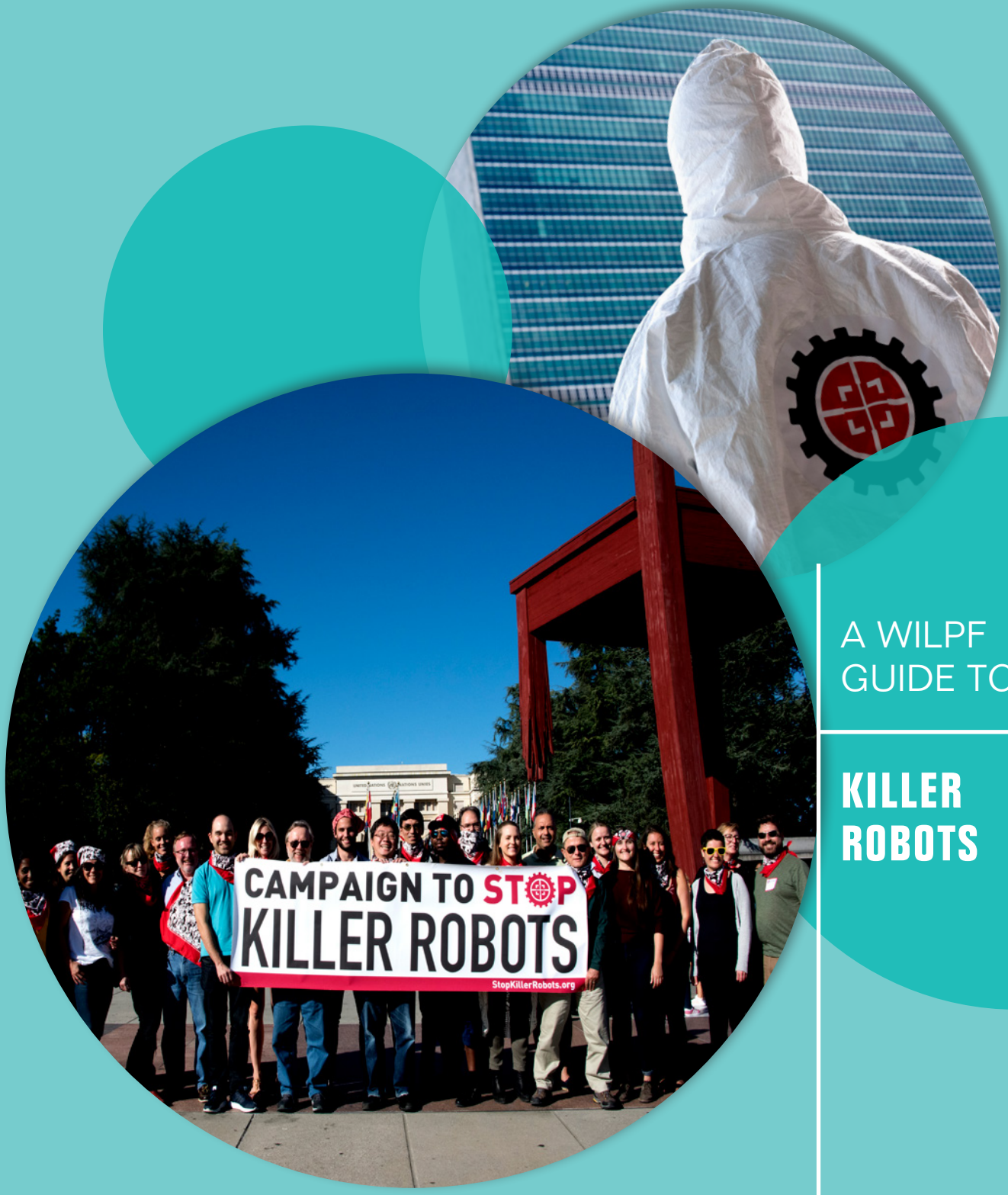

A WILPF GUIDE TO

KILLER ROBOTS

WOMEN'S INTERNATIONAL LEAGUE FOR PEACE & FREEDOM

Table of Contents

January 2020

A User Guide to Killer Robots

3rd Edition

14 pp.

Authors: Ray Acheson

Design: Nadia Joubert

Cover photo: © Clare Conboy, Campaign to Stop Killer Robots

www.wilpf.org | **www.reachingcriticalwill.org**

Background on Killer Robots

What are Killer Robots?

Killer robots are autonomous weapon systems. These are weapons that operate without meaningful human control.

