KILLER ROBOTS ARE NOT “INEVITABLE”

Jody Williams | Nobel Peace Prize Laureate, Chair of the Nobel Women’s Initiative

If the word “inevitable” were a mallet, we would long ago have been beaten to the ground by it. But that is part of the purpose of its relentless use. It is disempowering. Deadening. If something is inevitable, there is no point in challenging it.

The Campaign to Stop Killer Robots completely rejects the notion that autonomous weapon systems are inevitable. Just as blinding laser weapons were banned before they showed up on the battlefield, it is possible to prohibit autonomous weapons before they are used in conflict.

Along with the inevitability factor, we hear over and over again from some that it is “premature” to move beyond informal discussions in the CCW because such deadly technologies are so complex that states need more time to understand them in order to determine if or how they might be regulated. This is another diversionary tactic.

For example, I do not have to be a nuclear physicist to recognise the humanitarian and legal reasons that nuclear weapons must be banned now. And I don’t have to be an expert to see that the nuclear-armed states’ plan to “upgrade” their weapons instead of moving forward with eliminating them despite the desires of 80% of the world’s population.

Not surprisingly, it is some of the nuclear-armed states that have been in the lead in developing autonomous weapons technology that say they believe it is too early to do much more than continue talking about killer robots—even as the research and development steadily moves forward.

But here in the second series of meetings on these weapons in the CCW, it has been other voices in the ascendancy. These voices recognise the moral and ethical bankruptcy of moving toward creating autonomous weapon technologies that would be able to target and kill human beings. States have repeatedly stated the need for meaningful human control over weapons used to kill other human beings. Some have called for forward movement in the CCW beyond these informal sessions. Some are staunch proponents of prohibiting these weapons.

Just as killer robots are not inevitable, it is not inevitable that the CCW limit itself to being a forum for achieving the lowest common denominator and, in effect, make no meaningful impact on reigning in these weapons before it is too late. The CCW can demonstrate the same dramatic leadership it did when it banned blinding laser weapons in 1995 and exert “meaningful human control” over killer robots and take timely, forceful action to stop them.

This is a decision that we as human beings can make. It is a decision that we must make. We must not cede the power of life and death over other human beings to autonomous machines of our own design. It is not too late. Killer robots are not inevitable.

Chinese delegation argued that it would be better to take precautionary measures than deal with the aftermath of autonomous weapon systems. Mexico argued that weapons that conflict with IHL should be prohibited.

continued on next page
Cover article, continued

As the Irish delegation conveyed forcefully, the mandate of the CCW and its protocols “is to regulate or ban the use of specific categories of conventional weapons that have effects which trouble the conscience of humanity.” It remains to be seen how many delegations will be ready to move towards negotiations, based on their concerns that this technology could fall foul of humanitarian and human rights law. As the Dutch civil society group PAX noted in its statement, failing to take action on the development, production, and use of autonomous weapons also has ethical implications. It is time for action. Delegations should use this week to set out their proposals for how they intend to take forward concrete work on autonomous weapons in the CCW, the Human Rights Council, and in their national debates at home. •

EVENT: LISTEN TO THE DICTATES OF THE PUBLIC CONSCIENCE

Elisabeth Minor | Article 36

Today’s side event organized by the Campaign to Stop Killer Robots considered the fundamental ethical questions posed by increasing autonomy in weapon systems. Yesterday, the ICRC urged delegations to give their views on the question of whether it would ever be morally acceptable for a machine to make life and death decisions without human intervention. Many states have already emphasised that it would not.

Ryan Gariepy’s company Clearpath Robotics pledged in 2014 not to develop weaponised robots where these did not have a human in the loop. Governments, NGOs, and industry must unite on this issue so that development of fully autonomous weapons systems (AWS) is not funded or encouraged. Gariepy emphasised that prohibiting AWS would guide the development of beneficial civilian and military applications of emerging technology, rather than restrict it. He noted that as the flexibility of systems increases—as is required in the battlefield—predictability must be sacrificed. This is why meaningful human control must be retained in order for systems to be effective, ethical, and beneficial.

