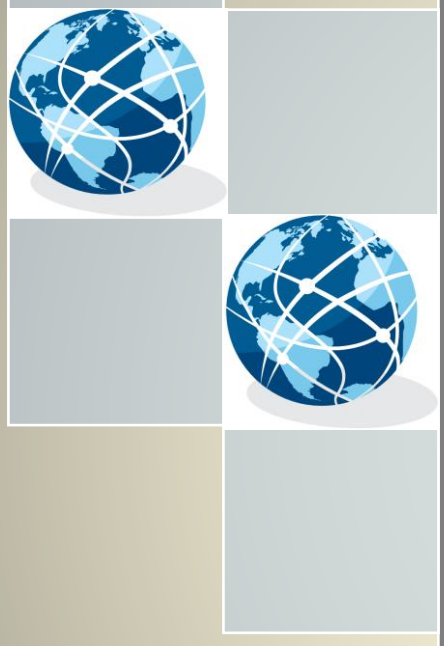
STRATEGIC REVIEW
Volume 1, Issue 3, May 2019



KEY STRATEGIC QUESTIONS RELATED TO CYBER-ENABLED AUTONOMOUS WEAPONS SYSTEMS

Caitríona Heinl

Research Network ‚European Security and Strategy‘

The Research Network on 'European Security and Strategy' examines strategic aspects of European security, defence and foreign policy. It brings together more than 45 researchers from 25 universities in Europe.

Network coordinators:

- Cornelia-Adriana Baciu, PhD Candidate in final year at the School of Law and Government, Dublin City University, Ireland.
- Delphine Deschaux-Dutard, Assistant Professor at the Centre d'Etudes sur la Sécurité Internationale et le Coopérations Européennes (CESICE), University of Grenoble, France.

Steering Committee:

- Antonio Calcara, Adjunct Professor at Vesalius College, Brussels & PhD Candidate in final year at LUISS Guido Carli, Rome and Vrije University of Brussels, Belgium.

You can contact us at ess.researchnetwork@gmail.com.

Website: https://ess-research.org/.

Twitter: @STRATEGIE21.

Caitríona Heinl is an expert in International Cyber Policy and Executive Strategist & Lead Strategist for Asia Pacific, EXEDEC. In 2019, her book chapter "Maturing autonomous cyber weapons systems: Implications for international security cyber and autonomous weapons systems regimes" will appear in the Oxford Handbook of Cyber Security, Oxford University Press, Ed. Prof. Paul Cornish (forthcoming).

Research Network
European Security and Strategy

Strategic Review
Vol 1(3), May 2019

Key Strategic Questions Related to Cyber-Enabled Autonomous Weapons Systems

Caitríona Heinl
Expert in International Cyber Policy
Executive Strategist & Lead Strategist for Asia Pacific, EXEDEC
Email: caitriona.heinl@exedec.net.

Introduction

While existing technological capabilities continue to conceptually challenge security strategies, emerging technologies and potential exponential leaps related to AI, machine learning, big data and computing power are maturing in a complex manner, calling human comprehension into question. The focus at UN level under the Convention on Conventional Weapons and its lethal autonomous weapons systems (LAWS) Group of Governmental Experts (GGE) has mainly surrounded physical, albeit AI-enabled, autonomous weapons systems (rather than cyber systems). Nonetheless, the former Chair of this group is now co-Lead of the UN High Level Panel on Digital Cooperation which has a mandate to consider cooperation across subject domains to safeguard against unintended consequences. Several strategic questions which stakeholders should still consider include, are: (1) How can commercial and civilian benefits vis-à-vis potential military use be guaranteed?; (2) What will these technological developments mean for global affairs and balance of power questions?; (3) What are the real opportunities and risks associated with these technologies?; and (4) In terms of machine learning, what do unpredictable outputs which can be produced by unpredictable inputs mean for strategic certainty?

This paper identifies several areas where future confidence building and collaborative initiatives could support international and regional stability. First, there is an identified need to consider the ethical and legal implications of developments in AI and robotics in the military context. Although major powers are unlikely to refrain from developing such military capabilities (which they do not deem to be contrary to international law), civil society and non-governmental groups continue to display strong convening powers. Convening these parties to build deeper common understanding may enable future progress.

Second, even where major state powers currently seem to broadly agree against banning autonomous weapons, there is potential for challenges to arise between these states. For instance,

although consensus is building around fully applying international humanitarian law to these weapons systems, future differences are likely to arise in relation to implementation. As it stands, France, Israel, Russia, the United Kingdom and the United States oppose negotiations for new international law on fully autonomous weapons, but China recently expressed its desire to negotiate a new Certain Conventional Weapons (CCW) protocol to only prohibit the use of fully LAWS.

This point relates to a third area, namely how principles such as meaningful human control can be implemented since it is now conceded that high levels of autonomy present new challenges to human control. A general principle is developing surrounding the need to retain human control over weapons systems and the use of force. There are, however, divergent perspectives on what meaningful human control constitutes, and how it can be achieved which require deeper examination. In addition, ICRC lawyers welcome renewed interest in weapons reviews, including the important emerging commonality of the legal requirement for States party to Additional Protocol I to the Geneva Conventions to conduct such reviews. A related question, however, is how and by whom, international standards for weapons review could be developed and how to monitor their implementation if such reviews were applied more broadly. This is an area for future collaborative work that is already identified by the LAWS Group of Governmental Experts (GGE) and there seems to be room for sharing national practices given the lack of guidance in Article 36 about how to conduct reviews.

