

VOL.9 NO.6

23 September 2021



IN THIS ISSUE:

- 1 Editorial: Constraining militarism and controlling weapons
- 3 Austria's conference on autonomous weapons offers bold support for a ban
- 12 The development of killer robots is a threat to human rights in Africa



Reaching Critical Will

www.reachingcriticalwill.org



www.wilpf.org

EDITORIAL: CONSTRAINING MILITARISM AND CONTROLLING WEAPONS

Ray Acheson | Women's International League for Peace and Freedom

On 15 September, UN High Commissioner for Human Rights Michelle Bachelet **called** for certain applications of artificial intelligence (AI) to be prohibited and for others to be subject to a moratorium, based on the risks they pose to human rights. This followed the publication by her office of a **report** that analyses how AI—including profiling, automated decision-making and other machine-learning technologies—affects people's right to privacy and other rights, including the rights to health, education, freedom of movement, freedom of peaceful assembly and association, and freedom of expression.

"We cannot afford to continue playing catch-up regarding AI—allowing its use with limited or no boundaries or oversight, and dealing with the almost inevitable human rights consequences after the fact," argued Bachelet. "The power of AI to serve people is undeniable, but so is AI's ability to feed human rights violations at an enormous scale with virtually no visibility. Action is needed now to put human rights guardrails on the use of AI, for the good of all of us."

This is a crucial and important step forward to protecting human rights from the grave risks posed by AI. The same measure of urgency and action is imperative for autonomous weapon systems (AWS). The recognition of how AI can undermine human rights must also be considered regarding the right to life and dignity.

At the **virtual conference** hosted by Austria on 15–16 September, New Zealand's Minister for Disarmament and Arms Control Phil Twyford **noted**, "If we don't create new binding rules on autonomous weapons, the risk is that we'll see a more dangerous and unstable world; that we'll undermine international humanitarian law." This precautionary approach matches that urged by Bachelet for AI. It is absolutely essential for ensuring that meaningful human control is maintained over weapons and the use of force. As

Austrian Foreign Minister Alexander Schallenberg and Gilles Carbonnier, Vice President of the International Committee of the Red Cross, argued at the Austrian conference, "death by algorithm" is unacceptable and must be prevented.

The urgency of action

The dangers posed by AWS for human rights, international humanitarian law, and peace and security have been discussed for years at the Convention on Certain Conventional Weapons (CCW), the Human Rights Council, and the UN General Assembly First Committee. We can already see the harms caused by other AI and machine-learning systems: bias and discrimination leading to incarceration, harassment, surveillance, denial of goods and services, refusal of entry at borders, misidentification of individuals, etc., as well as failures in software and sensors that lead to malfunctioning systems and misinterpretation of data. Weaponising this technology will inevitably lead to grave violations of human rights, human and environmental harm, and global violence.

The time to act against AWS is now, before they are developed and used. Autonomy in weapon systems is already increasing. We know that the handful of governments preventing progress on this issue at the CCW already have **advanced capacities** in designing, building, and even **using** these systems. But it is not yet too late to establish binding prohibitions and regulations against AWS. As Minister Twyford pointed out, not acting against these weapons now risks the development of a new political economy, in which these technologies are linked to powerful economic and military interests of violent countries. This will make these governments, and their military-industrial-tech complexes, even more resistant to the establishment of global rules and norms.

As journalist William M. Arkin **writes** in *The Generals Have No Clothes*, war persists because

the apparatus of people, bases, and weapons has grown so vast that it can no longer be understood, much less controlled; it has become “a gigantic physical superstructure” that “sustains endless warfare.” The perpetual war, Arkin contends, is “a physical machine, and a larger truth, more powerful than whoever is president,” and the result has been “hidden and unintended consequences, provoking the other side, creating crisis, constraining change.”

Standing up for humanity by constraining militarism

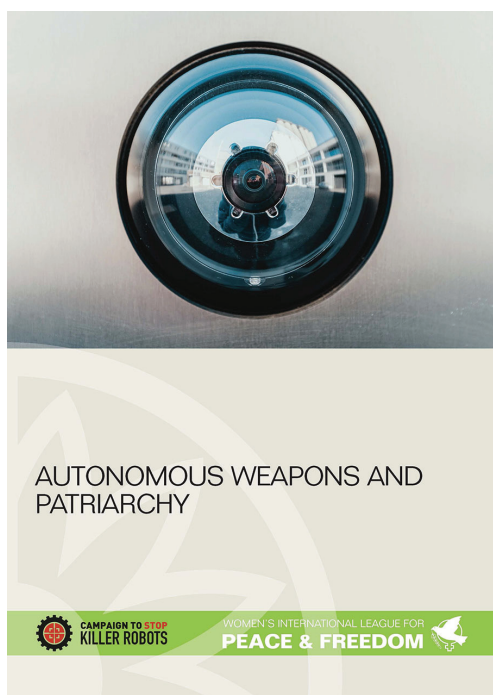
Even prior to the advent of AWS, we are already facing a crisis of control when it comes to restrictions and restraints on weapons, the use of force, and military aggression. As Dr. Patricia Lewis noted in her remarks at Austria’s conference, certain states feel entitled to shirk their responsibilities and commitments and to act however they feel suits their own interests,

rather than looking out for collective security and the good of humanity. This has posed a serious challenge to international law and norms that must be confronted. It is up to the so-called international community, if it wants to be an actual community, to stand up for the principles of humanity and the dictates of public conscience.