Peter Asaro of the International Committee for Robot Arms Control highlighted that all law is based on principles of humanity and the dictates of the public conscience. The Martens Clause enshrines consideration of these. He noted that increasing autonomy in weapons systems will disrupt and challenge existing legal and moral norms around accountability, acceptable risk to civilians, and human dignity. The CCW must anticipate this. States must engage in norm innovation, using the principle of meaningful human control to discuss and articulate what disturbs us about autonomous weapons, and codify this in law.

AJung Moon of the Open Roboethics initiative described her public opinion research on AWS. Public opinion surveys can assist in developing an understanding of the state of public conscience. Data gathered so far shows an overwhelming rejection of delegating life and death decisions to AWS. These voices should not be ignored. Transparency on AWS, which many states have called for, should include transparency to the public.

The hypothetical scenario of an AWS that would be absolutely certain not to harm civilians, and whether this would be acceptable, dominated subsequent discussion. The panelists emphasised the impossibility of absolute certainty in any system and observed that raw numbers of civilian casualties should not be our only consideration. Other fundamental questions such as ensuring accountability, responsibility, and human dignity must also be weighed up. •
The News in Brief provides quick highlights from presentations and discussions; it is not a comprehensive record.

**Technical issues, part II**

- Elizabeth Quintana suggested that autonomous weapon systems (AWS) are attractive because they might reduce costs, increase situational awareness, make better decisions, and, reduce costs of personnel; but she also argued they might increase the likelihood of military intervention, have potential for unlimited use, destabilise security environments, and risk an arms race.

- Professor Heather Roff summarised some pros and cons of increasing autonomy in land, sea, and air weapons. She noted that autonomy can offer expanded capabilities across these theatres of operation but that political scrutiny, high costs, interoperability issues, and unintended or inefficacious effects constitute serious risks.

- Colonel Wolfgang Richter highlighted some of the perceived advances and disadvantages to autonomy in warfare from a military perspective. He argued that the use of AWS would not be the same as tactical autonomy, because autonomous systems can destroy targets but cannot take or hold ground or make other tactical decisions.

- Dr. Darren Ansell highlighted the difficulties of thoroughly ensuring against reliability and vulnerability problems in AW software, cautioning that failure could lead to unintended fatalities. He noted that false or missing software requirements, incorrect algorithms or code, inadequate testing, incorrect or unexpected usage of the software, and also possible vulnerabilities to cyber attacks, expose systems to risk.

- Professor Frédéric Vanderhaegen discussed the instability of AWS due to dissonance, which affects knowledge, availability, or prescription in a system. Warning that the “enemy” can always find weaknesses in a system, he outlined some of the aspects of dissonance that could affect an autonomous weapon, including errors, voluntary violation of rules, ill will, sabotage, terrorism, new functionalities / alternations to the system, and incorrect learning.

- During the Q&A, some delegations questioned the implications of autonomous weapons for war and peace. Palestine argued that AWS would reduce the opportunity for dialogue and humanity, while India asked if AWS would introduce challenges to ending conflict.

- Cuba called for the prohibition of AWS through a legally-binding instrument.

- Col. Richter said to prevent arms races there is need for transparency around development of AWS and doctrines, article 36 processes, code of conduct, etc. He suggested developing new arms control process to ensure AWS are regulated sufficiently.

- Switzerland asked if the speakers if the development and use of AWS could affect existing command and control structures. Col. Richter said no, but Ms. Roff argued that the quest to make AWS survivable, cost effective, and operationally effective means putting targeting capacities on board the system, which means relinquishing certain powers that a commander would have to machines.

- Ireland asked if learning by an automated system can be predicted.

- Ms. Quintana argued that automatic systems are more reliable than human operators.

**Characteristics, part I**

- Maya Brehm described the characteristics of meaningful human control (MHC) and its relation to the use of armed force, noting that it requires control over who or what is harmed but also where, why, and how. She also described the process of “meaning making,” the social practices by which we interact with each other to make common sense of the world, and the relationship of these practices to moral judgments, values, accountability, and responsibility, all of which are critical for establishing the legitimacy and legality of organised armed violence.

- Dr. Marcel Dickow outlined some vectors of autonomy, noting that four useful criteria for evaluating AWS include time, space, sensors/recognition, and communication.

- Dr. Neil Davidson called for a focus on autonomy in the use of force, not on technical sophistication. He also outlined some of the ways in which MHC is context-dependent, including supervision of weapon, its freedom of action, etc.

