



CCW REPORT

Civil society perspectives on the CCW Group of Governmental
Experts on lethal autonomous weapon systems
27–31 August 2018



Reaching Critical Will

29 August 2018
Vol. 6, No. 8

IN THIS ISSUE

- 1 | Human control for human rights
- 3 | Free will and autonomous weapon systems
- 4 | News in brief

The CCW Report is produced by the Reaching Critical Will programme of the Women's International League for Peace and Freedom (WILPF). WILPF is a steering group member of the Campaign to Stop Killer Robots.

The views expressed in this publication are not necessarily those of WILPF or the Campaign.

Editor: Ray Acheson

www.reachingcriticalwill.org
info@reachingcriticalwill.org



HUMAN CONTROL FOR HUMAN RIGHTS

Ray Acheson | Reaching Critical Will, Women's International League for Peace and Freedom

Tuesday's discussions at UN talks on autonomous weapons moved to the Human Rights Council. This is significant, given that the UN began consideration of this issue in the Council in 2013 on the basis of the first report on the issue by Christof Heyn, the Special Rapporteur on extrajudicial, summary or arbitrary executions. The beige walls (and floors, and desks, and chairs) of the usual "disarmament" meeting room disappeared two floors beneath us, to be replaced by the famously vibrant ceiling created by Spanish abstract artist Miquel Barceló. "All of it is a sea upside down, but it is also a cave," Barceló said about his ceiling. "The complete union of opposites, the ocean surface of the Earth and its most concealed cavities." How fitting for discussions about killer robots, which are being developed quietly by a handful of countries while the world's governments come together in a room dedicated to "Human Rights and the Alliance of Civilizations" to wrestle with the philosophical, ethical, technical, and legal questions about the increasing mechanisation of violence and the further removal of human beings from accountability for the use of force.

Appropriately, Tuesday's focused primarily on the concept of human control—which many governments and activists tie directly to the protection of human rights and humanitarian law. Over the course of the last five years, the belief that meaningful human control must be maintained over the critical functions of weapon systems has emerged more or less as a point of consensus amongst participating governments. The question for most states is not if they "have a duty to control or supervise the development and/or employment of autonomous weapon systems, but how that control or supervision ought to be usefully defined and extended," as the Swiss delegation said.

Differences in opinion remain in regards to what constitutes "meaningful" control

and what limits on autonomy are required to ensure this control. On Tuesday delegations continued their examination of what stages of a weapon's life cycle is human control or intervention necessary. In April the Chair released a pie-chart (now apparently affectionately referred to as the "sunrise" diagram) indicating the potential phases in which human control could be relevant in the life of a weapon system. These phases include research and development; testing, evaluation, verification, validation, and review; deployment, command, and control; and use and abort. Some delegations have since suggested additional phases; the UK working paper for this session, for example, adds "national policies" and "battle damage assessment / lessons learned" to the beginning and end of the sunrise.

Selection and engagement of targets seems to be the most common definition of critical functions of a weapon system that require human control. Some, like Japan, thought that autonomy in selecting targets would be fine, but that human control is necessary to initiate an attack. Others believe humans must control both in order to ensure the protection of human dignity and compliance with international law.

Some states have also expressed concern with processes in the development stages of these weapons. Ireland expressed concern about bias in the programming of a weapon system, highlighting the potential for the perpetuation and amplification of social bias, including gender bias, at the programming stage. The International Committee of the Red Cross (ICRC) has argued that humans must maintain control over programming, development, activation, and operational phases of a weapon system, because international humanitarian law "requires that those who plan, decide upon and carry out attacks make certain judgments in applying the norms when launching an attack."

continued on next page

Editorial, continued

Overall, there seems to be convergence around the idea that fully autonomous weapons would not be acceptable, because they would not be able to comply with international law or ethical frameworks. There also seems to be a majority view that a fully autonomous weapon is one that can select and engage targets autonomously, without human intervention (as distinct from, for example, armed drones, which are controlled remotely by humans). It is precisely these critical functions that most states, together with the ICRC and the groups associated with Campaign to Stop Killer Robots, believe must not be left to programming and algorithms.