When it comes to the dangers posed by AWS, we cannot say we didn’t know. We cannot say, as many scientists did about the atomic bomb, that they didn’t realise the horrors they were unleashing upon the world until it was too late. Tech workers, scientists, roboticists, activists, and human rights organisations have warned for years now that this technology will harm humanity in a multitude of ways. And given what we know already about the political economy of war, it is never too early to act against investments in new weapons. It’s time to take a stand on what Minister Twyford described as “one of the great moral issues of our time,” and develop new legally binding rules prohibiting AWS.

WILPF PUBLICATIONS ON AUTONOMOUS WEAPONS

DOWNLOAD FROM WWW.REACHINGCRITICALWILL.ORG



AUSTRIA'S CONFERENCE ON AUTONOMOUS WEAPONS OFFERS BOLD SUPPORT FOR A BAN

Ray Acheson and Allison Pytlak | Women's International League for Peace and Freedom

From 15–16 September 2021, the Austrian Ministry for European and International Affairs hosted a virtual conference called “Safeguarding Human Control over Autonomous Weapon Systems,” with a view to contributing to the work of the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts (GGE).

The conference focused on the legal, ethical, and security challenges posed by increasing levels of autonomy in weapon systems. Panelists spoke to how humans can maintain meaningful control over these weapon systems, so as to preserve human dignity and to prevent unintended legal, ethical, and security consequences. They also addressed the implications of new and emerging technologies for armed conflict and on the use of force, which requires urgent political attention.

Elevating the perspectives of tech workers, as well as marginalised groups that will be most impacted by these weapon systems, panelists strongly advocated for new international rules and norms to prevent the development, possession, and use of autonomous weapon systems (AWS), and to maintain meaningful human control (MHC) over the use of force. Overall, the risks and dangers posed by increasing autonomy in weapon systems cannot be denied or ignored. Recognising how other developments in artificial intelligence (AI) and remote war technologies are already undermining human rights is crucial to taking the urgent action needed to prevent the weaponisation of related technologies and the “death by algorithm” scenario this presents.

High-level opening panel: Global AI governance—the necessity and challenges of regulating emerging technologies

The high-level opening panel set a highly positive and encouraging tone for the conference. The speakers all emphasised similar themes: the need for strong governmental action in ensuring human

control over weapon systems; the diverse risks presented by AI, autonomous weapons, and other related technologies; and the necessity of involving civil society and other non-governmental actors in discussions and action on these issues. There were multiple strong calls for the development of a legally binding instrument on AWS and demonstration of political leadership to move in that direction.

The panel took the form of a moderated discussion managed by Hannelore Veit, in which opening remarks were followed by a series of follow-up questions. Two video messages were delivered at the conclusion.

Opening remarks

H.E. Alexander Schallenberg, Federal Minister for European and International Affairs of Austria, outlined in his opening remarks why Austria is organising this conference and why AWS are a concern to the international community. He explained that Austria has been on the “vanguard” of many disarmament, non-proliferation, and arms control issues, and now we are facing a new challenge: the new frontier of AI. He underscored how technology is racing ahead of governance. While technology can be beneficial and an enabler of rights, it can also be misused in multiple ways. Schallenberg explained that AWS, in which an algorithm makes life or death decisions, is unacceptable for reasons of ethics, morality, and law. A core question is: how to ensure that humanitarian law, human rights, fundamental freedoms, and human dignity will be safeguarded into the future? From his perspective, an important point is to not be technophobic and to get industry involved in efforts toward a legal norm to ensure human control. However, he acknowledged that human control cannot be left to self-regulation within industry. He hoped that this conference would be a first step in a longer process leading toward a treaty.

H.E. Izumi Nakamitsu, the United Nations Under-Secretary-General and High Representative for Disarmament Affairs, reinforced several of Schallenberg's points around the positive potential of technology, including with aiding in disarmament initiatives. Speaking also of information and communications technologies (ICTs) as well as AWS, Nakamitsu identified key challenges facing the international community: some technologies are being used to carry out harmful or malicious acts that occur beneath the threshold of armed attack; ICTs and AI can lead to instability because they introduce unpredictability during existing times of tension; attributing responsibility for use or attacks is complicated; and non-state actors can more easily acquire and exploit new and emerging technologies. Nakamitsu referenced statements from the UN Secretary-General calling for a ban on AWS, noting a call for limitations is included in the UNSG's recently released Common Agenda, as well as in the 2018 Securing Our Common Future: An Agenda for Disarmament.

