, existing in all 'smart' objects; our phones, televisions, laptops and cars (Bohn et al., 2010, p. 763). This presents a dual-use issue whereby hardware could be appropriated 'into a weapon with autonomous targeting functions,' despite its intended purpose (Sauer & Rosert, 2020, p. 3). In consequence, export control of LAWS hardware is significantly more challenging than, for example, monitoring the purchase of 'asphyxiating' or 'poisonous' gasses, or the sale of powerful lasers to be used for Blinding Laser Weapons, both of which are banned under IHL and don't have everyday civilian functions (Moon & Lin, 2021, 25; ICRC, 2021, p.3; CCW, 1995, *Additional Protocol IV*).

This dual-use dilemma – that could hamstring the effectiveness of the hard governance of LAWS – extends to software and data. Machine learning algorithms (MLAs) that could be used to train autonomous weapons functions are largely open source nowadays (iPRAW, pp. 4-5). Open-source software can be legally copied and modified to encourage collaborative 'development and community through education and advocacy' (OpenSourceInitative, 2022). Military software has been 'mostly proprietary in the past,' now, regulation is at the behest of the person manipulating open-source code with malintent (iPRAW, 2020, p. 5). This has the capacity to further undermine safety mechanisms that could be implemented into code to restrict their use or simply the decision to input any restrictions on the code at all (Moon & Lin, 2021, p. 25). Scholars acknowledge that the 'dual-use of robotic devices […] by non-state actors would pose a threat' to 'national security' (Hillas, 2020, 823).

Echoing Securitisation Theory, we see how 'securitising actors are not limited to politicians' (Eroukmanhoff, p. 107). Despite the CCW's guiding principles on LAWS (development and deployment), the ability to effectively securitise LAWS as 'extreme security issues to be dealt with urgently, depends on the behaviour of the 'audience' – in this case, software developers (Eroukhmanoff, 2018, pp. 105-106). This issue of dual-use software is exacerbated by the nebulous nature of digital platforms. Multiple characteristics of the web render its regulation particularly futile. Firstly, the 'construction and installation' of LAWS is subject to the ethical choices of 'a technocracy not democracy' (O'Hara, 2017, 99). Many of these programming communities are 'on the pay of a few oligopolistic corporations directly accountable to no external party' in comparison to the exposure of an elected lawmaker (Ibid., 100; Noto & Diega, 2018, 16). Therefore, to an extent – within the confines of the law, the decisions to implement safety mechanisms in MLNs is at the prerogative of the programmer. Consequently, the ability to control the export of code that is complicit with IHL on LAWS regulation is problematic.

In addition, with the characteristic 'openness' of the internet, digitisation has undermined law enforcement on digital platforms (Finger, 50). Digital anonymity is a factor that could limit deterrence of illegal dealings with autonomous weaponry, including the trade of hardware,

software, training expertise and data. This would pose a further issue for enforcing the hard governance of LAWS. For instance, cryptographic techniques 'behind trustless property systems and transfers,' such as Bitcoin, create a state of 'crypto anarchy' that 'promises liberation from state and institutional oversight' (Baker, 2015, 371-373). Baker uses the analogy of fruitless efforts to thwart 'the Napster generation from pirating music,' whereby web surfers are not deterred by the risk of illegal file sharing (Ibid., 371). When applied to the context of controlling autonomous weapon development, the 'infamous black market trading platform called the "Silk Road,"' shows the more dangerous application of digital anonymity (Ibid., 373). This deep-net site hosted an 'utopic libertarian drug market,' allowing 'users to trade only in Bitcoin to promote anonymity and avoid law enforcement' (Ibid., 375; Templeton, 2014). In 2014, despite big busts by the FBI and National Crime Association – leading to 17 arrests, deep-net websites continue to resurface – Silk Road 2.0 replacing its predecessor – in a 'perpetual whack-a-mole scenario' (Ibid., 2014). This creates an almost extralegal realm of security, whereby securitisation – in this case, the securitisation of LAWS through the regulation of digital platforms – is 'characterised by competition' between political securitisers and the receivers of their securitising moves (Eroukhmanhoff, p. 107).