The term "meaningful human control" refers to control over the selection and engagement of targets, that is, the "critical functions" of a weapon. This means when, where, and how weapons are used; what or whom they are used against; and the effects of their use.

Autonomous weapon systems are different than armed drones. Drones, or "uncrewed aerial vehicles" (UAVs), are remotely piloted by humans. Human operators fly the drones by remote control, select targets, and choose when to fire upon those targets. An autonomous weapon would be programmed so that once it is deployed, it operates on its own. It would be able to select and fire upon targets all on its own, based upon its algorithms and data analysis programming.

In essence, this means that machines would have the power to kill human beings or destroy buildings or other infrastructure, without a human making the decision.

Why Does Anybody Want to Develop Killer Robots?

Some military personnel or other government officials believe that killer robots would give them an advantage in battle while saving their own soldiers' lives.

Proponents of autonomous weapon systems argue that these weapons will keep human soldiers in the deploying force out of danger. They also argue that autonomous weapons will be more "precise" in hitting their targets. They say robots will make calculations and decisions more quickly than humans, and that those decisions—in targeting and in attack—will be more accurate than those of humans.

They also argue that the weapons will not have emotional responses to situations—they will not go on a rampage out of revenge; they will not rape. In this sense, they purport that killer robots will comply better with the mission they have been programmed to carry out.

Who is Developing Killer Robots?

> **More than 380 partly autonomous weapon systems have been deployed or are being developed in at least 12 countries, including China, France, Israel, Republic of Korea, Russia, the United Kingdom, and the United States.[1]**

The Republic of Korea deploys mechanised sentries in the demilitarised zone, while Israel's Iron Dome detects and destroys short-range rockets. The **United States'** missile-defence systems like the Patriot and Aegis are semi-autonomous, and the US military has completed testing of an autonomous anti-submarine vessel, which is able to sink other submarines and ships without anyone on board. The **United Kingdom** is developing Taranis, a drone that can avoid radar detection and fly in autonomous mode. **Russia** has built a robot tank that can be fitted with a machine gun or grenade launcher, and has manufactured a fully automated gun that uses artificial neural networks to choose targets. **China** is developing weapon "swarms"—small drones that could be fitted with heat sensors and programmed to attack anything that emits a body temperature. If this trend continues unconstrained, humans will eventually be cut out of crucial decision-making.

What Are Some Concerns About Killer Robots?

> **While the countries that want to develop killer robots say they could have positive benefits, many roboticists, scientists, tech workers, philosophers, ethicists, legal scholars, human-rights defenders, peace-and-disarmament activists, and governments of countries with less-advanced militaries have called for an international ban on the development of such weapons. They are concerned that these weapons will result in more civilian deaths, be unable to comply with international humanitarian law or human rights law, make war more likely, encourage an arms race, destabilise international relations, and have moral consequences such as undermining human dignity.**

Here are some key objections to the development of killer robots:

The use of force has already become too disengaged from human involvement with the use of armed drones. Autonomous weapons go beyond remotely-controlled drones, devolving life and death decision-making to software and sensors. As machines, fully autonomous weapons would lack the inherently human characteristics such as compassion that are necessary to make complex ethical choices.

Algorithms would create a perfect killing machine, stripped of the empathy, conscience, or emotion that might hold a human soldier back. Proponents of autonomous weapons have argued that this is exactly what would make them better than human soldiers. They say machines would do a better job of complying with the laws of war than humans do, because they would lack human emotions. But this also means they would not possess mercy or compassion. They would not hesitate or challenge a commanding officer's deployment or instruction. They would simply do as they have been programmed to do—and if this includes massacring everyone in a village, they will do so without hesitation.

Any decision to use force should be made with great care and respect for the value of human life. From a moral perspective, the power to come to such a decision should rest with humans because

they are endowed with reason and possess "prudential judgment,"[2] the ability to apply broad principles and past experience to particular situations. Because the exercise of prudential judgment depends on more than numeric analysis of data on lawful and unlawful attacks, it would be very difficult for a fully autonomous weapon, no matter how much data it could process, to exercise this sort of judgment.

Autonomous weapon systems cannot be relied upon to comply with international humanitarian law or human rights. Robots programmed to kill might accidentally kill civilians by misinterpreting data. They would also lack the human judgment necessary to evaluate the proportionality of an attack, distinguish civilian from combatant, and abide by other core principles of the laws of war. Many tech workers, roboticists, and legal scholars believe that we will never be able to programme robots to accurately and consistently discriminate between soldiers and civilians in times of conflict.