Gilles Carbonnier, Vice President of the International Committee of the Red Cross (ICRC), was asked to respond to what the ICRC sees as potential challenges to international humanitarian law (IHL) in this context. He outlined that since its inception, the ICRC has focused on the humanitarian impacts of new weapons. Advances in military technology do not occur in a legal vacuum, he observed, but while IHL applies to new weapons, there is need to clarify how, and maybe develop new rules. Carbonnier expressed that legally binding rules are needed on AWS because of the unique challenges they present. In May, the ICRC put forward a series of recommendations to states on the form and substance of future limitations and prohibitions for AWS that respond to their unpredictability, indiscriminate effects, targeting issues, and the environment in which they may be used.

Honourable Phil Twyford, Minister for Disarmament and Arms Control of New Zealand, was asked about his unique role as one of very few (if any) disarmament ministers in the world and why this is so important for New Zealand. It is a moral and ethical issue, he responded, outlining a long history of antinuclear and pro-disarmament campaigning by the people of New Zealand. In his

view, AWS are emerging as a great moral issue of our time. He reflected that as technology is transforming our world in positive ways, there is a dark side to it which many in the tech industry and civil society have been warning about for a while. Twyford explained that New Zealand is undertaking a comprehensive policy process with its military and technology sectors, as well with consideration for economic development interests, to set out an "all of government" policy that will be a platform from which to push for new legally binding rules.

Moderated discussion

Veit posed various questions to panellists around the following themes: the importance of political leadership and of multilateralism; the role of the UN; prohibition versus regulation of autonomous weapons; and the price of inaction.

Below are highlights from the panelists' responses:

- Schallenberg said that AWS are part of a larger issue of morality and ethics in the midst of a technological revolution. Humans must be at the centre of this revolution or we risk "digital dehumanisation," a point supported by Nakamitsu.
- Schallenberg and Twyford referenced past positive experiences working with civil society in the processes to ban landmines and cluster munitions. Twyford referred to the Mine Ban Treaty as an "extraordinary act of moral leadership."
- Successful "public private partnerships" were stressed by Schallenberg in which everyone is at the table. He clarified that the tech industry is important to this conversation in order to ensure that policymakers understand what is being discussed.
- Twyford likened the threat of autonomous weapons to the climate emergency and sustainable development, because they require a multilateral approach in which all are mutually accountable.
- Twyford noted that there is a tension between the interests of humanity and the interests of the so-called great military powers who are already developing AWS because of the

advantages it will bring them, and which they do not want to give up. He also warned about the risks of a “new political economy” in which new technology is linked to powerful economic and military interests of these “powerful” states. Such an economy would be resistant to global civil society and governments coming together to create new rules later on, he argued.

- Nakamitsu distinguished between threats posed by AWS and emerging technologies, and more traditional weapons issues like nuclear disarmament, in which there are a limited number of possessors. She feels it is unhelpful to think about an AI arms race in the way we understand the Cold War arms race, because the fundamental technologies here are being developed in the private sector and not by national militaries and are also more easily accessible by other actors. Similarly, Schallenburg distinguished between the traditional arms industry and arms dealers, and the technology companies of today.
- Nakamitsu identified two pathways for action in the UN: exchange of national practices regarding IHL, and seeking to agree to limits and effective regulations on certain capabilities. Both can be complemented by cooperation relating to new weapons reviews. She also urged “redoubling and tripling” efforts at the Convention on Certain Conventional Weapons (CCW) and views the upcoming CCW Review Conference as an “inflection point” for deliberations on this issue. Twyford said he would like to give the CCW every possible opportunity for progress.
- Nakamitsu supports the CCW Group of Governmental Experts (GGE) Chair on his “forward looking” paper as a basis for agreement in the GGE and asserted that any path chosen can only be effective if it is supported and implemented by a broad majority of states.
- Carbonnier described in more detail how the ICRC arrived at its position of supporting prohibitions on some kinds of AWS, and limitations on others. “Targeting humans amounts to death by algorithm,” he observed, which is why the ICRC recommends a prohibition on AWS that target humans. He

also explained that the complexities of modern conflict environments, often urban, present greater risks of violations. He views this moment as an opportunity to “draw a clear normative line.”

- Carbonnier also observed that risks are increasing rapidly and there is an urgent need for states to take collective action and move toward the adoption of legally binding rules.

Video messages

A video message from Belgium reiterated that new technological developments in the area of AWS raise serious questions about the risks posed to respect for IHL on future battlefields. This risk touches on core values, and human control must be maintained. This is why Belgium insisted on the development of a guiding principle on this topic at the GGE. It has now joined with other countries to draft a working paper on the GGE Guiding Principles that require more development. Belgium is now chairing the GGE on LAWS and is encouraged to see growing convergence among states for a dual track approach that could be translated into a normative framework. In this framework, some AWS would be banned while other systems would be regulated to ensure conformity with IHL. There is also a role for national legal reviews of new weapons. The message concluded by stating that the CCW is the right forum to discuss AWS, with all stakeholders around the table.