The lack of deterrence complicates the issues of successfully securitising LAWS from a governmental position. This issue is furthered by software communities' commitment to the ultra-liberalism of open-source and 'crypto culture' itself, a form of 'rebellious political expression […] developed in the vacuum of created by repeated breaches of trust by the traditional institutions that surround us' (Baker, 2015, 371). Software communities' commitment to 'political agnosticism' has been demonstrated by the limited success of open-source activism and of ethics-based licences – small-scale initiatives of a 'a contributor modifying their project's licence to restrict who can use their code' (Moon & Lin, 2021, 25-26). Here, we see the subversion of state-level strategy at the behest of software culture. Again, Securitisation Theory provides insight into how speech is important to the quantification of a threat. If state-level actors want to convince these communities to comply with their agendas, it must be 'articulated' in a sufficiently persuasive way to abandon their political ideologies (Eroukhmanhoff, 2018, p. 104).

Importantly, the manipulative dual-use of technologies applicable to LAWS is not confined to non-state actors. The ability of software developers to thwart the use of their code for military purposes has led to protests from corporate employees against their employer's collusion with the US government. Over 3,000 Google employees signed a petition in protest against company's involvement with 'a U.S. Department of Defense artificial intelligence (AI) project that studies imagery and could eventually be used to improve drone strikes in the battlefield'; also known as Project Maven (Shane & Wayakabashi, 2018). Despite this prompting Google's CEO Sundar Pichai to state publicly that Google will not pursue AI applications in 'weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people,' Google built its software on 'TensorFlow, an open-source software library' and so the DoD can still apply the code to serve their purposes (Pichai, 2018; Anderson, 2018). Here, we see how government and military officials possess the parallel capacity to resist 'proprietary constraints on the code,' due to the fact that the Pentagon now has access to a trainable algorithm 'it can continue to develop and refine its object-recognition as it chooses' (Anderson, 2018).

In addition, Google employees have also criticised the company's organisational structures from keeping rank-and-file employees in the dark over what they are in fact developing,

therefore, limiting the ability of software developers to ensure that their programmes are compliant with IHL in relation to LAWS (Warren & Hillas, 2020). In the case of Dragonfly, a censored Chinese search engine developed by Google, a signatory spoke at a Group of Government Affairs (GGE) side event at United Nations Office for Disarmament Affairs (UNODA), stating how programmers don't know the 'true nature of the project they are working on' due to the fact that 'technology gets [opaquer] through layers of automation, the cloud, and the artificial intelligence and machine learning technologies' (Ibid., 836).

Similarly, Palantir Technologies have faced criticism for providing the software behind US Immigration and Customs Enforcement's (ICE) 'Investigative Case Management systems, which the agency has used to plan and carry out immigration arrests' (Del Valle, 2019). Again, despite deleting the Ruby code – Chef Sugar, earlier iterations of the open-source software exist, permitting their continued appropriation for military use (Ibid., 2019). Less covert uses of code for military purposes have also sparked protest in the US. Palantir Technologies contract with the ICE – 100,000 signatories called on the company to 'pull out of its contract with ICE' (Cox, 2019). Programmers and researchers also face the hurdle of successfully exposing these truths – with it taking 'years to build a social media following that makes [outreach] worthwhile' (Russell et al., 2015, 415). Together, these issues of dual-use technologies reveal a special problem with the hard governance of LAWS. The ability to enforce traditional means of export control and copyright law enforcement on these 'enabling' components are largely subverted (iPRAW p. 1, Finger, p. 50). Again, this is a two-sided problem, with government and non-state actors undermining the ability to securitise LAWS in line with IHL.

The aforementioned cases reveal how there is in fact a dialectical relation between programmers and the state-level securitising actors. Whilst we have seen how non-state adversaries could manipulate dual-use civilian hardware and software to undermine international agendas, such as ensuring 'meaningful human control' over autonomous technology, governments are able to exploit open-source software for purposes incongruent with their programmers' intentions. This competitive, often adversarial, security landscape communicates how multiple securitising actors – e.g. software developers and defence companies – compete over 'the 'right' knowledge over the threat' and 'competition over the 'right' solution' (Eroukhmanhoff, p. 107).

