



AI will not set its own limits in warfare. States must.

October 15, 2026, Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Autonomous Weapons / IHL / New Technologies

13 mins read



Cordula Droege

Director of International Law,
Movement and Partnerships,
ICRC



Photo source: Unsplash.

Artificial intelligence (AI) is being developed at extraordinary speed, prompting warnings from within the technology industry itself that advances may be outpacing our ability to understand and control its risks. Nowhere are the stakes higher than in armed conflict. As militaries increasingly explore the integration of AI into weapons and military decision-making, technological capability can easily be mistaken for inevitability: if a system can be built, the assumption follows that it will eventually be deployed. History suggests otherwise. States have made choices before about which technologies of warfare to prohibit, constrain or refrain from using – and they face such a choice again today.

In this post, ICRC Director of International Law, Movement and Partnerships Cordula Droege argues that abstract debates about the need to address the risks of AI should not distract from a concrete and achievable task already before states: establishing legally binding rules on autonomous weapon systems (AWS), including those incorporating AI. Drawing on the realities of deploying complex technologies in the unpredictable conditions of armed conflict, she explains why technical safeguards alone cannot resolve

the humanitarian, legal and ethical concerns raised when weapons select and apply force without human intervention. Ahead of the Seventh Review Conference of the Convention on Certain Conventional Weapons in November, she calls on states to turn more than a decade of work into clear prohibitions and restrictions on AWS, while there is still time to decide what role machines and algorithms should play in war.

ICRC Humanitarian Law & Policy Blog - AI will not set its own limits in warfare. States must.

Last month, American technology company CEOs [made the case for somewhat slowing down the development of AI](#). Not for the first time, they clearly warned that the technology may be advancing faster than our ability to understand its consequences and put appropriate safeguards in place.

One week earlier, in the early hours of Saturday 5 September, in Geneva, the [Convention on Certain Conventional Weapons' Group of Governmental Experts on Lethal Autonomous Weapons Systems](#) (CCW GGE on LAWS) closed its final session after three years of work.

On the face of it, these events concern quite different questions. One is about the pace of AI development, the other about international rules for weapons. But the two are connected. [Trends indicate](#) a push to integrate more complex algorithms into weapons, to increase their autonomy. The risks posed by AI – and especially the risks flagged in recent discussions that AI escapes the prediction of those who launch it – are perhaps the most consequential when it is integrated into weapons that can directly affect whether people live or die.

Both events therefore concern the question of whether the international community will act effectively to prevent potentially catastrophic harm from AI, or delay and be left to witness it once it occurs. At the recent [UN Security Council meeting on 23 September](#), calls were made for “governance”, “frameworks” and “norm-setting”. These are welcome, but if they stay at this level of generality such calls will have no impact in reducing the suffering of people living through armed conflicts.

At the International Committee of the Red Cross (ICRC), we always look at new technologies of warfare through the lens of their human cost in armed conflict and of international humanitarian law (IHL). We consider how these technologies might be used (and misused) in warfare, and how they risk affecting civilians and combatants, and undermining crucial legal protections.

The insights of industry are crucial. AI developers and researchers have a key role in identifying risk, building safeguards, testing honestly, and saying plainly what a system can and can't do. But naming a risk isn't the same as deciding what to do about it.

States have a responsibility – to their own populations, and to those in other states who will face the consequences of the proliferation and use of AI-enabled weapons and by conflict escalation – to set clear limits. A company cannot determine what international law, including IHL, permits. No safety testing, voluntary commitment or corporate policy substitutes for international rules where a new technology of warfare creates humanitarian risks that the world cannot accept. That is a decision for states, made together, in public, and one they can be held to. Several states at the Security Council did name autonomous weapons as a concern. None, however, made reference to the work of the [128 High Contracting Parties to the CCW](#) on such systems for over a decade now. Yet, [they have a clear opportunity in November this year](#) to turn this work into binding rules.

What happens when technology meets the battlefield?

War is one of the most high-stakes and dangerous contexts that can be imagined. The consequences of getting things wrong can be tragic. In a commercial environment, if a system produces an unexpected result, it may cause inconvenience or financial loss. On a battlefield, an unexpected result can mean that a civilian is killed or injured, a medical clinic is destroyed. Weapons and other systems used in warfare are therefore necessarily “safety-critical”: their performance can directly affect whether people live or die.