Fourth, some experts are advocating transparency and confidence building measures (TCBMs) to prevent risks and unintended consequences such as a new arms race or proliferation to non-state actors. It is not clear, however, whether traditional solutions for conventional conflict will be sufficient for these new risks. Nonetheless, there currently seems to be a willingness to share experiences in national policies and good practices (potentially including the scientific and commercial communities). This could be a useful starting point for initial confidence building efforts. Moreover, undertaking related initiatives to better clarify this field could also be folded into the initial parts of a confidence building strategy given the ongoing need to develop better understanding of this subject area, which continues to be a major stumbling block.

Lastly, in terms of international security challenges, a number of related contemporary matters have long-term strategic implications. While countries such as China and the United States seem to broadly agree against a ban of LAWS within these GGE discussions, this belies the intense present-day strategic competition and risks to regional and international stability that are connected to acquiring future weapons systems' component technologies, tools and tradecraft in areas such as AI, machine learning, big data, computing power and cloud technologies. Recent trade disputes have deeper present-day and long-term strategic ramifications that must be addressed. This type of strategic thinking can take some of the future leaning LAWS disputes out of the distant future realm where experts struggle to find working examples and case studies for future autonomous systems. This would link, for example, to current debates about Huawei connections to the Chinese government, Chinese law as it pertains to its corporations' national security obligations, and global 5G bids. The United States Executive Order on AI now includes, for instance, the goal of protecting critical AI technologies from acquisition by strategic

competitors and adversarial nations. It is possible that economic interdependence could bring these leading powers to the table, but this requires closer examination. To conclude, any further work in these fields should consider that these types of conversations about leading "AI" powers should ideally be couched in wider geopolitical trends such as the Thucydides trap and expert observations that Americans have suddenly grown fearful of Chinese power.

Strategic Implications Related to Cyber-Enabled Autonomous Weapons Systems

Although some of the emerging and future technologies under examination in the area of lethal autonomous weapons systems (LAWS) may arguably not even exist in future, this field continues to be important given the risk that leaps in technological change can potentially be exponential rather than linear.[1] This is especially worrisome where linear approaches are currently employed for strategic thinking.[2] Existing technological capabilities are already conceptually challenging for the international security architecture and security strategies such as deterrence and arms control for reasons that include the speed of technological change, increasing complexity, and rising vulnerabilities. This situation is, however, exacerbated by technologies such as AI, machine learning, big data capabilities, and computing processing powers, among others, that are maturing in such a way that they have potential to become increasingly independent from human control and their levels of complexity are either already, or predicted to be, beyond human comprehension.

There already seem to be simple forms of highly automated and autonomous cyber-enabled weapons that are independent of human control today, albeit defensive and "generally limited to systems that are supervised by humans that protect vehicles and military bases from attacks"[3]. Some systems are being deployed in offensive roles such as certain missiles and loitering munitions, and it is argued that Stuxnet was an autonomous weapon on account of its capabilities to learn and adapt.[4] The component technologies of these physical autonomous systems are not at the same stage of development, however. AI, for example, is infamous for both periods of little progress as well as periods of heightened attention. In fact, some political scientists currently warn that innovation is slowing down and future trends may comprise the use of low tech threats such as hybrid and information operations to overthrow adversaries rather than the use of such high tech – they too caution about the current level of hyperbole, including the now infamous warnings within the 2015 open letter from AI and robotics researchers.[5]

Nonetheless, there is significant uncertainty and red flags continue to be raised within the technological and scientific community about whether highly autonomous systems can in fact be controlled and whether it will always be possible to control these systems. Full verification of their safety and behaviour may be difficult. On the other hand, some argue that it should be possible to fully examine such systems. These types of challenges that are associated with higher levels of autonomy in weapons systems, while difficult, are likely to be surmountable. Nonetheless, the ways in which these emerging technologies might mature is not clear yet nor is the time frame certain, especially for fully autonomous technologies that are not said to exist yet.

It is this aspect – a sense of the intangible – that can make present-day discussions among states, civil society and industry even more difficult. At times, there are more questions than answers in this field. A point evidenced by much of the international community's focus so far on clarifying concepts and reaching basic understanding about the actual nature of these emerging technologies in the area of LAWS. There is a spectrum of technologies that could end up ranging from advanced automation to advanced autonomy to full autonomy, and these terms are sometimes used interchangeably. In order to conduct meaningful discussions and to make concrete progress in this field, key stakeholders must continue to clarify the meaning and understanding of key concepts.