The prospective development of LAWS is unlikely to decrease either. Commercially-driven development in AI is likely improve MLAs and their availability – software that is necessary to the functioning of LAWS. The McKinsey Institute predicts that automation will disrupt 15% of the workforce by 2030, underlining the growing capital market of AI (Horowitz, 2018, 50). Moreover, this development is state sponsored across the globe. President Xi has pledged to build a $150b AI industry by 2030 (Wired). Likewise, western governments are closely tied to Defence contractors. Returning to Palantir, the analytics firm secured a 'role in $823m government contract in 2020' (Novak, 2013). Due to widespread lobbying, the US have a vested interest in defence expansion. Defence companies contributed $27m to candidates and political action committees during the 2012 campaign cycle, whilst, in 2015, 900 representatives of 266 companies lobby Congress (Rahall, 2015). Indeed, military purposes must be compliant with the guiding principles of LAWS and IHL. However, in combination, the commercial and military first-mover advantages of AI developments will likely foster a more destabilising environment for the regulation of LAWS as the AI-military nexus deepens and more advanced AI comes into fruition.

Limitations

The nature of this paper has been largely speculative. Principally, it is inferring how regulatory protocols of a weapons technology – that is yet to exist – are challenged at present and how they could be in the future. Moreover, it employs recent case studies to consider how digital platforms could invite problems for controlling the export of code that could apply to LAWS. Further, it speculates on the trajectory of AI development to buttress these inferences. Due to the speculative nature of this discussion, it should be noted I speak with no emphatic certainty, nor proof of the forthcoming conclusion. However, these admissions do not undermine the validity of the discussion for, at present, all existing literature surrounding LAWS development is restricted by an equal degree of speculation.

A secondary limitation of this paper is a product of my limited technical expertise in the sphere of Artificial Intelligence that LAWS programming is beholden to. With best intentions, this paper has sought to employ relevant scientific discourses on AI and autonomous technology to support its inferences about LAWS and how its regulation could be undermined. It should be noted that due to a limited technical knowledge, this paper has not addressed the know-how required to construct a LAWS from its constituent parts – a factor that could alter the capabilities of non-state actors to develop and deploy LAWS.

Conclusion

The ability to regulate LAWS in compliance with protocols of hard governance, such as UN CCW Guiding Principles, is less straightforward than previous prohibitions of weapons under the CCW. According to IHL, the maintenance of meaningful human control must be considered in the development and deployment of autonomous weaponry. However, the ability to enforce this objective is likely to be bankrupt. Due to the ever-growing domain of Artificial Intelligence, compounded with exposure to dual-use and digital anarchy, guaranteeing that weaponry capable of selecting and engaging targets autonomously – without human input – will never materialise, is unlikely.

It is a valid objection to inquire as to whether IHL enforcement of LAWS is any different from the law enforcement of any other crime. Indeed, the capacity to regulate LAWS is comparable to all forms of IHL enforcement. The law is only effective as long as those who live under it, abide by it. However, what this paper has sought to communicate is that the process of regulating LAWS through hard governance will be markedly challenging. In digital spaces, deterrence is largely ineffectual, and the growth of the US military-industrial complex and global AI sectors certainly do not appear to be conducive to a military environment without autonomous weaponry.

Significantly, what also separates the regulation of LAWS from other IHL crimes is its imperative to do so. LAWS threaten to radically alter international security, pervading all ‘five sectors of security: economic, societal, military, political and environmental’ (Eroukhmanhoff, 2018, p. 105). Whilst hard governance of LAWS may prove ineffectual, other soft forms of governance will be required.

Bibliography

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016, October). Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security* [online]. Available at: https://doi.org/10.1145/2976749.2978318 (Accessed: 25 April 2022), 308-318.

Altmann, J., & Sauer, F. (2017) 'Autonomous weapon systems and strategic stability,' *Survival*, *59*(5) [online]. Available at: https://doi.org/10.1080/00396338.2017.1375263 (Accessed: 25 April 2022). 117-142.

ANDERSON, K. (2016) 'Why the Hurry to Regulate Autonomous Weapon Systems - But Not Cyber-Weapons,' *Temple International & Comparative Law Journal,* 30(spring) [online]. Available at: https://heinonline.org/HOL/Page?handle=hein.journals/tclj30&div=6&g_sent=1&casa_token=&collection=journals (Accessed: 25 April 2022), 16-41.