A video message from Gabriela Ramos, Assistant Director-General for the Social and Human Sciences, of the United Nations Educational, Scientific and Cultural Organisation (UNESCO) emphasised that AI plays a role in billions of people’s lives. While AI has extraordinary value and potential for development, this is not the case with AWS. UNESCO member states have recently agreed on a draft text for a series of recommendations on AI and ethics. While this is mainly focused on civilian uses, Ramos stated that AI systems cannot replace human responsibility and accountability.

Expert Panel 1: International law and safeguarding human control

The four experts in this panel focused mainly on international humanitarian law (IHL) and outlined its applicability, as well as limitations, for maintaining human control in the area of emerging technologies, with an emphasis on autonomous weapon systems (AWS).

Dr. Elisabeth Hoffberger-Pippan, of the International Panel on the Regulation of Autonomous Weapons (iPRAW) and the German Institute for International and Security Affairs, began by explaining that IHL does not explicitly apply to AWS or require the maintenance of human control over the use of force. However, human control in abstracto deviates from various IHL rules by implication. In particular, she referenced the IHL principles of distinction and proportionality; command responsibility; and the Marten's Clause. Hoffberger-Pippan went into greater depth in explaining the Marten's Clause and how it can be used to direct obligations onto parties to a conflict, as well as to interpret and apply other norms of IHL. She cited the 1996 Advisory Opinion of the International Court of Justice on nuclear weapons, including a Dissenting Opinion, as an example of where the Marten's Clause was invoked both as a means to address the rapid evolution of military technology, and to express that, human rights must be taken into account when interpreting the Marten's Clause. Hoffberger-Pippan concluded by stating that human control in abstracto is not only politically desired, but legally required. A normative and operational framework is needed to regulate in more detail the operationalisation of human control, in concreto.

Dr. Vincent Boulanin of the Stockholm International Peace Research Institute (SIPRI) reinforced many of the points made by Dr. Hoffberger-Pippan. His remarks drew from a June 2021 report, *Autonomous Weapon Systems and International Humanitarian Law: Identifying Limits and the Required Type and Degree of Human-Machine Interaction*. Boulanin outlined that there are limits under IHL on the development and use of AWS, and that an AWS would be considered unlawful depending on some of its characteristics, such as if it is indiscriminate by nature, or would cause

superfluous injury or unnecessary suffering. He explained that respect for IHL presupposes the fulfillment of three conditions: the ability to foresee; to administer and trace the operation; and to trace the operation, performance, and effect of AWS. Boulanin ended by raising a series of questions which, in his view, require further exploration by states and the international community. The report recommends action in three areas: 1) clarifying what states must do to respect and ensure respect for IHL in the use and development of AWS; 2) deepening discussion on what respecting and ensuring IHL means for AWS; and 3) sharing views and experiences of practical measures that could enhance respect for IHL.

Laurent Gisel of the International Committee of the Red Cross (ICRC) spoke to some of the specific characteristics of AWS that present compliance challenges for IHL and underscored the importance of ethical considerations. In this context, he outlined the ICRC position that AWS that are designed or used in ways that their effects cannot be understood, predicted, or explained should be prohibited. For those where the effects can be reasonably predicted, then there is a need for regulation to ensure human involvement and control. Limits and requirements on use should, in the ICRC's view, include limitations around targeting; duration of an attack; contexts where no civilians are present; and imposing requirements for human-machine interaction. He encouraged states to clarify their discussions further and urged that any development or use of AWS be in full compliance with IHL. Gisel noted that this is the role of a national legal review, but also stated that if states agree that some AWS should be prohibited and others limited, then this should be enshrined through new and legally binding rules. New rules would offer clarity, Gisel explained, citing blinding laser weapons as an example.

Richard Moyes of Article 36 reinforced points made in the opening panel that there have been significant developments over the last year in terms of shaping an international response to this issue. He noted that while not all participants in the GGE are yet in agreement, they have capacity to understand each better now in a conversation about AWS, which he sees as a basis for moving toward a legal instrument. Moyes then raised

several points relating to how states and other actors engaged in the GGE on AWS can orientate toward law. First, he observed that ongoing development and further codification is needed, as evidenced by the development of additional protocols to the CCW overtime, but that simply “repeating” existing law isn’t sufficient. Second, there are already legal developments within wider society that relate to the relationship between automation and people. Moyes gave the General Data Protection Regulation (GDPR) as an example of where a significant number of states have recognised that technological developments and the relationship to people has required a specific legal response. Third, Moyes encouraged states to see that many of the existing legal proposals are meant to be protective of the existing legal regime and framework. Finally, he urged a better understanding of the context and technology being used, as a basis for evaluating legal compliance. Moyes ended by reiterating that civil society looks forward to working with states and others to develop practical and flexible instruments to preserve the role of humans as appliers of the law, but that also draw a clear line against machines killing people.

Panel moderator **George-Wilhelm Gallhoffer** of Austria then opened for questions from the audience.