- Professor Nehal Bhuta suggested there seems to be general understanding that MHC ought to be preserved except perhaps in circumstances when a weapon’s deployment are narrow temporally and geographically. He also highlighted the relationship between normativity and legality in the concept of MHC. And he cautioned that article 36 reviews are in danger of being formalistic without sufficiently high standards.

- During the Q&A, delegations including Australia, Netherlands, and Poland asked questions about MHC.

- Switzerland suggested that MHC should be discussed in relation to a clearly identified goal.

- France expressed a preference for discussing autonomy rather than MHC.
The principle of ensuring meaningful human control over the use of weapon systems has been asserted frequently by some states, international organisations, and civil society. It continues to be a key point of discussion at the CCW this year, referenced in delegations’ statements, expert panels, and pre-circulated documents. Members of civil society and some others have proposed that the principle of requiring meaningful human control over every individual attack is used as a basis for agreement by states to draw the necessary prohibitions on certain weapon systems. What are the prospects for agreement on this principle moving forward the CCW’s consideration of autonomous weapon systems?

Prior to the CCW expert meeting, the “food-for-thought” paper circulated by the German chairperson suggests questions for considering what meaningful human control means and whether the level of human control helps distinguish autonomous weapons from other weapon systems. The paper asks what level of human supervision of a weapon system is implicit in the rules of international humanitarian law (IHL) and whether, as autonomy increases, IHL will need to be clarified. Regarding the way forward, the paper asks whether there is general recognition that some level of human control of a weapon system is necessary and if so what can be done on that issue. Austria devoted a working paper ahead of the meeting to meaningful human control and IHL, noting that “proportionality thus requires a distinctly human judgement”. Japan also highlighted meaningful human control in its paper.

So far, seventeen states in their statements have explicitly supported the position that weapons must remain under meaningful human control (see Vol. 2, No. 2 of the CCW Report for a list). Yet more states have referred in some way to human control more broadly. Fourteen, including some that have not explicitly endorsed the principle of meaningful human control such as Canada and Poland, see the concept as key to discussions moving forward. Five have highlighted the need for discussion on where the limits of meaningful human control lie, including the Czech Republic and South Africa. A few more explored the possibility that it may provide the basis for placing limits on the development of autonomous weapons, or endorsing its inclusion in legal parameters, including Chile, Croatia, and Mexico. Some states have also observed that convergence around the concept sets the direction for moving discussions forward.

Meaningful human control was made the subject of a specific discussion in the annotated programme of work, during Tuesday’s session on characteristics of autonomous weapons. All four expert presentations focused on the principle, prompting detailed interventions and questions on developing the concept from several states. The ICRC noted on Monday that states should now “turn their attention to agreeing a framework for determining what makes human control of a weapon meaningful or adequate,” grounding discussion on consideration of “current and emerging weapon systems that are pushing the boundaries of human control.” In-depth discussion on elaborating the content of this principle is developing, and it is up to states to continue this in order to make the principle operational.

News in brief, continued

• France also said it has no plans for autonomous weapons that deploy fire.
• The UK said it sees no military utility in a fully AWS but that if they existed the principles of targeting would be the same as for humans and noted that such decisions are best made with human judgment.
• Ms. Brehm asked the UK if we accept automatic target recognition and say we will have MHC, what does that mean for operation for that software—e.g. what limitations must be placed on where and how it is used? She argued that MHC requires every individual attack be supervised.
• China identified convergence on concept of MHC but argued this doesn’t mean states that haven’t made a statement are also in favour of this concept.
• Belarus supported exploring MHC and called for an agreement before AWS exist.
• Cuba cautioned that while AWS don’t yet exist there are many questions about accountability and legality around their use that must be answered.
• Dr. Dickow argued that self-learning capabilities have potential to undermine MHC.
• ICRAC noted MHC is to be a normative not a technical standard and asked about the ideal tool of governance for MHC.
• India and Pakistan expressed confusion over MHC, noting that if there is MHC then there will not be AWS.
• Dr. Bhuta explained that MHC is indeed inconsistent with autonomy; this makes it clear what is called for in a ban. He argued that the prohibition of AWS is legally sound and ethically obvious.