> Although progress is likely in the development of sensory and processing capabilities, distinguishing an active combatant from a civilian or an injured or surrendering soldier requires more than such capabilities. It also depends on the qualitative ability to gauge human intention, which involves interpreting the meaning of subtle clues, such as tone of voice, facial expressions, or body language, in a specific context.
>
> – Bonnie Docherty of Harvard Law School and Human Rights Watch[3]

This delegation of violence also has implications for accountability and liability. Who is responsible if a robot kills civilians or destroys houses, schools, and marketplaces? Is it the military commander who ordered its deployment? The programmer who designed or installed the algorithms? The hardware or software developers? We cannot lock up a machine for committing war crimes—so who should pay the penalty? This accountability gap would make it is difficult to ensure justice, especially for victims.

Autonomous weapons could be used in other circumstances outside of armed conflict, such as in border control and policing. They could be used to suppress protest or to prop up regimes. Force intended as non-lethal could still cause many deaths.

Autonomous weapon systems risk lowering the threshold for war. They present a perception of "low risk" and "low cost" to the military deploying the weapon. This perception increases the scope for the deployment of weapons into situations and to carry out tasks that might otherwise not be considered possible. Replacing troops with machines could make the decision to go to war easier. The implications of having an amoral algorithm determine when to use of force means that we will likely see more conflict and killing, not less. As we have seen with armed drones, remote-controlled weapons have made war less "costly" to the user of the weapon. Operators safely ensconced in their electronic fighting stations thousands of miles away do not face immediate retaliation for their acts of violence. While this is obviously attractive to advanced militaries, which do not have to risk the lives of their soldiers, it arguably raises the cost of war for everyone else. It lowers the threshold for the use of force, especially in situations where the opposing side does not have equivalent systems to deploy in response. In the near future, autonomous weapon systems are not likely to result in an epic battle of robots, where machines fight machines. Instead, they would likely be unleashed upon populations that might not be able to detect their imminent attack and might have no equivalent means with which to fight back. Thus the features that might make autonomous weapons attractive to technologically advanced countries looking to preserve the lives of their soldiers will inevitably push the burden of risk and harm onto the rest of the world.

These features also fundamentally change the nature of war. The increasing automation of weapon systems helps to take war and conflict outside of the view of the deploying countries' citizenry. If its own soldiers are not coming home in body bags, will the public pay attention to what its government does abroad? Does it care about the soldiers or the civilians being killed elsewhere? From what we have seen with the use of drones, it seems that it is easier for governments to sell narratives about terrorism and victory if their populations cannot see or feel the consequences themselves.

Autonomous weapons will likely be more susceptible to cyberattacks or hacking, as well as problems associated with machine-learning or artificial intelligence. Introducing these technologies into weapon systems risks making them more unpredictable and less controllable.

Notes

1. Mattha Busby, "Killer robots: pressure builds for ban as governments meet," The Guardian, 9 April 2018, https://www.theguardian.com/technology/2018/apr/09/killer-robots-pressure-builds-for-ban-as-governments-meet.

2. "Killer Robots and the Concept of Meaningful Human Control," Human Rights Watch, April 2016, https://www.hrw.org/news/2016/04/11/killer-robots-and-concept-meaningful-human-control.

3. Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban, Human Rights Watch and International Human Rights Clinic, 2016, https://www.hrw.org/sites/default/files/report_pdf/arms1216_web.pdf.

Sylvie Ndongmo of WILPF Cameroon and Katrin Geyer of WILPF's Disarmament Programme, November 2018
Photo © Clare Conboy, Campaign to Stop Killer Robots

WILPF's Work on Killer Robots

Why Does WILPF Want to Prevent Killer Robots?

> **WILPF has opposed war and the development of technologies of violence since its founding in 1915. We have condemned high levels of global spending on militarism and conflict rather than on the benefit of humankind and the promotion of human security. While WILPF opposes all war and violence, there is something especially cynically abhorrent in the idea of human beings assigning killing to a technological creation. The taking of life requires human accountability, determined by morality and law. Without that we shirk our responsibilities and betray our common humanity.**

WILPF has delivered several statements to UN meetings on autonomous weapons and has provided analysis and reporting on all relevant international discussions. We have used these opportunities to express WILPF's position against the development of killer robots, including our belief that the laws of war and protection of human beings require that humans must be the ones to make decisions about the use of force.