A first question noted there is not an agreed definition of AWS within the GGE and asked how much of a challenge that is for legal discussions. Moyes said that a definition is not vital and the focus should be on adopting a process of work that moves into a negotiating mode. Hoffberger-Pippan agreed and noted that it’s important to bear in mind that AWS are not a single platform but are multiple systems with other capabilities involved, which complicates developing a definition. She suggested that when speaking about definitions it is important to remain open to future technological developments, and to focus instead on the role humans play. Boulanin agreed and highlighted that working definitions, such as that provided by the ICRC, are sufficient for now.

A second question asked how the ICRC and SIPRI understand and define “unpredictability”. Gisell reinforced that unpredictable weapons are already

prohibited because it cannot be guaranteed that their use can be limited by IHL.

A third question asked what panelists would say to those who disregard the Marten’s Clause as irrelevant because it’s not legally binding. Hoffberger-Pippan responded to say that it is enshrined in various IHL documents, which are binding. She clarified that in her presentation, she had meant that the Marten’s Clause might not be used to confer direct obligations on states.

The final question asked how to define “human control”. Moyes acknowledged this as a good question and one that delegates have posed in the GGE. He feels it is important to not try to define it first and then work out what the legal rules are. Rather, Moyes said it is more important to bring together a group of states who are prepared to take leadership and develop a legal instrument which will embody a collective orientation to human control.

Expert Panel 2: Ethical Dimension and Human Dignity

Moral and ethical objections to the impending risks of “death by algorithm” described during the previous day of the conference were the focus of the ethics panel. Featuring perspectives from the military, tech workers, and regional and international bodies, the speakers all highlighted the connections between the technical limitations of relevant technologies and the challenges these pose for ethical and moral principles, as well as for human rights and dignity. The need to prohibit and regulate autonomous weapon systems (AWS) now, before these technologies are developed and used, was a clear message from the panel.

Colonel (General Staff) Dr. Markus Reisner, PhD, Austrian Armed Forces Military Academy, grounded his remarks in the context of existing weapon systems, describing some of the weapons currently in use that push the current boundaries of autonomy. He explained that human control over weapons is exercised via a network of structures in cyber space, arguing that the issue of AWS is not just a matter of robotics or hardware but also about computer networks and operations. We can already see how militaries are trying to advance

the speed of decision making and the operation of weapon systems in warfighting, through sensors, swarms, and targeting tools. Yet as this work goes on, questions about how humans should be in the loop, how they should be connected to the weapon and to the system overall, have gone unanswered. Robotics software, argued Reisner, operates on a purely mechanical basis, not an ethical basis. An AWS is a fighting machine—in simulations, an AI pilot “beat” a human F-16 pilot, because the machine was more aggressive than the human. Machines can only be stopped by technical failure, or by a command from the outside—if that is possible. Thus, software must be designed so that humans have possibility to intervene at any time in a machine’s operations. While humans can violate international law, he noted, AI’s violations could have even more devastating impacts.

Wanda Muñoz, Victim Assistance and Disarmament Consultant Member of Human Security Network in Latin America and the Caribbean (SEHLAC) and Global Partnership on Artificial Intelligence, highlighted the growing and clear public conscience against AWS. Since 2013, she explained, a growing number of states and international organisations have called for a legally binding instrument on AWS. Public surveys have found increasing opposition to AWS. More than 4500 AI and robotics scientists have signed a letter opposing the development of AWS. Comments from the UN Secretary-General, as well as the common agenda he released recently, support action against AWS. National and regional bodies are developing standards for ethics and artificial intelligence (AI). Muñoz highlighted the recent UNESCO recommendations, which say that life and death decisions must not be given to AI systems and raise fundamental ethical concerns about bias in AI technologies. She also noted the report released by the UN High Commissioner for Human Rights and the call from Michelle Bachelet to prohibit some AI systems that will violate human rights.

These and other ethical commitments undertaken by various groups and countries should be taken as a minimum ethical standard for discussions about AWS, Muñoz argued, otherwise we will develop an incoherent international framework. Furthermore, she noted, ethics is not a single monolithic guideline—it evolves to incorporate

different perspectives, including from marginalised groups. When it comes to AWS, we should prioritise the ethical perspectives of those affected by conflict and the humanitarian organisations that respond to conflict every day—these perspectives will be different from those of the governments that produce weapons. A feminist approach brings to the fore the lived experience of affected people, who should be considered “experts” alongside those traditionally understood as such. With this in mind, it is not premature to call for a ban on AWS, given all of the knowledge of lived experience we have with other methods of remote warfare, which AWS will make much worse. Thus, the international community should:

1. Negotiate a legally binding instrument that prohibits AWS that function without meaningful human control (MHC) and that target humans, and regulate other AWS. If the Convention on Certain Conventional Weapons (CCW) fails to agree to this, like-minded countries should initiate a process outside the CCW. Continuing discussions without action only benefits those that want to produce these weapons at the expense of those who will be victims.
2. Ensure human rights are at the forefront of efforts against AWS.
3. Build upon and ensure coherence with commitments countries have already undertaken on AI and ethics.
4. Focus on AWS, not Lethal AWS, as lethality is not a criterion under international humanitarian law (IHL).
5. Ensure measures for accountability and responsibility for consequences of any AWS use.