ANDERSON, K., WAXMAN, M.C. (2013) 'Debating Autonomous Weapon Systems, Their Ethics, and Their Regulation Under International Law,' *The Oxford Handbook of Law, Regulation, and Technology,* 14(553) [online]. Available at: https://ssrn.com/abstract=2978359 (Accessed: 25 April 2022), 1097-1117.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). 'Concrete problems in AI safety,' *arXiv preprint arXiv:1606.06565* [online]. Available at: https://arxiv.org/pdf/1606.06565.pdf (Accessed: 25 April 2022), 1-29.

ASARO, P. (2012) 'Banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making,' *International Review of the Red Cross,* 94(886) [online]. Available at: doi:10.1017/S1816383112000768 (Accessed: 25 April 2022), 687-709.

ATHERTON, K. (2018) 'Targetting the future of the DoD's controversial Project Maven initiative,' *C4ISRNET,* 27 July [online]. Available at: https://www.c4isrnet.com/it-networks/2018/07/27/targeting-the-future-of-the-dods-controversial-project-maven-initiative/ (Accessed: 25 April 2022).

BAKER, E. (2015) 'Trustless property systems and anarchy: how trustless transfer technology will shape the future of property exchange,' *Southwestern Law Review*, 45(2) [online]. Available at: https://heinonline.org/HOL/Page?handle=hein.journals/swulr45&div=16&g_sent=1&casa_token=&collection=journals (Accessed: 25 April 2022), 350-376.

Bartáková, V. (2019). Lethal Autonomous Weapon Systems [online]. Available at: https://www.uniuni.edu/wp-content/uploads/2021/06/sird-thesis-veronika-bartakova.pdf (Accessed: 25 April 2022), 1-51.

Bills, G. (2014). LAWS unto themselves: Controlling the development and use of lethal autonomous weapons systems. *Geo. Wash. L. Rev.*, *83* [online]. Available at: https://

heinonline.org/HOL/Page?handle=hein.journals/gwlr83&div=7&g_sent=1&casa_token= (Accessed: 25 April 2022), 176.

Bohn, J. et al. (2010) 'Living in a World of Smart Everyday Objects—Social, Economic, and Ethical Implications,' *Human and Ecological Risk Assessment: An International* Journal, 10(5) [online]. Available at: https://doi.org/10.1080/10807030490513793 (Accessed: 25 April 2022), p. 763-785.

Booth, B. (2018) '"Autonomous weapons are among the world's dumbest ideas": A.I. CEO,' *CNBC*. Available at: https://www.cnbc.com/2018/03/15/autonomous-weapons-are-among-the-worlds-dumbest-ideas-a-i-ceo.html (Accessed: 25 April 2022).

Buzan, B., Wæver, O., De Wilde, J. (1998) *Security: A New Framework for Analysis,* Lynne Rienner, London.

Collingridge, D. (1981) *The Social Control of Technology*, Palgrave Macmillan, London.

COX, J. (2019) '"Everyone Should Have a Moral Code," Says Developer Who Deleted Code Sold to ICE,' *VICE*
, 20 September [online]. Available at: https://www.vice.com/en/article/mbm3xn/chef-sugar-author-deletes-code-sold-to-ice-immigration-customs-enforcement (Accessed: 25 April 2022).

DeSario, J., & Langton, S. (1984). Citizen participation and technocracy. *Review of Policy Research*, *3*(2), Available at: https://doi.org/10.1111/j.1541-1338.1984.tb00116.x (Accessed: 25 April 2022), 223-233.

Eroukhmanoff, C. (2018) Securitisation Theory: An Introduction, *E-International Relations*. Available at: https://www.e-ir.info/2018/01/14/securitisation-theory-an-introduction/ (Accessed: 25 April 2022).

Fierke, K.M. (2015) *Critical Approaches to International Security*, 2nd ed. Polity Press, Cambridge.

GALLAGHER, R. (2019) 'Google Is Conducting A Secret "Performance Review" Of Its Censored China Search Project,' 27 March [online]. Available at: https://theintercept.com/2019/03/27/google-dragonfly-china-review/ (Accessed: 25 April 2022).

Garcia, D. (2015). 'Killer robots: Why the US should lead the ban.' *Global Policy*, *6*(1) [online]. Available at: https://doi.org/10.1111/1758-5899.12186 (Accessed: 25 April 2022), 57-63.