The ICRC has worked in armed conflict for more than 160 years. Our colleagues see what weapons and methods of warfare actually do – they witness the after-effects; they speak to the people in temporary shelters who cannot return because their homes have been destroyed. In conflict after conflict, our colleagues encounter the consequences of decisions made far from the people who bear them.

This experience has shown us that the reality of the battlefield is often more difficult and unpredictable than descriptions from those developing or using weapons suggest, and the gap between expected and actual performance can have a direct human cost.

[In warfare, information is often incomplete](#). Civilians can behave and move erratically, and be present in unexpected places. And there is an enemy force constantly seeking to gain an advantage through techniques of deception and manipulation.

These conditions can be [near impossible to replicate in a testing environment](#). They are difficult conditions for any human decision-maker to assess and to foresee, and they are likely even more challenging for AI systems that rely on quality, repeatable data and statistical patterns to interpret what they are seeing. In such volatile conditions, the performance of AI-enabled weapons can rapidly degrade, producing unpredictable outcomes.

This is why, when we consider AI-enabled weapons, we are not only concerned with whether a system is technically capable, but more broadly whether it can be used in compliance with IHL in the conditions of armed conflict. To answer this question, militaries will need to know the data that the underlying algorithms were trained on and whether that data was representative of the operating environment. They should be able to say how the system will behave when information is incomplete or misleading. They should know what will happen when the system invariably encounters something it has not seen before.

These sound like very technical questions, but they have practical consequences, especially when AI is integrated into an [autonomous weapon system](#). With such a system, the weapon itself selects and applies force to a target, without a person selecting that specific person or object, or the specific time or place. So these questions will ultimately impact whether a civilian, a house, a school or a hospital is hit, or whether a fighting combatant is targeted rather than one who is wounded or surrendering.

Capability is not the same as deployment

AI can process vast quantities of information and generate outputs – whether they be analyses, recommendations or predictions – much faster than a human. We understand that this is a key driver behind the push to integrate AI in military operations. We also understand that [speed can have an operational benefit; but this should not be conflated with a humanitarian benefit](#). In our experience, and in the experience of militaries over decades, increased operational tempo increases the risk of miscalculation and of errors, and these can put both civilians and a military’s own force in danger.

The issue of speed is also relevant to Dario Amodei’s warning, that “[building AI too fast is reckless](#).” Continuing to do so is a choice. Continuing to do so, despite such a warning, is a reckless choice.

How fast states and militaries move to put AI to use in an actual conflict is also a choice, despite the narrative of inevitability about the integration of AI across all aspects of warfare, including in weapons. States have made choices about the use of new technologies on the battlefield before, and the historical record shows that the consequences have cut both ways.

By World War II, chemical weapons had become more sophisticated and lethal. Yet [chemical weapons were not used as a method of warfare on European battlefields](#). A combination of legal rules and deliberate restraint played a key role, since the [1925 Geneva Protocol](#) had been joined by all major parties to the conflict except the US and Japan.

Aerial bombing tells a different story. In 1923, a commission of jurists drafted the Hague Rules of Air Warfare, which sought to restrict bombing of civilian areas and prohibit attacks directed against civilians. [The rules were never adopted as a binding instrument](#). By the 1940s, strategic bombing campaigns had subjected cities to destruction on a scale that those rules had sought to prevent, and in such a way that “[one might have wondered whether \[the customary rule requiring protection of civilians and civilian objects\] still existed](#).”

The fact that a technology exists doesn’t have to be determinative of what today or tomorrow’s wars look like, of how much suffering they cause. It can exist without anyone being ready to accept what happens when it’s used in war. And the key here is that **states have previously been prepared to prohibit, constrain and uphold, even under enormous pressure**.

A similar choice is in front of states now, with autonomous weapons.