In very broad terms, several ongoing key strategic questions include the following issues. First, how can stakeholders work to ensure that the benefits that can be derived from these technologies for commercial/civilian needs as well as enhanced security and defence are not impeded by their future dual-use? It is likely that increasing pressure will continue for maturing autonomy to be developed for cyber defence and cyber resilience. It is likely too that the importance of these autonomous intelligent systems will continue to grow for other defence purposes. Moreover, it is highly probable that the component technologies driving cyber-enabled LAWS (such as AI) can be used for both civilian and military purposes given their often dual-use nature. In this current phase of AI "hyperbole", it is the private sector that is driving many developments in parts of the world. It is thus a difficult problem for decision-makers to balance where they must be cautious that policy restrictions do not impede innovation while facing the danger that technologies for civilian use could be used for lethal weapons. The EU position, for example, is that given the dual use of emerging technologies, policy measures should not hamper civilian research, including AI.[6] The 2018 LAWS GGE report similarly notes the delegation agreement that future policy measures should not hinder progress in or access to peaceful uses of intelligent autonomous technologies. In fact, the organisers of workshops on safety and control for AI for a White House public report also explain that many technical leaders think that the main limits on deriving benefits from AI are in the confidence in the safety of these "smart systems".[7]

Second, an open-ended question for future study and state consideration is whether, and how, increasingly autonomous technologies can change the balance of state power. Statements, such as Putin's now infamous remarks that the nation that leads in AI "will be the ruler of the world" naturally require closer examination for what this means for global affairs.[8] Third, numerous assertions continue to be made relating to the benefits and risks associated with the military and economic use of these technologies (for example, assertions that such technologies may reduce the risk to military lives), but these statements are not always founded on strong evidence-based analysis. Such stated benefits and risks require unpacking and deeper research to be conducted together with state and non-state stakeholders. Fourth, in terms of strategic certainty, there is a concern that unlike systems reliant on obvious codes or rules, machine-learning systems cannot be asked why a decision was made. There is no clear code branch to explain the output – neural networks are considered to be black boxes where outputs are unpredictable. A key question then

becomes whether transparency can in fact be increased – is transparency as a potential policy solution (as suggested by the Global Commission for Cyber Stability on AI questions) and traditional cooperative measures between states even possible? It is even more destabilising if there is a risk that such systems are vulnerable to unforeseen applications or tactics and unexpected ways of use – there is apparently no way of knowing the consequences of the manipulation of machine-learning systems. In other words, *what unpredictable outputs will unpredictable inputs produce?*

It seems that the current trend in the civilian sector is a move away from rules based systems to machine learning based systems which allows such systems to become fully autonomous by optimising functions such as minimal harm and risk.[9] There is thus a concern about the black box nature of some of these decision-making algorithms and their scalability, and unanswered questions include whether advances in "explainability" could address this concern.[10]

In terms of discussions that have so far taken place among states and non-governmental organisations at the international level, the focus has mainly surrounded physical (but cyber-enabled) LAWS platforms such as land, sea, and air (cyber weapons are excluded) under the rubric of the Conference of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects (CCW). In 2016, the Fifth Review Conference of the High Contracting Parties to the Convention on CCW established an open-ended GGE on "Emerging Technologies in the Area of Lethal Autonomous Weapons Systems" which held its first meeting in November 2017.

The High Contracting Parties then agreed in 2018 that the GGE on LAWS would meet again in March and August 2019 for a shorter duration of seven days.[11] Ljupco Jivan Gjorgjinski, Minister Counselor and Chargé d'Affaires at the Permanent Mission of the Republic of Macedonia in Geneva replaces former Ambassador Amandeep Singh Gill of India as the GGE Chair. That said, Amandeep Singh Gill is currently serving as Executive Director and co-Lead of the UN's High Level Panel on Digital Cooperation which has a mandate to consider cooperation across subject domains in order to realise the potential of digital technologies while safeguarding against risks and unintended consequences. This more recent focus on identifying cross-cutting subjects and safeguarding against unintended consequences is particularly helpful where component technologies behind cyber-enabled LAWS may be relevant to discussions in other multilateral forums such as the former UN GGEs on cyber.

In terms of the LAWS GGE deliberations, while these discussions have so far failed to properly agree upon a definition of the problem set, they have made some progress by working to better clarify key concepts. More recently, it seems that consensus is building within the GGE process for principles such as meaningful human control and ethical acceptability as agreeable standards. One of the most contentious questions to now explore between state and non-state stakeholders, however, is how these standards or principles, such as meaningful human control, can in fact be applied.

In short, the LAWS GGE under the CCW is recognised as a forum that can constructively help states surmount these challenges by sharing best practices and gathering inputs from experts, academics and civil society.[12] In terms of the current status quo, the 2018 GGE affirmed a number of guiding principles, which the 2019 group is expected to build upon, and the Chair then outlined a number of outstanding issues, which include, among others, the following points:

(1) International law, in particular the United Nations Charter and international humanitarian law (IHL), as well as relevant ethical perspectives should guide the continued work of the Group – IHL "continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems";

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

(3) When developing or acquiring new weapons systems based on emerging technologies in the area of LAWS, physical security, appropriate non-physical safeguards (including cybersecurity against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered;

(4) Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems;

(5) Consideration should be given to the use of emerging technologies in the area of LAWS in upholding compliance with IHL and other applicable international legal obligations;