Human beings are fallible. We can be violent, we can break laws. But we have traits that machines do not have, and likely cannot be programmed to have: moral reasoning, empathy, compassion, mercy. Giving machines power to target and kill human beings crosses a moral line. It points to an increasing remoteness and abstraction of violence. It suggests the further erosion of the value of human life and dignity. In addition to the concerns with killer robots noted above, WILPF has also raised the following key points:

Weapons symbolise power. Whether it is small arms or atomic bombs, weapons have been developed and used to dominate others. The production and proliferation of weapons also mean profits for corporations and their leaders. The potential development of autonomous weapon systems must be seen in the context of power and profit. Corporations will be seeking to make money off the development of these weapons, and high-tech countries will use autonomous weapons to oppress and occupy others.

Countries of the global south may not be the ones to develop and use autonomous weapons, but they will likely become the battlegrounds for the testing and deployment of these weapons. It will be the rich countries using these weapons against the poor—and the rich within countries using it against their own poor, through policing and internal oppression.

If fully autonomous 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 WILPF's 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.

At the end of the day, the answer is simple: **Weapons must be under human control.** We already experience far too much violence among human beings. How can we risk further automating this violence? Fighting to retain human control over violence is not just about preventing mechanised death and destruction; it is also about calling ourselves to account for the violence we already perpetuate. Maybe this can be a wake-up call for us all—one that we would do well to heed now, before it is too late.

What Do Gender and Feminism Have to Do With Killer Robots?

> **WILPF brings a unique gender analysis to questions of disarmament and arms control. In our work on killer robots, we highlight the ways in which autonomous weapon systems exacerbate discriminatory gender norms and can be used to commit acts of gender-based violence; and we promote a feminist approach to opposing emerging technologies of violence.**

Autonomous weapons are being developed in the context of the established norms of gender and power. These norms can and do affect how we think about weapons, war, and violence. Throughout history, we have seen that weapons symbolise power. The association of weapons with power comes from a dominant understanding of masculinity, in which ideas like strength, courage, and protection are equated with violence. This is not to say that all men agree with or perpetuate this idea, but that this is widely considered the norm or standard for masculinity. It is a masculinity in which the capacity and willingness to use weapons, engage in combat, and kill other human beings is seen as essential to being "a real man".

Turning men (and women) into warfighters has tended to require breaking down their sense of ethics and morals and building up a violent masculinity that is lacking in empathy and glorifies strength as violence and physical domination over others portrayed as weaker. **Autonomous weapons would be the pinnacle of a fighting force stripped of the empathy, conscience, or emotion that might hold a human soldier back.** While armed drones are on this trajectory towards mechanised violence, autonomous weapons would complete the separation of the body and mind from the battlefield and the process of dehumanisation of warfare.

This process of dehumanisation also often includes sexual – or other kinds of gender-based violence. Some people who support the development of killer robots have argued that these weapons will be better than human soldiers because they will not rape. This is a myth. An autonomous weapon could be programmed to inflict terror on a population through sexual violence, which is already ordered by states and armed groups using human soldiers. An autonomous weapon, if programmed to rape, would not hesitate to do so. It is also important to consider the broader culture of rape in relation to weapons and war. Rape and sexual violence are used as weapons in conflict, and is also heightened during and after conflict. War destabilises communities and exacerbates already existing gender inequalities and oppression of women, LGBTQ+ people, and others who do not conform to societies' standards of gender norms.

WILPF is also concerned about the development of autonomous weapons programmed with "target profiles" and algorithmic bias in terms of gender, race, socioeconomic status, ability, and sexual orientation. Facial recognition software struggles to recognise people of colour; voice recognition struggles to respond to women's voices or non-North American accents; photos of anyone standing in a kitchen are labelled as women; people's bail is denied because a program decided that a woman of colour was more likely to reoffend than a white woman; trans people are surveilled on the basis of the clothing they wear. Imagine this kind of bias being programmed into a weapon system designed to target and fire upon targets without any meaningful human control, without any human judgment to counteract that bias. Or, imagine that bias being deliberately programmed into the killer robot, to kill all people of a certain ethnicity, religion, or sex.