Only a small minority of states seemed skeptical that maintaining sufficient human control would necessitate limiting a weapon system's autonomous functions. Not unsurprisingly, those arguing against restrictions on autonomous weapons these are the same governments that are engaged in their research and development. The US delegation, for example, argued that states and civil society must not "stigmatise new technologies" or set new international standards, but instead work to ensure "responsible use of weapons".

The rest of the participants at this meeting, in contrast, seem keen to move ahead with delineating limits on autonomy and corresponding rules and mechanisms for meaningful human control now, as a matter of growing urgency. As the Austrian delegation said, the best way to settle the aforementioned differences in opinion on human control and critical functions is to engage in negotiations of new international law.

The majority of governments, including 26 countries that have called explicitly for a prohibition on fully autonomous weapons, have previously supported this call. The tech, academic, and scientific communities engaged in work on autonomous technologies and artificial intelligence (AI) also broadly support this call. A side event hosted by the Campaign to Stop Killer Robots on Tuesday featured three individuals who have helped organise these communities to give voice to their opposition to the development of autonomous weapons. "We keep hearing questions about what the private sector is saying," remarked Peter Asaro of the International Committee on Robot Arms Control (ICRAC), who helped coordinate a letter by 1200 academics in support of Google employees demanding their company cancel its Project Maven contract with the Pentagon. "The private sector is asking states to take action and ban killer robots." Likewise, Amr Gaber of the Tech Workers Coalition pointed out that tech workers are actually taking the lead in this work, but they cannot do it alone. As a self-described "low-

level software engineer" with Google, he helped coordinate the letter signed 4000 Google employees who demanded the company cancel Project Maven and institute a policy against taking on future military work. Banning autonomous weapons is the first step to curbing the potential harm of technology, he argued, but states and civil society need to do more to make sure technology remains in the service of human rights, not repression and violence. "We're tired of hearing this it too hard to solve," he said. "We are standing up for justice."

At the end of the day, the ask is simple. Weapons must be under human control. The alternative, as ICRC President Peter Maurer has written, is a future where humans are "so far removed from wartime choices that life-and-death decision making is effectively left to sensors and software," and "where wars are fought with algorithms, with uncertain outcomes for civilians and combatants alike." This is why the Campaign to Stop Killer Robots and a growing number of governments are calling for the negotiation of a legally binding treaty to prohibit fully autonomous weapons. It is also why we believe, as Amr Gaber noted, that it will be important to hear not just from governments, or even from tech workers and scientists, but also those who will be impacted by these technologies in the future. If these weapons are developed, human beings around the world will suffer. Human rights will be undermined. They will be used to repress, to harm, to kill. It is the protection of human rights and dignity that has motivated our previous work for humanitarian disarmament. It is this that motivates our work for stronger laws and norms to prevent the increasing abstraction and mechanisation of violence. A different future is possible and we have the capacity to build it. •



Image: Russell Christian/Human Rights Watch

FREE WILL AND AUTONOMOUS WEAPON SYSTEMS

Alimzhan Akhmetov | Center for International Security and Policy (Kazakhstan)

A prominent philosopher of the 20th century, Nikolai Berdyaev, believed that the key question of philosophy lies in the problem of the relationship between the free will of humankind and the free will of God. Humanity's free will ends at a place where the freedom of God's will begins. Yet where does that line start?

The rapid development of artificial intelligence shows the acuteness of this problem from another perspective, especially in the area of the development of lethal autonomous weapons systems.

Undoubtedly, this is a direct contradiction to international humanitarian law and, in particular, the Martens clause. However, this is only the tip of the iceberg.

The origins of the problem lie much deeper. The philosophical foundation of the challenges in developing technology and artificial intelligence that humankind face were uncovered by Nikolai Berdyaev in 1931 in his work, *The Spiritual Condition of the Contemporary World*.

The problem he articulates is that humanity took on God's burden but failed to carry it.

Now, the question facing humankind, in the development of artificial intelligence or autonomous weapons is as follows: where does the freedom of a machine's will and autonomy end, and the will of humanity begin?