(6) In crafting potential policy measures, emerging technologies in the area of LAWS systems should not be anthropomorphized;

(7) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies;

(8) The CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of LAWS within the context of the objectives and purposes of the Convention, which seeks to strike a balance between military necessity and humanitarian considerations;

(9) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system;

(10) Accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control; and

(11) Technical characteristics related to self-learning (without externally fed training data) and self-evolution (without human design inputs) *must be further studied*. Some delegations considered differentiating autonomy from semi-autonomy or automation, while others consider autonomy on a spectrum noting that it is not an "on/off phenomenon and there is lack of a clear line beyond which human control is lost or autonomy becomes problematic".[13] There is thus a gap that requires the international cyber and AWS policy communities to deal with the impact of very high levels of autonomy in cyber weapons – for example, should the LAWS GGE principle of meaningful human control not apply by extension to fully autonomous cyber weapons, and if so, is it even possible to achieve such meaningful human control for these systems? Supporting such a study or, alternatively, convening an informal group of both cyber and LAWS experts to deeply examine the policy implications surrounding the technical characteristics related to self-learning and self-evolution would add much-needed analysis to both these fields. Moreover, it would straddle the demarcation between the separate UN forums on cyber and LAWS, potentially helping to avoid future obstacles.

Key Protagonists: Identifying Current and Future Friction Points Between Stakeholders

Civil society groups have so far played a very active role in the debates shaping the development of LAWS. Different groups continue to be active in this space – for example, a small demonstration against the use of self-automated weapons took place in Brussels before a European Defence Agency conference in November 2018 on AI/unmanned military systems.[14] On a larger scale, groups such as the Campaign to Stop Killer Robots have been instrumental in shaping and driving the public discourse related to bans and moratoriums for fully autonomous weapons systems. Furthermore, given the open-ended nature of the LAWS GGE under the CCW, such non-governmental organisations can participate in these international discussions. Particularly vocal groups include the International Committee of the Red Cross (ICRC), the Campaign to Stop Killer Robots, and Human Rights Watch (86 NGOs are calling for a ban of fully autonomous weapons). The Campaign to Stop Killer Robots advocates, for instance, that meaningful human control over targeting and attack decisions should be retained by prohibiting development, production and use of fully autonomous weapons – it calls for all countries to commit to creating a new ban treaty.

Non-governmental organisations that participated in the work of the 2018 LAWS GGE include:

- Campaign to Stop Killer Robots, Amnesty International, Article 36, Association for Aid and Relief Japan, Committee of 100 in Finland, Center for International Security

and Policy, Facing Finance, Future of Life Institute, Human Rights Watch, ICT4Peace Foundation, International Committee for Robot Arms Control (ICRAC), Mines Action Canada, Nobel Women's Initiative, Norwegian Peace Foundation, PAX, Pax Christi Ireland, Pax Christi Vlaanderen, Project Ploughshares, Protection, Pugwash Conferences on Science and World Affairs, Rete Italiana per il Disarmo, Seguridad Humana en Latinoamérica y el Caribe (SEHLAC), Women's International League for Peace and Freedom (WILPF), Conscious Coders, International Action Network on Small Arms, Pax Christi International, the Centre for a New American Security (CNAS) and Zonta International.

In addition to such groups, the UN Secretary General António Guterres recently called for States to ban autonomous systems that could independently target and attack human beings at the November 2018 Paris Peace Forum.[15] Twenty-six countries including a group of African States, the Holy See and the Non-Aligned Movement group of states as well as the European Parliament are also calling for a ban or moratorium for fully autonomous weapons.[16]

On the other hand, states such as the United States, China, South Korea, Russia and the UK are apparently developing weapons systems with significant autonomy in the critical functions of selecting and attacking targets.[17] In short, all P5 members of the UN Security Council - the United States, China, the United Kingdom, France and Russia - seem to share the broad position that these weapons systems should not be banned. Israel has also rejected an outright ban.

The UK, France and United States focus upon aspects related to human control over weapon development, deployment and use in their submission of working papers for the 2018 GGE.[18] While China did not advocate for an outright ban in 2018, it did call for a *ban on the use* of fully autonomous lethal weapons systems. This is interpreted to mean that China is unlikely to stop building its own such systems.[19]

The Chief of General Staff of the British Army explained last year that while the imminent dangers from killer robots are over-hyped and catastrophic, there is a genuine need to consider the ethical and legal implications of development in AI and robotics when they are applied in a military context.[20] These are at least some broad starting points for potential collaborative examination among state parties and non-governmental organisations. Such major powers are unlikely to refrain from developing these types of military capabilities, especially where their officials currently posit that these systems are in line with international law. However, the non-governmental organisations involved in this field have exhibited strong convening powers which may cause unwanted future friction. There may, therefore, be an important space for confidence building to be conducted between countries like the United Kingdom, France, Russia, China and the United States together with think tanks/research institutes and other non-governmental organisations to build deeper common understanding about the most contentious questions. For example, organisations like the ICRC, CNAS, and the Future of Life Institute could be a good starting point given their field experience and research rigour. On the states' part, government officials may also agree that a meeting of minds is much needed in order not to prevent future progress.