In this way, an autonomous weapon contributes to further dehumanising people by turning them into target profiles and expendable targets. A good example of this is how already in conflict, civilian men are often targeted—or counted in casualty recordings—as militants only because they are men of a certain age. While men are not necessarily targeted only because they are men, taking sex as a key signifier as identity and exacting harm on that basis constitutes gender-based violence. This erodes the protection that civilians should be afforded in conflict and violates many human rights, including the right to life and due process. It also has broader implications in the reinforcement of gender norms, including violent masculinity. Assuming all military-age men to be potential or actual militants or combatants entrenches the idea that men are violent and thus targetable. This devalues male life—it suggests men are relatively more expendable than women. It increases the vulnerability of men, exacerbating other risks adult civilian men face such as forced recruitment, arbitrary detention, and summary execution.

The gendered culture of violent masculinities that surrounds the development of autonomous weapons, likely to be embedded within the technology and its use, will create new challenges for preventing violence, protecting civilians, and breaking down gender essentialisms or discrimination. Understanding how autonomous weapons are likely to be perceived in a gendered way by their developers, operators, and their victims is crucial to developing policies that can help break the cycle of violence. This could include an understanding that the operation of weapons without meaningful human control, weapons programmed to target and kill based on pre-programmed algorithms of who is considered to pose a threat, used without consent in foreign lands or in the streets of local cities, will result in civilian casualties, psychological harm, and destruction of civilian infrastructure. That this in turn will result in a violent masculine response from affected communities, reinforcing gender inequalities and oppressions.

Such understandings should have significant implications for our thinking about and approach to the development of autonomous weapons. We can think about how this kind of analysis and argumentation could help tech workers and policy experts see the need for meaningful human control over weapon systems. We can also see how autonomous weapon systems could facilitate the profiling, targeting, dehumanisation, and death of various marginalised communities and as well as activists or others confronting power. In a context where weapons are treated as tools of power, violence, and subordination of others, increasing the remoteness and abstraction of violence is not the answer. Dealing with violence and conflict as a social institution, rather than a technical challenge to be "solved" with new weapons technology, is imperative. Understanding these dimensions of both violence and technology could help contextualise our work against weapons in a broader context of social justice.

What Does WILPF Think the Solution is?

> **WILPF believes that the best solution is a legally binding international treaty to prohibit the development, production, and use of fully autonomous weapons.**

This is the best way to ensure that meaningful human control is retained over all targeting and attack decisions. This is the main demand of the **Campaign to Stop Killer Robots.**

The international treaty would also require **national legislation** from each country, to ensure against the development of killer robots.

WILPF also believes that **technology companies, tech workers, scientists, academics, and others involved in developing artificial intelligence or robotics should pledge** to never contribute to the development of fully autonomous weapons.

Financial institutions such as banks and pension funds should also pledge not to invest money in the development or manufacture of killer robots.

What Has WILPF Done So Far?

WILPF is a founding member of the **Campaign to Stop Killer Robots** (www.stopkillerrobots.org). WILPF is on the Steering Committee of the Campaign and contributes to the Campaign's advocacy with governments and analysis of international meetings.

Through its disarmament programme **Reaching Critical Will** (www.reachingcriticalwill.org), WILPF has monitored and reported on all of the UN meetings dealing with autonomous weapons. We produce a daily report at each meeting with analysis and advocacy, deliver statements, participate in events, and work with other campaigners to encourage states to support a ban on autonomous weapons and the retention of meaningful human control over all weapon systems.

WILPF members from various Sections have attended UN meetings on autonomous weapons, and now several Sections and Groups in Africa have received small grants from the Campaign to Stop Killer Robots to conduct national-level advocacy in support of the prohibition of autonomous weapons.

How can you get involved?

Sign up to receive news from the Campaign to Stop Killer Robots: **www.stopkillerrobots.org/act/**

Encourage your government to support the negotiation of a treaty banning autonomous weapons and the retention of meaningful human control over weapon systems.

Encourage tech companies, tech workers, university students, academics, scientists, and financial institutions to pledge to not be involved in the development of killer robots.

What is Happening Now?

What is Happening in the World in Relation to Killer Robots?

> **There is growing momentum amongst governments, activists, and the tech and scientific community against the development of autonomous weapon systems. The Campaign to Stop Killer Robots has been active in advancing the prohibition treaty at the United Nations and in working with technology and academic communities to ensure that weapons without meaningful human control are never developed or deployed.**

So far, 29 states have called for a prohibition on fully autonomous weapons: Algeria, Argentina, Austria, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Djibouti, Ecuador, Egypt, El Salvador, Ghana, Guatemala, Holy See, Iraq, Jordan, Mexico, Morocco, Namibia, Nicaragua, Pakistan, Palestine, Panama, Peru, Uganda, Venezuela, and Zimbabwe.