In this regard N. Berdyaev aptly notes: "Man [sic] is shaken and crushed by the might of technology, making all his life topsy-turvy. Man himself has created it, it is the product of his genius, of his reason, of his inventiveness, it is a child of the human spirit. Man has succeeded in unlocking secret powers of nature and using them for his own ends, of intro-

ducing a teleological principle into the activity of mechanical-physical-chemical powers. But to master the results of his work man has not succeeded. Technology has come to seem more powerful than man himself, it subjugates him to itself. The crisis of our time is to a remarkable degree begotten by technology, which man lacks the strength to deal with."

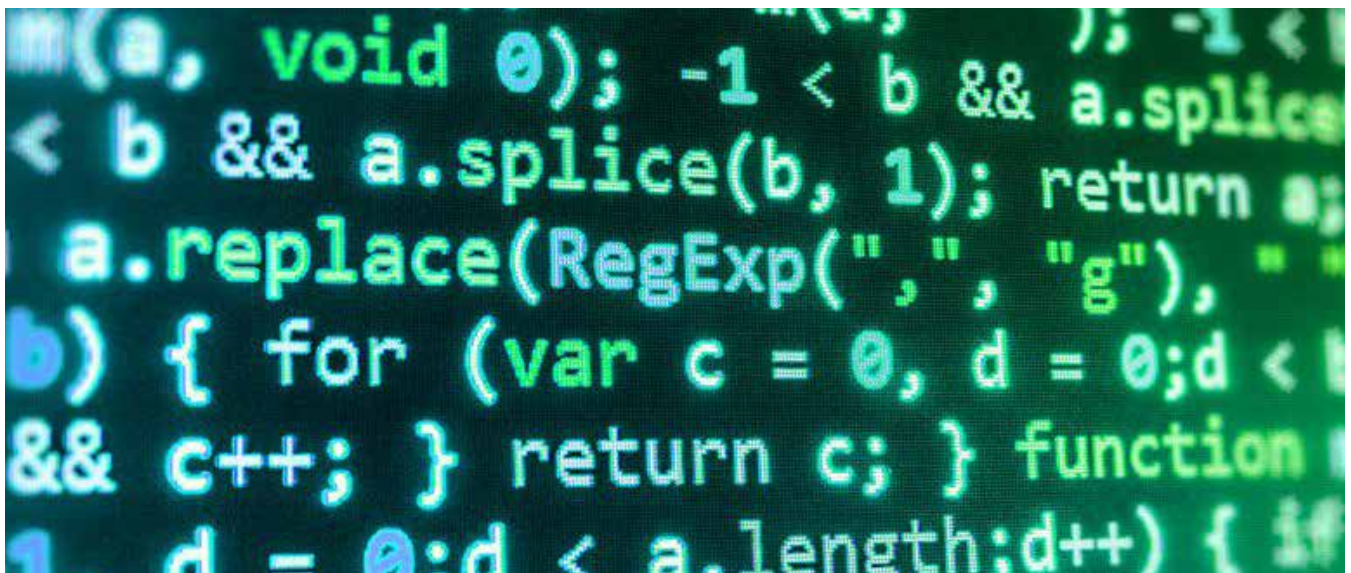
Further he says: "Technology has ceased to be neutral... Man has succeeded in creating a new world. Within the machine is present the reasoning power of man, within it operates a teleological principle. Technology creates an atmosphere, saturated with energies, which earlier were hidden within the depths of nature. And man has no assurance that he is in a condition to breathe in the new atmosphere. Into the hands of man technology puts a terrible and unprecedented power, a power, which can be to the destroying of mankind."

Going back to initial question of freedom of will, we need to remember two key characteristics of artificial intelligence.

Nobody can guarantee that a machine, which is being used in peaceful way, will not "decide" to act against a human at a certain point.

The development of such systems has far-ranging unpredictable consequences.

We still have time and opportunity to draw a clear boundary between the machine and humanity, and ensure the machine remains under meaningful human control. This boundary exists in the unconditional recognition of life and the right for life as the highest values, and the recognition of the principles of morality and ethics to be solely the sphere of responsibility of a person. •



NEWS IN BRIEF

Allison Pytlak and Katrin Geyer | *Reaching Critical Will*, Women's International League for Peace and Freedom

The following is a summary, not a comprehensive report, of key discussion points of Tuesday's talks.