For example, the UK's position is such that efforts should be undertaken to identify and mitigate the risks posed by new technologies and ensure they are used in compliance with IHL and in accordance with "national policies which are often more exacting."[21] Moreover, *in those states which develop weapons responsibly*, control measures on weapons procurement and use exist. In the case of the UK, it is felt that developments in AI mean it is possible that an unpredictable system could give a better outcome than human decision-making (recognising that such a system would not be making a decision in the same way that a human does since it has no moral agency despite tendencies to anthropomorphise them).[22] Here this is understood to mean that without such a moral agent, it is not possible to ensure compliance with International Humanitarian Law (IHL) – so IHL already prohibits the use of such systems. Moreover, the Chief of General Staff does not see any tactical value in allowing automated execution of unpredictable outcomes when applying lethal force. Nonetheless, it is conceded that *high levels of autonomy present new challenges to human control*.

In terms of longer term strategic thinking, future differences and friction points that could potentially warrant confidence building could arise *between* major state actors, even where they now seem to broadly agree against banning autonomous weapons. These areas could be identified through a closer examination of their current position papers, including working papers on national policies and positions as well as civil society/industry contributions which can be found in Annex 1 of the 2018 GGE Report. For example, China has now expressed a desire to negotiate and conclude a new protocol for the CCW to ban *the use* of fully autonomous lethal weapons systems (rather than the systems).[23] The country is described as continuing to maintain a degree of strategic ambiguity and "apparent preference for optionality" in its diplomatic posture on AWS, with its latest position paper characterising AWS rather narrowly with many exclusions.[24] This compares, for instance, to British perspectives that states must apply already existing international legal and ethical obligations in the development and use of new weapons technology.[25]

Thematic Approach: Identifying Key Subject Areas for Collaborative Efforts Under the Framework of the LAWS GGE Agenda

There is potential to identify subject areas for confidence and trust building under the framework of the current LAWS GGE work agenda which is divided into four sections, namely (a) The potential challenges posed by emerging technologies in the area of LAWS to IHL; (b) The human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of LAWS; (c) Potential military applications of related technologies in the context of the Group's work; (d) Characterisation of these systems in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the Convention; and (e) Possible options for addressing the humanitarian and international security challenges posed by emerging technologies in the area of LAWS.

The below sections identify several divisive areas which will likely warrant parties' common understanding in future. Non-governmental actors could potentially contribute their insights and

experiences to this Group's deliberations by submitting working papers or registering interest in upcoming workshops. Notably, the LAWS GGE has been open to non-state parties, non-governmental parties and international organisations from the outset since it falls under the umbrella of the CCW. This is markedly different to previous cyber GGEs that have so far been held under the UN's first committee (disarmament and international security) where they only comprised governmental experts in the past. This may mean that LAWS negotiations under the CCW forum could evolve in a rather different manner.

(a) Potential challenges posed by emerging technologies in the area of LAWS to IHL

As described in the above sections, even where consensus seems to be building around IHL applying fully to these weapons systems, it is likely that challenges will arise in relation to future implementation which could likely warrant the need for the facilitation of confidence building. As it stands, five states (France, Israel, Russia, the UK and the United States) are opposing moves to negotiate new international law on fully autonomous weapons, while China recently expressed its desire to negotiate a new CCW protocol to only prohibit the use of fully autonomous lethal weapons systems.[26]

As it stands, a number of options are being presented (which apparently may not be considered as mutually exclusive), namely (1) a legally binding instrument; (2) a political declaration; and (3) clarity on the implementation of existing international law obligations, in particular IHL with discussion on the human-machine interface. Others argue that IHL is fully applicable to potential LAWS which means that no further legal measures are needed.

(b) The human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of LAWS and accountability

First, a general principle is developing surrounding the need to retain human control over weapons systems and the use of force. There are, however, divergent perspectives on what meaningful human control constitutes, and how it can be achieved which require much deeper examination. For example, how can meaningful human control be applied over autonomous swarms, or systems that overload human operators with information with little time to make decisions?[27] There is therefore space for future discussions to be held to find common understandings on the extent and quality of the human-machine interaction in the different phases of the weapon system's life cycle, including clarifying the accountability threads in these phases.[28] Given that this gap is already at the forefront of stakeholders' minds, there may be a timely opportunity to add value by convening key state (and potentially non-state stakeholders) to examine and build common understanding on these specific questions.

Second, ICRC lawyers welcome renewed interest in weapons reviews, including the important emerging commonality of the legal requirement for States party to Additional Protocol I to the Geneva Conventions to conduct such reviews. A related question, however, is how and by

whom, international standards for weapons review could be developed and how to monitor their implementation if such reviews were applied more broadly.[29] Moreover, concerns are being raised about verification of weapons reviews that could be perceived as interference in a state's national affairs. An area for future collaborative work that is already identified by the group is the space for sharing national practices given the lack of guidance in Article 36 about how to conduct reviews.[30]

(c) Characterisation of these systems in order to promote a common understanding on concepts and characteristics

There still seems to be a number of different conceptual approaches to characterisation (a list of different attributes and characteristics so far mentioned can be found within the 2018 GGE report). First, some delegations believe that a working definition of LAWS is necessary in order to properly address the risks. Other parties find that the inability to agree upon a definition should not impede discussions within the CCW, although it was noted that a key obstacle is the absence of working samples and a common understanding on a working definition. Future collaborative work must continue be organised to address this gap in common understanding on issues related to emerging technologies in LAWS, including education and the deepening of collective understanding as already identified in previous GGEs. This type of work could be an initial "quick win" and first step for organisations which may be interested in collaborating with stakeholders within the field.