Liz O’Sullivan, CEO of Parity, offered her own lived experience with AI systems. She described her work at Clarify, a tech start-up that initially eschewed military contracts but eventually became involved in Project Maven, an infamous Pentagon effort to “increase precision” in drone strikes through AI. This work caused harm to many of her colleagues working on the project, as well as the wider company, introducing surveillance, spyware, and extreme security measures to the firm. Many

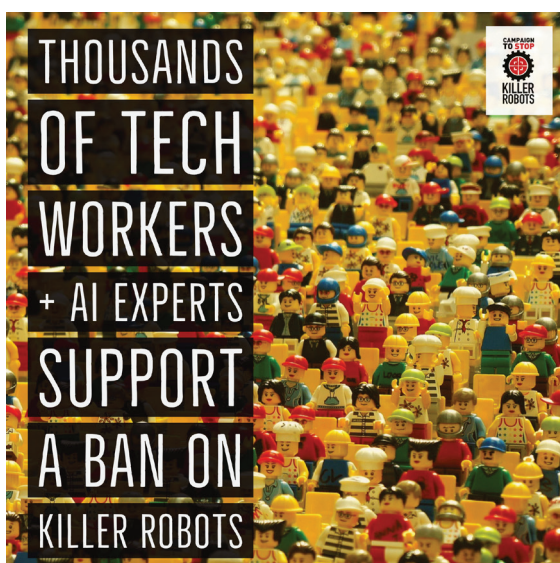
people asked to be transferred once they found out what they were working on, but then realised that any data the company collected on other projects could end up with the military. The broader public has no idea how artefacts of their online lives may become embroiled in military pursuits and weapon development. In her own work, O'Sullivan experienced first-hand the problems with bias in AI, in particular on marginalised people. She saw the various ways this technology would fail and the unpredictability and inexplicability of that failure. Once she realised it was never possible to fully debias this technology, O'Sullivan she organised internally a resistance to Clarify's involvement in military contracts. When the company refused to commit to not building weapons, she left her job. When it comes to such employee activism, she noted, what is visible is only the tip of the iceberg. Among the tech community, there is nearly unanimous solidarity in understanding this tech will fail.

Charlotte Stix, Technology Policy Expert with specialisation in AI governance and PhD Student Fellow at Cambridge University, talked about the European Union (EU) strategy on AI, published in April 2018. The group of experts developing the strategy issued ethical guidelines, policy and investment recommendations, and assessments for "trustworthy" AI—which it defines as lawful, ethical, and robust AI. The strategy sets out a number of principles, including respect for human autonomy, prevention of harm, fairness, and explicability, which feed into the requirements needed for an AI system to be considered

trustworthy. These requirements include: human agency and oversight; technical robustness; safety; privacy and data governance; transparency; diversity, non-discrimination, and fairness; societal and environmental well-being (including for future generations; and accountability. Stix also outlined a number of lessons learned from the process to develop the AI strategy, including, among other things, the necessity of a participatory multistakeholder process that includes multiple iterations, diverse expectations, and backgrounds; and a broad lifecycle approach that asks, what does it do to put this AI system into the world? As a concluding remark, Stix emphasised that it's important not to just engage in "ethics washing" for questionable technology but to actually assess its real-world impacts.

Irakli Khodeli, Programme Specialist for Ethics of Science and Technology, Social and Human Science sector, United Nations Educational, Scientific and Cultural Organization (UNESCO), explained that UNESCO works not just for the absence of war but to build active structures for peace, cooperation, and education. It has used this mandate to promote ethics of science and technology since its foundation. Khodeli agreed with Muñoz about the need for a global normative framework to harmonise the growing number of national and regional AI strategies. For regulating or outlawing military uses of AI, he argued, we need a global normative framework to capture civilian use of AI—that's how we build a general global background understanding among people to push back against military use and support efforts for regulation.

He suggested that UNESCO offers an optimal platform for establishing and promoting such a framework for ethics of AI, given its global reach. Through multiple consultations with diverse representation over two years, UNESCO worked to elaborate the first global standard-setting instrument on the ethics of AI. Among other things, some of the principles for such an instrument include proportionality and do no harm; human oversight and determination; transparency and explainability; and responsibility and accountability. The choice to use AI systems—and the decision about what systems to use—must be justifiable. It must be appropriate and proportionate, must not infringe upon the values set out in the document,



including the protection of human rights, should not have irreversible effects, and, most importantly for AWS, must not involve life and death decisions.

The Q&A, moderated by **George-Wilhelm Gallhoffer** of Austria, addressed questions about bias and how civilian principles can help shape military use.

In relation to bias, O'Sullivan reiterated that there is no such thing as a fully unbiased model, especially in relation to warfare. The definition of a battlefield is that it will be unpredictable, she pointed out. Those who think data bias is a limited factor are wrong. Tech workers are trying to improve AI systems continuously, but it is the most marginalised people that will be impacted by these systems.