Geist, E. M. (2016) 'It's already too late to stop the AI arms race—We must manage it instead,' *Bulletin of the Atomic Scientists*, *72*(5) [online]. Available at: https://doi.org/10.1080/00963402.2016.1216672 (Accessed: 25 April 2022), 318-321.

Gerring, J. (2017) 'Qualitative Methods,' *The Annual Review of Political Science*, 20(15) [online]. Available at: https://www.annualreviews.org/doi/pdf/10.1146/annurev-polisci-092415-024158 (Accessed: 25 April 2022).

Gibbs, S. (2017) ‘Elon Musk leads 116 experts for outright ban on killer robots,’ *The Guardian*. Available at: https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war (Accessed: 25 April 2022).

Gill, A. S. (2019). Artificial intelligence and international security: the long view. *Ethics & International Affairs*, *33*(2) [online]. Available at: doi:10.1017/S0892679419000145 (Accessed: 25 April 2022), 169-179.

Grace, K. (2015). The Asilomar Conference: A Case Study in Risk Mitigation. *Technical report 2015–9*. [online]. Available at: https://intelligence.org/files/TheAsilomarConference.pdf (Accessed: 25 April 2022), 1-68.

HAMBLING, D. (2021) ‘Australian Army Getting Bulletproof Swarming Attack Robots,\ *Forbes*, 4 November [online]. Available at: https://www.forbes.com/sites/davidhambling/2021/11/04/australian-army-gets-bulletproof-attack-robots/?sh=4857e00e463b (Accessed: 25 April 2022).

HAMBLING, D. (2020) ‘U.S. To Equip MQ-9 Reaper Drones With Artificial Intelligesnce,’ *Forbes,* 11 December [online]. Available at: https://www.forbes.com/sites/davidhambling/2020/12/11/new-project-will-give-us-mq-9-reaper-drones-artificial-intelligence/?sh=3ae3150a7a8e (Accessed: 25 April 2022).

Horowitz, M. C. (2018) ‘Artificial intelligence, international competition, and the balance of power,’ in Krieger et al. *American Defense Policy. 2018*, *JHU Press*.

Horvitz, E., & Selman, B. (2009). AAAI Presidential Panel on Long-Term AI Futures: Interim Report from the Panel Chairs. *Association for the Advancement of Artificia l Intelligence (AAAI)*.
[online]. Available at: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/11/panel_chairs_ovw.pdf (Accessed: 25 April 2022), 1-5.

Klincewicz, M. (2015) Autonomous Weapons Systems, the Frame Problem and Computer Security, Journal of Military Ethics, 14:2 [online]. Available at: DOI:10.1080/15027570.2015.1069013 (Accessed: 25 April 2022), 162-176.

Lawson, B., Samson, D., & Roden, S. (2012). Appropriating the value from innovation: inimitability and the effectiveness of isolating mechanisms. *R&D Management*, *42*(5) [online]. Available at: https://doi.org/10.1111/j.1467-9310.2012.00692.x (Accessed: 25 April 2022), 420-434.

Lin, C., & Moon, A. (2021). Can Open Source Licenses Help Regulate Lethal Autonomous Weapons?. *IEEE Technology and Society Magazine*, *40*(3) [online]. Available at: DOI:10.1109/MTS.2021.3101832 (Accessed: 25 April 2022). 25-27.

Mattingly-Jordan, S. (2017). The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [online]. Available at: https://standards-qa21.ieee.org/wp-content/uploads/import/documents/other/becoming_leader_global_ethics.pdf (Accessed: 25 April 2022).

Maas, M. M. (2019). Innovation-Proof Global Governance for Military Artificial Intelligence?: How I Learned to Stop Worrying, and Love the Bot. *Journal of International Humanitarian Legal Studies*, *10*(1) [online]. Available at: https://doi.org/10.1163/18781527-01001006 (Accessed: 25 April 2022). 129-157.

Maggetti, M. (2015) 'Hard and soft governance,' *Research Methods in European Union Studies*, Palgrave Macmillan, London [online]. Available at: https://link.springer.com/chapter/10.1057/9781137316967_16 (Accessed: 25 April 2022).