The Non-Aligned Movement, the largest bloc of states operating in the UN, has called for a legally binding instrument stipulating prohibitions and regulations of autonomous weapon systems.

In 2018, **Austria, Brazil, and Chile** collectively tabled a recommendation for states "to negotiate a legally-binding instrument to ensure meaningful human control over the critical functions" of weapon systems.

A few others, mostly European states, expressed their interest in other mechanisms, such as a **political declaration** proposed by France and Germany. They envision a declaration to be a good vehicle to outline principles for the development and use of autonomous weapon systems, such as the necessity of human control in the use of force and the importance of human accountability. Some also suggested the development of a **code of conduct** on the development and use of autonomous weapon systems could be useful in this context.

Only Australia, Israel, Republic of Korea, Russia, and United States are blocking any of these initiatives from going forward. They argue that negotiations of a treaty or a political declaration or other mechanisms are "premature". They believe that autonomous weapons will be beneficial to their militaries and want to keep their options open.

The UN discussions on killer robots are currently taking place within the Convention on Certain Conventional Weapons (CCW) in Geneva. Unfortunately, this treaty body operates on the basis of consensus, which means that even though a tiny handful of states oppose negotiating a ban or developing other instruments, they have been able to block progress so far.

But tech workers and others are rising up against killer robots. In April 2018, Google employees sent a letter demanding the company commit to never build "warfare technology". In response, Google executives said they would not renew the company's participation in a US Department of Defense contract. **More than 800 scholars, academics, and researchers who study, teach about, and develop information technology released a statement in solidarity with the Google employees.** They called on the company to support an international treaty to prohibit autonomous weapon systems and commit not to use the personal data that Google collects for military purposes. **In February 2019, Microsoft employees demanded the company cancel its work on a contract with the military.** More than 100 employees signed a letter sent to the CEO and President criticising the company's plans to build HoloLens AR tech for the military, stating, "We did not sign up to develop weapons, and we demand a say in how our work is used." **Several tech workers have left their jobs in protest of these types of developments and some have joined the Campaign.**

At a major world congress of leading artificial intelligence researchers in July 2018, over 200 technology companies and organisations from more than 36 countries and 2,600 individuals issued a pledge committing to "neither participate in nor support the development, manufacture, trade, or use of lethal autonomous weapons." Citing moral, accountability, proliferation, and security-related concerns the pledge finds that "the decision to take a human life should never be delegated to a machine."

In May 2018, UN Secretary-General António Guterres pledged to support states to elaborate new measures such as "legally binding arrangements" to ensure that "humans remain at all times in control over the use of force." He has said killer robots are "politically unacceptable, morally repugnant and should be prohibited by international law."

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."

In July 2018, **the European Parliament** adopted a resolution calling for the urgent negotiation of "an international ban on weapon systems that lack human control over the use of force." The day before, **Belgium's national parliament** adopted a resolution that calls for a ban on fully autonomous weapons. In May 2019, the **Dutch parliament** adopted a resolution calling for a legally-binding instrument on new weapons technologies, including autonomous weapons. In July 2019, parliamentarians from the **Organisation for Security and Co-operation in Europe (OSCE)** urged states to support legally binding rules on autonomous weapons. The Campaign to Stop Killer Robots has received **pledges of support from parliamentarians and the public** at events in capitals including Abuja, Helsinki, London, Pretoria, Seoul, Tokyo, and Yaoundé.

In December 2018, a survey of 26 countries by Ispos found that 61 percent of people oppose the development of killer robots. Most said these weapons would "cross a moral line because machines should not be allowed to kill."

This is a guide about autonomous weapons, also known as killer robots. It provides information on the current status of related technologies as well as international efforts by governments, activists, and tech workers to prevent the development of these weapons. It outlines some of the key legal, moral, political, and technical concerns about autonomous weapons and it explains, from a feminist perspective, what the gendered impacts of these weapons could be.

Our intention is that this publication can help inform those concerned with autonomous weapons to take action to help prevent their development, including through the global Campaign to Stop Killer Robots and other international and local initiatives.

WILPF Geneva
Rue de Varembé 1
Case Postale 28
1211 Geneva 20
Switzerland

T: +41 (0)22 919 70 80

E: info@wilpf.org

WILPF New York
777 UN Plaza, New York
NY 10017 USA

T: +1 212 682 1265

WOMEN'S INTERNATIONAL LEAGUE FOR
PEACE & FREEDOM