In relation to influencing military considerations of AWS, Stix suggested that broadening the dialogue about AWS in the public and media is an important step towards shaping the world's understandings of what can go wrong, or what can't simply be solved, when it comes to AWS, and that will help shape the military dialogue. Reisner agreed, stressing the importance of having these conversations now. With nuclear weapons, he argued, it took people a long time to realise what had been created and the damage it would do to our world. Before we create fully AWS, he urged, we should start now to think about what will come in the future. Democratic processes should come up with red lines to limit or prevent the development and use of this technology.

Expert Panel 3: Maintaining Human Control and International Peace and Security

"Slow down," was a key message from the panel on MHC and international peace and security. The desire to speed up decision making and responses to attacks or uses of force is not in anyone's best interest, the panelists argued. The technical capacity to turn warfighting over to machines is already here, but we must stop this breakneck race into danger and instead build international rules and norms against delegating life and death decisions to machines. The tech community, as the panel also discussed, stands behind this endeavour to prohibit and regulate AWS to protect human life, rights, and dignity and to avoid escalation of conflict and erosion of barriers to war and violence.

Dr. Frank Sauer, Bundeswehr University Munich, Senior Research Fellow and member of the International Committee for Robot Arms Control (ICRAC), described how networked weapon systems already operate fairly autonomously, from surveillance to command and control to targeting. In such systems, humans are only responsible for greenlighting an attack and are not involved in other aspects of the "kill chain"—which has operated at lightning speed. The capability, then to operate weapons fully autonomously is here now; having a human greenlight attacks is only a matter of doctrine, not technical capacity. Sauer then put this in the context of the Cuban missile crisis. While military officials were advising US President Kennedy to launch air strikes, he didn't want to be rushed. He stalled for time. This demonstrates the importance of allowing time to de-escalate. There is extreme danger in accelerating decision-making time beyond a point where human cognition can keep up.

There are strategic risks due to compression of time and the possibility for algorithmic accidents and inadvertent escalation of conflicts. Humans make mistakes, Sauer noted, but humans don't all make the same mistake at the same time at lightning speed—this is what a machine does.

Thus, Sauer argued, rules must govern who or what makes decisions about what, when, and where in relation to use of force. Without this, the risks of weapon autonomy will outweigh any possible benefits. No matter how much we train AWS, it won't fully understand consequences of its actions. Human insight and consideration are imperative to understand consequences and mitigate collective risks. These kinds of political implications of AWS are just as important as the legal issues, Sauer argued, and it's up to the militarised states pursuing this technology to address this.

Dr. Patricia Lewis, Chatham House, Research Director International Security agreed with the need to slow down, not speed up, when it comes to decisions around the use of force. Meaningful human control (MHC) means we have to stop, engage, understand, and decide. We know the risks, she said. We are already living in a situation where nuclear weapons are on alert status. We are aware

of how close we've come to inadvertent nuclear war many times. But for the chance decisions of individuals, we would have been tricked into nuclear war. We must use this, and also the crisis of COVID-19, to think through what could go wrong with AWS—to think about the importance of decision-making, of disparities and differences between countries, and more.

Lewis also noted how highly reliant we are already on AI, especially in cyber conflict. Cyber security attacks, whether they are criminal attacks, espionage, or attacks against critical infrastructure such as industrial control systems or military systems, are already taking place on a daily basis. And it is machine-on-machine responses. We do not have MHC in our cyber networks, other than in relation to the decisions made on how to deploy the AI. Right now, this acts as a constraint but once machines are making decisions about how to respond, and to respond with attacks, we will be in a different state of play, she warned. This is not the future—this is now. And international regulation is the only way to deal with it. But at the same time, the value of international law is being challenged. For some countries, the memory of the significance of control has been lost—some think they have the right and entitlement to move beyond these controls and constraints. This is a huge problem for those trying to use international law to deal with dangerous developments in weapon systems. We need a normative approach, but this is difficult when norms are being eroded and undermined. That said, Lewis argued, we are already drifting into a situation where we are putting human beings into impossible situations, where they have no choice but to either trust or not trust machines. We must act now to prevent this, or regret it later.

Emilia Javorsky MD, MPH, The Future of Life Institute (FLI), highlighted the growing support from tech workers and scientists to prevent the situation described by Lewis. Thousands have signed the FLI pledge abdicating from work on AWS and calling on governments to take action against developing AWS. In surveys, FLI has found that 74 per cent of researchers are opposed to AWS—to the point where about one-third would resign from their jobs if forced work on such systems, and one-quarter would speak publicly to the media to blow the whistle on their companies. Javorsky

noted that this is largely because tech workers believe AI has the potential to transform society for the better, but that to do so, there must be a red line that decisions to take human life must never be delegated to AI. Chemists don't want to build chemical weapons, she argued, and tech workers don't want to build AWS. Tech workers have also identified several key challenges with AWS, including their proliferation risks, the erosion of geopolitical stability, lowering the threshold for war, acquisition by non-state actors, and the escalation of the risk of unintentional conflict and reduction of the ability to de-escalate conflict.

The Q&A, moderated by George-Wilhelm Gallhoffer of Austria, addressed questions about international law and norms, decision making, and arms racing.