Policy responses and process

- Costa Rica called for a legally binding instrument to retain meaningful human control over weapon systems and the use of force.
- Austria agreed that it is time to begin elaborating a legally binding instrument on this.
- Brazil noted the correlations between the discussion about characteristics, the human element, and possible options as leading to a legally binding instrument in form of an additional protocol to the CCW.
- Republic of Korea (ROK) maintained that IHL applies to autonomous weapons systems (AWS) but this does not necessarily mean they should be prohibited.
- Argentina said there is a legal vacuum because of the lack of clarity if AWS will comply with principles under IHL.
- Germany referenced its draft joint political declaration with France that assures a commitment to uphold the principles of human control.
- United Kingdom (UK) opposed the implementation of a legal instrument regulating AWS as the debate is "not yet ripe".
- Switzerland sees the merit in extending discussion on human control in future meetings and is interested in proposals to define "minimal levels" of human control and to discuss if this is possible.
- United States (US) is not in favour of stigmatising new technology or seeking new international standards but advocates for responsible use by implementing holistic and proactive review processes guided by principles of the laws of war.
- Russia noted that AWS do currently not exist and that current military systems with high levels of automation should not be placed in a separate category that would require specific legal rules or any prohibitions or restrictions that the category may entail.
- Pakistan stated that the outcome of this GGE must be to confirm that weapons with autonomy stay under control of humans. It said that it is a positive development that there is general understanding that weapons with autonomous functions must remain under direct control of humans at all times.
- India supported working toward promoting common understanding that will enable flexibility.

Approach to/utility of definitions

- China said that there has to be a clear definition of AWS, and it suggests further detail and clarifications on formulations by giving concrete *modus operandi* instead of general concepts.
- Russia believes it to be necessary to have a clear and common understanding and basic characteristics of a working definition of AWS. It proposed that AWS be defined as "unmanned technical means other than ammunition, that are designed to carry out combat and support tasks without any participation of an operator."
- International Committee for Robot Arms Control (ICRAC) urged simple definitions and a focus on targeting, the application of force, and meaningful human control.
- Pakistan observed that concepts like meaningful human control have gained traction, but they provide an approach and not a solution to the technical, legal, moral, and regulatory questions being raised.
- International Committee of the Red Cross (ICRC) said that no matter which words are used (i.e. "minimum" "sufficient" "effective") human agency must be preserved. Austria expressed a similar view.
- The US agreed that it is not necessary to arrive at a specific definition to continue discussing how IHL applies to emerging technology.
- Mexico said the future work of the GGE could focus on identifying what is meaningful human control as related to the critical characteristics of AWS.

Human-machine interaction, degrees of autonomy, and meaningful human control

- Russia believes the notion of meaningful human control being promoted by some states includes significant risks of politicising discussions and future work. It said that the elaboration of a working definition should not lead to artificial divisions of "good" and "bad" weapon systems.
- ROK stated that meaningful human control must be maintained in AWS in their development, deployment, operation, especially in critical functions of selecting, and attacking the target. It noted that humans are not perfect and that autonomous technology can complement human decisions.
- Germany noted that any human-machine interaction has to be subordinate to human operators.
- India stated that human control must be maintained over all weapon systems, including AWS,