(d) Addressing the humanitarian and international security challenges such as new arms races – "lots of questions without answers"[31]

The literature in this field and the 2018 GGE report highlight other humanitarian and international security risks. These include, for example, the following issues: (1) the risk of harm to civilians and combatants in armed conflict in contravention of IHL obligations; (2) exacerbating regional and international security dilemmas through new arms races and the lowering of the threshold for the use of force; (3) proliferation, acquisition and use by terrorists, (4) vulnerability of these systems to hacking and interference; and (5) the undermining of confidence in the civilian uses of related technologies.

Notably, some experts are advocating transparency and confidence building measures to prevent risks and unintended consequences such as a new arms race or proliferation to non-state actors. However, it is not clear that traditional solutions such as TCBMs will be sufficient for these new risks – 21st century solutions need to be further developed. In the field of cyber, for instance, there is a debate that transplanting traditional military CBMs for conventional conflict to the cyber field as cyber CBMs is not effective. Nonetheless, given that there currently seems to be a willingness to share experiences in national policies and good practices (potentially including scientific and commercial communities), this could be a useful starting point for initiating confidence-building efforts.

However, the need to develop better understanding of this subject matter continues to be a major stumbling block when trying to address these issues properly and pursuing concrete options.[32] Thus, undertaking measures and initiatives to better clarify the field could be folded into the first parts of a confidence building strategy. While this may seem to be a minor point, there is a real concern that future progress and international agreement on ways to address these challenges could be impeded by the inability to converge on a definition of the capabilities which cause concern.[33]

(i) Achieving progress beyond the confines of the LAWS and Cyber GGEs: Collaborating outside these structures on existing tensions related to LAWS' and strategic technologies/assets such as AI, machine learning, big data, computing power, and cloud technologies[34]

Under this last pillar of international security challenges, there could be space for parties to address related contemporary matters that have long-term strategic implications. While countries such as China and the United States seem to broadly agree against a ban of LAWS within these GGE discussions, this belies the intense present-day strategic competition and risks to regional and international stability that are connected to component technologies such as AI, machine learning, big data, computing power and cloud technologies through races to acquire companies, talent, IP and data through legal and sometimes illegal means. New policy instruments from the United States and the EU, which have been released over the past year, seem to be exemplify an awakening to other long-term strategic risks related to present day activities in these technology fields. This is especially the case in EU Member States such as Germany following, for example, warnings from United States' officials about the foreign acquisition of robotics companies.

Key areas for state collaboration are becoming clearer with the increasing national security emphasis on AI recently. For example, the EU released the EU Strategy on AI in December 2018 and its Global Tech Panel Members will provide input in the coming months so that the development of AI, which can be used in weapons systems, fully complies with international law, respects human dignity and opportunities are also harnessed.[35] China and the United States (not Russia) are regarded as the two frontrunners, with China recently announcing its ambition to become the global leader in AI research by 2030.[36] Moreover, many analysts warn that the United States may fall behind, especially where the Trump administration was seen to cut funding for science and technology research.[37] Future work in this field should therefore consider the role of these two major state powers.

This space is challenging even for bodies like the EU where European defence planning is apparently now facing "two major tectonic shifts": first, the end of Pax Americana and the rise of China as an economic, military and technological power, and second the digital revolution driven by robotics and AI combined with the increase in computer processing power and access to data.[38] Emphasis continues to be made to enhance cybersecurity for European infrastructures but there are now heightened warnings about the *dependence on strategic foreign technologies or harmful foreign technologies such as data analysis tools placed in strategic infrastructures such as cloud servers.*[39] In other words, what can sometimes seem like intractable trade disputes have

deeper present-day and long-term strategic ramifications which must be addressed. This approach and strategic thinking can take some of the future leaning LAWS disputes out of the science fiction or distant future realm where experts currently struggle to find working examples and case studies for future autonomous systems.

Conclusion

It is essential to understand the strategic context behind subjects such as LAWS and other future cyber capabilities or cyber-physical weapons systems in order to better grasp the significance and potential future risks related to long-term strategic economic and military ambitions that possibly underpin current debates (for example, the growing concern about the rise of foreign direct investments that can be used to take control of strategic assets).