Manyika, J. et al. (2017) 'Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages,' *McKinsey & Company* [online]. Available at: https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages (Accessed: 25 April 2022)

Meier, M. W. (2016). Lethal autonomous weapons systems (laws): conducting a comprehensive weapons review. *Temp. Int'l & Comp. LJ*, *30* [online]. Available at: https://heinonline.org/HOL/Page?handle=hein.journals/tclj30&div=13&g_sent=1&casa_token=&collection=journals# (Accessed: 25 April 2022).

Montero, J., & Finger, M. (2021). *The rise of the new network industries: Regulating digital platforms* Routledge.

Mordvintsev, A., Olah, C., & Tyka, M. (2015). 'Inceptionism: Going deeper into neural networks,' *Google Research Blog* [online]. Available at: https://research.google/pubs/pub45507/ (Accessed: 25 April 2022).

Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE conference on computer vision and pattern recognition* [online]. Available at: https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf (Accessed: 25 April 2022), 427-436.

Noto La Diega, G. (2018). Against the dehumanisation of decision-making–Algorithmic decisions at the crossroads of intellectual property, data protection, and freedom of information. *Against the Dehumanisation of Decision-Making–Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information (May 31, 2018)*
[online]. Available at: https://ssrn.com/abstract=3188080 (Accessed: 25 April 2022).

Novak, V. (2013) 'Defense: Background,' *Open Secrets* [online]. Available at: https://www.opensecrets.org/industries/background.php?cycle=2014&ind=D (Accessed: 25 April 2022)

O'hara, K. (2017). Smart contracts-dumb idea. *IEEE Internet Computing*, *21*(2) [online]. Available at: DOI:10.1109/MIC.2017.48 (Accessed: 25 April 2022), 97-101.

Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2017, April). Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security* [online]. Available at: https://

doi.org/10.1145/3052973.3053009 (Accessed: 25 April 2022).

Pichai, S. (2018) 'AI at Google: our principles,' *The Keyword*. Accessed at: https://www.blog.google/technology/ai/ai-principles/ (Accessed: 25 April 2022).

Pizza, M., Romanoff, M., Engelhardt, T. (2021) 'AI for humanitarian action: Human rights and ethics,' *International Review of the Red Cross,* 102(913), 145-180, Cambridge University Press, Cambridge.

Radin, S., & Coats, J. (2016). Autonomous Weapons Systems and the Threshold of Non-International Armed Conflict. *Temp. Int'l & Comp. LJ*, *30* [online]. Available at: https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/tclj30&section=14 (Accessed: 25 April 2022), 33.

Rahall, K. (2014). The green to blue pipeline: Defense contractors and the police industrial complex. *Cardozo L. Rev.*, *36* [online]. Available at: https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/cdozo36&section=49 (Accessed: 25 April 2022), 1785.

Rajkumar, R., Lee, I., Sha, L., & Stankovic, J. (2010, June). Cyber-physical systems: the next computing revolution. In *Design automation conference* [online]. Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5523280&tag=1 (Accessed: 25 April 2022), 731-736.

Russell, S., Hauert, S., Altman, R., & Veloso, M. (2015). Ethics of artificial intelligence. *Nature*, *521*(7553) [online]. Available at: https://doi.org/10.1038/521415a (Accessed: 25 April 2022), 415-416.

Rosert, E., & Sauer, F. (2021). How (not) to stop the killer robots: A comparative analysis of humanitarian disarmament campaign strategies. *Contemporary Security Policy*, *42*(1) [online]. Available at: https://doi.org/10.1080/13523260.2020.1771508 (Accessed: 25 April 2022), 4-29.

Rosert, E., & Sauer, F. (2019). Prohibiting autonomous weapons: put human dignity first. *Global Policy*, *10*(3) [online]. Available at: https://doi.org/10.1111/1758-5899.12691 (Accessed: 25 April 2022), 370-375.

RUSSELL, S., AGUIRRE, A., JAVORSKY, E., TEGMARK, M. (2021) 'Lethal Autonomous Weapons Exist; They Must Be Banned – It may not be too late to put the evil "Slaughterbots" genie back in the bottle, if the world acts now,' *IEEE Spectrum,* 16 June [online]. Available at: https://spectrum.ieee.org/lethal-autonomous-weapons-exist-they-must-be-banned (Accessed: 25 April 2022).

Schmitt, M.N. (2012) 'Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics,' *Harvard National Security Journal Feature*. Available at: https://ssrn.com/abstract=2184826 (Accessed: 25 April 2022).