continued on next page



News in brief, continued

- and that maintaining control over critical functions is essential.
- Belgium encouraged states to focus on ensuring sufficient, meaningful human control in the design and meaningful human control in the use of AWS.
 - The Netherlands agreed with the UK that two more “touch points” should be added to the “sunrise chart” drafted by the Chair in April 2018: political control, and analysis to see if a system is being used as intended.
 - Brazil referenced the sunrise chart provided in April and suggested that future work should focus on the phases represented by the last quarter of the graph.
 - Switzerland noted that human control can be exerted in different stages of a weapon systems’ life cycle and suggested the consideration of training in the phases of deployment and command and control, so to help ensure compliance with IHL.
 - Estonia said that the sunrise table is a helpful framework complemented by the indicative list of activities provided in the UK working paper. It said it would be useful to explore further how specific activities at different stages of the life cycle would contribute to this outcome.
 - The Netherlands maintained that AWS should be programmed to operate in certain pre-programmed conditions that cannot be altered by AWS themselves. It said that meaningful human control has to be exercised over the complete targeting cycle and the whole lifecycle of a weapon and humans must assess reliability and predictability.
 - European Union (EU) said that the deployment of a weapon system in armed conflict must be under human control in relation to the use of lethal force.
 - Japan said that further discussion is needed to determine at what stage meaningful human control or appropriate human judgment must be guaranteed. It said that the human element is crucial in regulating AWS.
 - Ireland said that human element must be retained in critical functions of weapon systems and that machine autonomy in use of force raises a myriad of legal, moral, ethical questions.
 - The US believes the use of autonomy in weapons systems can improve the degree of human control over the use of force and can reduce negative humanitarian impacts on civilians.
 - Switzerland stated that human control will always be required over deployment and use, regardless of the level of human control already embedded in the weapon’s design.
 - New Zealand said that a human centric-approach is needed that considers the life cycle of weapon systems and how human control is applied throughout in light of IHL.
 - Argentina agreed with the UK human-machine interface is cross-cutting and covers the entire life cycle.
 - France highlighted two areas where the human-machine interface is relevant and noted that IHL is always applicable.
 - ICRAC identified three conditions to be met to ensure meaningful human control: a) that human commanders or operators have full contextual awareness for each and every attack; b) there is active cognitive participation in every attack that allows for deliberation, if needed; c) there are means to suspend or abort an operation or activity. It referred to the guidelines it published for the April 2018 GGE.
 - Norway has not yet concluded a specific legal definition of “fully autonomous weapon systems” but stated that it generally uses this to refer to systems that can select and attack without meaningful human judgment and control and that have autonomy or elements of autonomy in their critical functions. It said the weapons described at the GGE, such as by Sweden, would fall under this category.
 - Pakistan expressed that autonomous weapons can never be predictable or reliable, and outlined many possible reasons for their failure.
 - The ICRC shared that a recent workshop revealed that the need for control increases with the complexity of an environment and that while “human on the loop” interaction can be useful to mitigate for problems, it is not sufficient to guarantee risks.
 - Austria noted that the targeting process requires complex analysis on legal, political, and ethical grounds, all of which are deeply human and cannot be transferred into algorithms.
 - The ICRC affirmed the inverse relationship between autonomy and meaningful human control while the United States believes that more autonomy brings greater control.
 - Chile suggested that understanding meaningful human control involves following levels of predictability and ensuring supervision, including

continued on next page



News in brief, continued

being able to intervene in light of changes in a mission or operation.

- India outlined that characteristics could include technological capabilities, physical attributes, autonomy, and lethality.

Accountability, international humanitarian law (IHL), international human rights law (IHRL)

- ROK stated that IHL is applicable throughout the full life cycle of AWS, including the development and application.
- Germany believes that humans must be accountable for the weapons they use and that humans, not machines, must maintain ultimate decision in matters of life and death.
- Argentina reiterated that the principles of respect for human life, proportionality and discrimination under IHL are legally binding.
- Belgium noted that rights and obligations apply to human, and not robotic agents, and the use of AWS would not exempt the human agent in any case.
- The EU said that the type and degree of human involvement needs to be discussed so to ensure compliance with IHL and IHRL and to respect ethical principles. It referenced national Article 36 reviews and encouraged to share information to contribute to confidence-building and transparency.
- Japan noted that appropriate human involvement is necessary to comply with IHL and IHRL, and for moral and legal responsibility and accountability.
- Ireland said that full compliance with IHL is only ensured by retaining human agency and human control must be ensured in critical functions of the weapon system.
- Switzerland stated that given the current state of robotics and AI, it is difficult to conceive of an AWS that could operate with full compliance with IHL.
- Russia stated that legal responsibility should be accorded to both states and individuals.
- Pakistan said that lethal autonomous weapon systems create an accountability vacuum because attribution of responsibility is difficult. It noted that the use of these systems is unethical and unlawful, and that its use in the battlefield against human soldiers would amount to one-sided killing.
- The ICRC pointed out that legal concerns will arise where autonomy in the design or use pre-

vents a commander or operator from making the judgments that are necessary under IHL. Ethical considerations also demand human control over weapon systems and use of force.