In terms of Chinese strategic ambitions, China's involvement with these UN groups is consistent with its commitments under the 2017 AI development plan which calls for China to strengthen the study of major international common problems and deepen international cooperation on AI laws and regulations.[40] The country emphasises the importance of AI to development, but the boundaries between military and civilian applications of AI technology are blurred, especially by its national strategy of civil-military fusion.[41] Experts suggest that China's emergence as an "AI powerhouse" may enable its diplomatic leadership on these questions, while enhancing its future military power.[42] Moreover, the civilian world is seen to have dominance over AI currently which explains why military establishments seek help from MNCs.[43] For this reason, many would argue that a country like China may be better placed in this field than its Western counterparts.

In addition, any work in this area should consider that these conversations about leading "AI" powers must be couched in wider geopolitical trends such as the Thucydides trap and the risk that when a rising power threatens to displace a ruling power, there is a high risk of collision. As Niall Ferguson wrote recently (based upon Graham Allison's book), what made war inevitable according to Thucydides was the growth of Athenian power and the fear this caused in Sparta – he writes that "in the space of barely a year, Americans have suddenly grown fearful of Chinese power. What was once the position of a few alarmists is the new orthodoxy in Washington shared by Republicans and Democrats, foreign policy wonks and technology nerds. We may not be destined for a hot war, but we certainly are on track for a cold one".[44] Ferguson writes that in the Cold War the launch of the Soviet satellite Sputnik in 1957 was the moment America woke up to the red menace, but he is not quite sure what the Chinese Sputnik moment was – suggesting the publication last year of Kai-fu Li's "AI Superpowers: China, Silicon Valley and the New World Order".[45] Dr. Kai-Fu Lee writes in his book, however, that AlphaGo's victory in 2016/2017 was China's sputnik moment for AI. He explains that the country is ramping up AI investment, research and entrepreneurship on a historic scale – following this sputnik moment in 2017 the Chinese central government issued an ambitious plan to build AI capabilities. It called for greater funding, policy support, and national coordination for AI development. It set clear benchmarks for progress by 2020 and 2025, and it projected that by 2030 China would become

the centre of global innovation in AI, leading in theory, technology, and application. He explains that by 2017, Chinese venture capital investors had already responded to that call, pouring record sums into AI start-ups and making up 48 per cent of all AI venture capital funding globally, surpassing the United States for the first time.

Notably, Ferguson now concludes that *"China bashing is no longer about unfair trade policies and the loss of manufacturing jobs, the trade war that Trump launched against China last year has morphed into a tech war over 5G networks, AI, online payments and quantum computing."*[46] Kai-Fu Lee thus argues that because of these developments in AI "dramatic changes will be happening much sooner than many of us expected" – *how these two countries will choose to compete and cooperate in AI will have dramatic implications for global economics and governance*. This is a hugely important point for all state and non-governmental actors to bear in mind while hoping to identify and build future collaborative and confidence building structures in these fields.

Ferguson mentions that China and America are economically intertwined which makes this "new cold war" different.[47] This could be a silver lining – such economic "intertwining" may be a catalyst to bring these leading powers to the table. Their mutual interests may at times converge to ensure that the economic benefits associated with such emerging technologies are secured, while weighing this up with related military and security interests. First, both states may not wish to undermine confidence in the civilian uses of related technologies. Second, they may not want to impede progress in or access to peaceful uses of intelligent autonomous technologies (and component technologies such as AI, machine learning, big data and cloud computing – deep learning/machine learning is underpinned by vast amounts of data that are needed). And third, they may prefer to avoid economic protectionism, balancing their ambitions for future military prowess associated with these technologies.

On the United States' part, the United States Department of Defense and parts of the intelligence community are now apparently "fully seized of the AI issue" and are actively pursuing a number of initiatives.[48] More recently, in the face of criticism, President Trump's Executive Order on AI recognises the implications for the economy and national security, stating that the United States must be an AI leader by sustaining and enhancing the scientific, technological and economic leadership position of the United States in AI R&D.[49] The Executive Order *addresses the link between AI and big data* including strategic objectives such as enhancing access to federal data and computing resources *while maintaining security, privacy and confidentiality protections*. Naturally questions are being raised about China's intended use for these technologies and the large volumes of data needed to support AI learning activities. Some experts note that China has a "disproportionate advantage" in terms of the volume of data about human behaviour to which its AI developers have access because of its "intensive (and repressive) collection of data about the activities of its very large population".[50] This thus links into current debates about Huawei connections to the Chinese government, Chinese law as it pertains to its corporations' national security obligations, and Huawei's 5G network bids in other parts of the world.

Key principles in the American Executive Order now include the goal of protecting critical AI technologies from acquisition by strategic competitors and adversarial nations, and minimising vulnerability to attacks from malicious actors.[51] Similarly, the European Commission proposed the EU Framework for screening foreign direct investment (which was agreed in November 2018) on the grounds of security and public order. By the end of 2018, the Commission was due to carry out an analysis of FDI flows into the EU, focusing on strategic sectors and assets such as key technologies, critical infrastructure and sensitive data. This is particularly the case when the investor is owned or controlled by a third country or benefits from significant state subsidies.

To conclude, this dilemma "[i]s a matter of economic growth, and it is a matter of security."[52] This means that informal initiatives, which intend to work on easing contemporary trade tensions related to technologies and the new "tech war" between major powers like the United States and China, would do well to consider the wider strategic lens and long term strategic implications beyond the trade domain as they relate to certain technologies. Beyond economic competitiveness, surveillance and questions related to civil liberties, future weapons systems will require these large data sets, AI technologies, and computing processing powers.