Shane, S. & Wakabayashi, D. (2018) ''The Business of War': Google Employees Protest Work for the Pentagon,' *The New York Times* [online]. Available at: https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html (Accessed: 25 April 2022).

Sparrow, R. (2016) 'Robots as "evil means"? A rejoinder to Jenkins and Purves,' *Ethics & International Affairs*, *30*(3), 401-403.
Basic Law for the Federal Republic of Germany (Basic Law 1949, last amended 2020, Article 1.1).

THOMPSON, N. (2018) 'The AI Cold War That Threatens Us All,' *Wired*, 23 October [online]. Available at: https://www.wired.com/story/ai-cold-war-china-could-doom-us-all/ (Accessed: 25 April 2022).

Warren, A., & Hillas, A. (2020). Friend or frenemy? The role of trust in human-machine teaming and lethal autonomous weapons systems. *Small Wars & Insurgencies*, *31*(4) [online]. Available at: https://doi.org/10.1080/09592318.2020.1743485 (Accessed: 25 April 2022), 822-850.

Vincent, J. (2017) "Putin says the nation that leads in AI 'will be the ruler of the world",' *The Verge*. Available at: https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world (Accessed: 25 April 2022).

'An Open Letter: research priorities for robust and beneficial artificial intelligence,' *Future of Life Institute* [online]. Available at: https://futureoflife.org/2015/10/27/ai-open-letter/ (Accessed: 25 April 2022).

'LAWS and Export Control Regimes: Fit for Purpose? iPRAW Working Paper – April 2020,' *International Panel on the Regulation of Autonomous Weapons* [online]. Available at: https://www.ipraw.org/wp-content/uploads/2020/04/iPRAW_WP_ExportControls.pdf (Accessed: 25 April 2022), 1-5.

'AI Existential Safety Community,' *Future of Life Institute* [online]. Available at: https://futureoflife.org/team/ai-existential-safety-community/ (Accessed: 25 April 2022).

'Top Myths and Facts on Human Cotnrol of Autonomous Weapons,' *Future of Life Institute* [online]. Available at: https://futureoflife.org/2020/10/10/top-myths-and-facts-on-human-control-of-autonomous-weapons/ (Accessed: 25 April 2022).

'FLI's Position on Lethal Autonomous Weapons Systems,' *Future of Life Institute* [online]. Available at https://futureoflife.org/2020/06/05/flis-position-on-lethal-autonomous-weapons/: (Accessed: 25 April 2022).

(2021) 'A Shared Movement,' *Stop Killer Robots*. Available at: https://www.stopkillerrobots.org/a-global-push/a-shared-movement/ (Accessed: 25 April 2022).

(2021) 'ICRC position on autonomous weapons,' *International Committee of the Red Cross*. Available at: https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems (Accessed: 25 April 2022).

South Korea, "Opening Statement," UN Office at Geneva, Discussions on Emerging Technologies in the Area of LAWS, 2015.

(2021) '1993 Chemical Weapons Convention,' *International Committee of the Red Cross*

[online]. Available at: https://www.icrc.org/en/document/1993-chemical-weapons-convention (Accessed: 25 April 2022).

Del Valle, G. (2019) 'College Kids Are Vowing Not to Work for Palantir Because of Its ICE Contracts,' *Vice News* [online]. Available at: https://www.vice.com/en/article/ywa3kw/college-kids-are-vowing-not-to-work-for-palantir-because-of-its-ice-contracts (Accessed: 25 April 2022).

ESR, (2018) 'Non-discrimination is a core value of open source,' *Armed and Dangerous: Sex, software, politics, and firearms. Life's simple pleasures…* [online]. Available at: http://esr.ibiblio.org/?p=8106 (Accessed: 25 April 2022).

(2017) 'Focus on Computational Methods in the Context of LAWS,' *International Panel on the Regulation of Autonomous Weapons,* 2 [online]. Available at: https://www.ipraw.org/wp-content/uploads/2017/11/2017-11-10_iPRAW_Focus-On-Report-2.pdf (Accessed: 25 April 2022), 1-22.

(2022) *OpenSourceInitiative* [online]. Availble at: https://twitter.com/OpenSourceOrg (Accessed: 25 April 2022).