- The US recommended that accountability not be viewed as “handing decision-making ability over to machines”. It said a better framing is that machines have been engineered to make decisions in the environment that they are meant to be deployed in.
- Estonia asserted that humans must retain ultimate control over decisions of life and death, which is a requirement that follows from IHL.
- Chile noted that preservation of human control is a key link for the attribution of legal responsibilities. Limiting unpredictability is important.
- Estonia stated that commanders and operators have a special role in compliance and should only rely on a weapon system with autonomous functions if they are certain that the weapon and the situation will not lead to breaches of international law.
- Mexico stated that human control being maintained at all phases is of crucial importance for identifying legal, including criminal, responsibility.

Artificial intelligence (AI)

- Costa Rica mentioned the UN Secretary General’s presentation to the Conference of Disarmament that called on all stakeholders to work together and ensure the development of peaceful applications of AI.
- France believes that the application of AI does not replace human command but helps humans to process quantities of information and can improve decisions that are taken by humans.

Current national practice

- Sweden presented its sensor-fused munitions as a weapon system that has autonomous functions. Human control is exercised when analysing target areas’ character, and the warhead is then autonomously directed against the target. Once the autonomous function is launched, the mission cannot be stopped. Sweden highlighted that this weapon system passed the Article 36 review.
- The US presented about its naval mines counter-measures programme as an example of how autonomous systems can mitigate highly complicated threats.
- Sweden raised the question if there is a risk of electronic warfare disrupting communication channels, and if the system could also be used

continued on next page



News in brief, continued

in offensive situations. The US replied that the system is not designed to be offensive, it does have similar technology but would require more research.

- The University of North Carolina challenged the notion that the absence of civilians makes the use of AWS more permissible. It pointed at the major military powers having forces at the sea that could risk rapid escalation of the use of AWS in these environments.
- Russia explained that both international and national laws are taken into account when developing new weapons. The process further includes

an environmental impact assessment and a test cycle that is comprehensive, includes testing in “real life” conditions, and ensures a situation where military and political leadership is fully aware of potential consequences of weapons under consideration.

Gender

- Ireland noted that the degree of autonomy in weapons may be affected by numerous factors, including algorithmic foundations. It stated that there is the possibility of perpetuating and amplifying social bias, including gender bias, which therefore requires careful consideration. •

Photo: Clare Conboy/Campaign to Stop Killer Robots





PROJECT
PLOUGHSHARES



Reaching Critical Will



What's Gender Got to Do With It? Feminist Approaches to Disarmament and LAWS

Meeting of the Convention on Certain Conventional Weapons
Group of Governmental Experts on Lethal Autonomous Weapons Systems

SIDE EVENT BRIEFING

Thursday, 30 August 2018
13:15 - 14:45

Conference Room XXIV
United Nations Geneva

Light refreshments will be provided

Is there a role for gender analysis in the Convention on Certain Conventional Weapons (CCW)?

This side event says yes. This discussion will contribute to issues considered at the second 2018 meeting of the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems (LAWS). It will chart a path towards stronger consideration of gender in the context of LAWS and in the CCW's work more broadly. Speakers will address gender diversity and equality in disarmament negotiations and discussions; gender norms and the development and use of weapons; gendered impacts of existing weapon systems; how the issue of gender has been taken up in other disarmament forums; and the importance of feminist foreign policy approaches in relation to disarmament and arms control.

Moderator:

Ambassador Rosemary McCarney, Ambassador and Permanent Representative of Canada to the United Nations and the Conference on Disarmament in Geneva

Speakers:

- Ms. Ray Acheson, **Reaching Critical Will**
- Ms. Erin Hunt, **Mines Action Canada**
- Dr. Elke Schwarz, **International Committee for Robot Arms Control**
- Ms. Sylvie Jacqueline Ndongmo, **WILPF Cameroon**
- Dr. Branka Marijan, **Project Ploughshares**

For more information about this briefing event, please contact Erin Hunt, Mines Action Canada, Mobile +1 613 302-3088 or erin@minesactioncanada.org.

Sponsored by the Government of Canada

Canada