---

[1] This first section was previously published at: Caitríona Heinl, "A short primer behind the upcoming open-ended group of governmental experts meeting on emerging technologies in the area of lethal autonomous weapons systems", https://caitrionaheinlcyberpolicywatch.wordpress.com/2019/03/18/a-short-primer-behind-the-upcoming-open-ended-group-of-governmental-experts-meeting-on-emerging-technologies-in-the-area-of-lethal-autonomous-weapons-systems/, 18 March 2019.

[2] See Paul Cornish and Kingsley Donaldson, 2020: World of War, July 2017.

[3] https://blogs.bard.edu/dronecenter/cnas-primer-on-laws-for-un-delegates/

[4] CNAS, 4. Scharre, Horowitz, Sayler, 1.

[5] Rid, New America.

[6] https://eeas.europa.eu/topics/instrument-cooperation-industrialised-countries/51679/international-security-and-lethal-autonomous-weapons_me

[7] Carnegie Mellon University Workshop.

[8] https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world

[9] Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, Presentation Comments by Professor Pascale Fung, Director of the Centre for AI Research, Hong Kong University of Science and Technology, 23 October 2018.

[10] Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, 23 October 2018.

[11] The first session of the GGE on LAWS in 2019 took place on 25-29 March. The second session is scheduled for 20-21 August 2019.

[12] https://www.weforum.org/agenda/2018/05/banning-autonomous-weapons-is-not-the-answer/

[13] 2018 GGE Report.

[14] https://www.theparliamentmagazine.eu/articles/opinion/ai-unmanned-military-systems-finding-opportunities-amidst-challenges

[15] https://www.stopkillerrobots.org/2018/11/unban/

[16] https://www.stopkillerrobots.org/2018/08/sixthmeeting/

Also see under the 2018 GGE report, "General Principles on Lethal Autonomous Weapons Systems, Submitted by the Bolivarian Republic of Venezuela on behalf of the Non-Aligned Movement (NAM) and Other States Parties to

the Convention on Certain Conventional Weapons".

[17] https://www.stopkillerrobots.org/learn/

[18] Accessible under the 2018 GGE Report: (1) Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles. Submitted by the United Kingdom of Great Britain and Northern Ireland; (2) Human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems. Submitted by France; and (3) Human-Machine Interaction in the Development, Deployment, and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. Submitted by the United States of America.

[19] https://www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems
[20] https://www.weforum.org/agenda/2018/05/banning-autonomous-weapons-is-not-the-answer/
[21] Ibid.
[22] Ibid.
[23] https://www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems
[24] Ibid.
[25] https://www.weforum.org/agenda/2018/05/banning-autonomous-weapons-is-not-the-answer/
[26] https://www.stopkillerrobots.org/2018/04/convergence/
[27] https://www.weforum.org/agenda/2018/05/banning-autonomous-weapons-is-not-the-answer/
[28] Chairman's notes, 2018 GGE report.
[29] Knut Dörmann, Head of the Legal Division and Chief Legal Officer, ICRC, 2018 GGE Report.
[30] Chairman's notes, 2018 GGE report.
[31] Kai Fu Lee.
[32] Chairman's notes, 2018 GGE report.
[33] https://www.weforum.org/agenda/2018/05/banning-autonomous-weapons-is-not-the-answer/
[34] See the draft book chapter: Caitríona Heinl, "Maturing autonomous cyber weapons systems: Implications for international security cyber and autonomous weapons systems regimes", In the *Oxford Handbook of Cyber Security*, Oxford University Press, Ed. Prof. Paul Cornish (forthcoming).
[35] https://eeas.europa.eu/topics/instrument-cooperation-industrialised-countries/51679/international-security-and-lethal-autonomous-weapons_me
[36] https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world
[37] Ibid.
[38] European Commissioner for Digital Economy and Society, Mariya Gabriel, https://www.theparliamentmagazine.eu/articles/opinion/ai-unmanned-military-systems-finding-opportunities-amidst-challenges
[39] European Commissioner for Digital Economy and Society, Mariya Gabriel, https://www.theparliamentmagazine.eu/articles/opinion/ai-unmanned-military-systems-finding-opportunities-amidst-challenges
[40] https://www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems
[41] Ibid.
[42] Ibid.
[43] Professor Mary Cummings, Duke University, Fellow of AIAA and Co- Chair of WEF's Council on Artificial Intelligence and Robotics, 2018 LAWS GGE Report.
[44] https://www.thetimes.co.uk/edition/comment/in-this-cold-war-between-trump-and-china-beware-the-enemy-within-zz87wk2js

[45] Ibid.
[46] Ibid.
[47] Ibid.
[48] https://www.lawfareblog.com/president-trumps-executive-order-artificial-intelligence
[49] Ibid.
[50] Ibid.
[51] Ibid.

[52] https://www.defensenews.com/global/europe/2018/11/30/eu-members-seek-common-ground-on autonomous-weapons/