Asked to elaborate on her comments about erosion of the rules-based order and international norms, Lewis noted that after the horror of World War II, the international community created the UN and tried to constrain itself. But over time, some have forgotten this history, or think it does not apply to them. We need to remind everyone of what happens when you don't put on constraints or control behaviour, when you think you can ride above that and win. We need to build up restrictions before another horrific event, she argued. The importance of law is clear when we look at what's happening in cyber space. All governments have agreed that international law applies in cyber space, yet some are trying to create a special set of laws and acting as if existing law doesn't apply to them.

Sauer noted that in the context of AWS, there is growing convergence on the need for new law. Some specific applications, such as targeting human beings or weapons that operate without MHC, will need to be prohibited, while other applications will need to be regulate. He stressed again that decision-making in social situations requires human thinking. The human brain is a fantastic cognitive apparatus, he argued, we shouldn't remove it from the equation. Instead of barrelling headlong into an arms race, it's time to pump the breaks and think rationally about where we want to use this tech and in which way, and where we want to draw specific lines.

THE DEVELOPMENT OF KILLER ROBOTS IS A THREAT TO HUMAN RIGHTS IN AFRICA

Guy Blaise Feugap | WILPF Cameroon

Since April 2018, WILPF Sections and Groups have been campaigning against killer robots in Africa, with WILPF Cameroon participating in CCW meetings in Geneva and other Groups and Sections joining a workshop on autonomous weapons at the WILPF International Congress in Accra. The challenge from the beginning has been to convince more African countries to endorse a preventive ban on autonomous weapon systems (AWS). While the Campaign to Stop Killer Robots has sufficient resonance with governmental actors and regional institutions, concrete actions to officially support the campaign remain few and far between. This is why campaigners in Africa are working to further expose the dangers and risks associated with the development and use of killer robots.

A **webinar** organised by WILPF in Africa on 20 September 2021 focused on the issue of human rights associated with emerging technologies and the case of killer robots, with a look at the work of the African Commission on Human and Peoples' Rights. This year, the Commission, at its 31st Extraordinary Session held virtually from 19 to 25 February, adopted Resolution 473 on the need for a study on human and peoples' rights and artificial intelligence (AI), robotics, and other new and emerging technologies in Africa. This concern of the African Union brought together on a call campaigners in Africa, experts in AI, lawyers, security experts, and media representatives, convinced that the possible horrors of AWS cannot be allowed in a continent already scarred by armed conflict.

Killer robots are a violation of human rights, yet robots were originally intended to assist human beings. According to Jacques EONE, CEO of Sparte Robotics, humankind has reached the point where robots reduce human tasks daily. As science evolves, the work is now left to the robots and their use will no longer be limited. Thus, in the arms industry, this technology is being innovated on a daily basis and allows the manufacture of AWS. But

this technology is meeting with ethical and legal resistance. The technology tramples on the robotic laws laid down by Assimov, who said that robots should not harm human beings.

Resolution 473 of the African Commission on Human and Peoples' Rights on the need for a study on Human and Peoples' Rights and Artificial Intelligence, Robotics and other new and emerging technologies in Africa, provides a basis and perspective on the urgency of such a study. Dr. Thomson Chengeta, who is a researcher at the European Council on Drone Violence and the Ethics of Artificial Intelligence, explained why this is important. Indeed, it does not conceive of robotics and other new and emerging technologies that are not "compatible with the rights and duties in the African Charter and other regional and international human rights instruments". Furthermore, the African commission "appeals to State Parties to ensure that all AI technologies, robotics and other new and emerging technologies which have far reaching consequences for humans must remain under meaningful human control in order to ensure that the threat that they pose to fundamental human rights is averted."

As the webinar of 20 September 2021 was a continuation of the actions of the Campaign to Stop Killer Robots in Africa, it allowed some of the invited participants to realise that the issue is of paramount importance. New actors, including from the law enforcement and security sector, have shown solidarity with the Campaign and are available to accompany WILPF in the continuation of the advocacy to turn into action the call of the African Union Commission "to work towards a comprehensive legal and ethical governance framework for AI technologies, robotics and other new and emerging technologies so as to ensure compliance with the African Charter and other regional treaties." This logic takes on a truly broad meaning in the global demand for a binding instrument against AWS.

CCW REPORT

Reaching Critical Will (RCW) is the disarmament programme of the Women's International League for Peace and Freedom (WILPF), the oldest feminist peace organisation in the world. RCW works for disarmament and the prohibition of many different weapon systems; confronting militarism and military spending; and exposing gendered aspects of the impact of weapons and disarmament processes with a feminist lens. RCW also monitors and analyses international disarmament processes, providing primary resources, reporting, and civil society coordination at various UN-related forums.



Reaching Critical Will

www.reachingcriticalwill.org



www.wilpf.org

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 to Stop Killer Robots.

CCW REPORT

Vol. 9, No. 6

23 September 2021

Editors: Ray Acheson and Allison Pytlak

Contact: disarm@wilpf.org