Autonomous weapons: a choice to set clear limits now

Some of the risks from military AI may be able to be mitigated through measures such as careful design, better data, rigorous testing, clear use cases and thorough user training. But autonomous weapons raise a more basic question: should some decisions be left to machines at all? Effectively delegating life or death decisions to a machine is a fundamentally dehumanizing process for the person targeted, and [one that removes the moral agency of the person responsible](#). “[The decision to use lethal force cannot be delegated to opaque or automated processes](#).” These concerns cannot be solved by making the technology better; this is a [moral redline](#) that we must not cross. And so this is where the international community needs to make a decision about what uses should be permitted at all.

To address the fundamental ethical, legal and broader humanitarian concerns, [we have called for legally binding rules that would:](#)

[Prohibit those autonomous weapons that pose the greatest threat](#), namely:

- “Unpredictable” autonomous weapons, whose effects cannot be sufficiently understood, predicted and explained. If humans cannot predict what a weapon will do, they cannot exercise the control needed to use it consistently with IHL; and
- Autonomous weapons designed or used to target humans directly, also known as anti-personnel autonomous weapons systems. Here the concern is both legal and ethical (as outlined above): The signals that indicate a person’s status in armed conflict – e.g. whether they are surrendering, injured, or directly participating in hostilities – are highly contextual, and this status can change rapidly. The consequences of getting it wrong will generally be tragic; and
- Place clear restrictions on all other autonomous weapons.

What needs to slow down

The ICRC is not calling for research in military applications of AI to stop. Not every use of AI in military activity carries the same risk. But there are circumstances where slowing deployment is the responsible choice, to allow time to assess existing law, safeguards and risks, to test systems and understand their limitations.

AI-enabled autonomous weapons are a clear example. If weapons that integrate complex algorithms to select and engage targets are widely deployed, setting meaningful international constraints gets much harder. *The opportunity is now* – to set those limits before battlefield experience makes the need for them undeniable – rather than after.

The moment for states to act

There is now a clear avenue for states to set these limits. The GGE on LAWS has finalized its mandate, having produced a significant corpus of work, including a [final report](#) affirming that control and human judgment must be retained in the use of force, and that accountability for decisions involving the use of autonomous weapons cannot be transferred to a machine.

Now, these principles must be turned into clear, binding rules setting clear limits.

Over [70 states already favour negotiation](#). In mid-November states will meet for the Seventh Review Conference of the CCW. This Conference is only held once every five years. *States must not miss that moment; they must use it to decide to launch negotiations* for a treaty on autonomous weapons with clear prohibitions and restrictions.

Otherwise, there is a risk that abstract discussions about governance, standards and safeguards may give the appearance of action, without being it. And that they might distract from the concrete decisions that states can already take, now. **On autonomous weapons, the risks are clear. The work has been done, and a path to binding international rules is available: states must take it.**

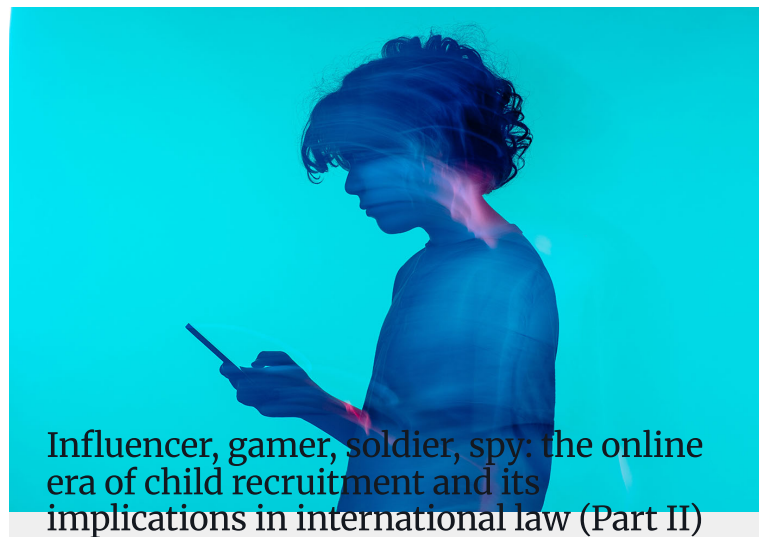
The debate about AI is often framed as a race: between technological development and our ability to regulate it. But there is another important framing, about when and how we choose to use technologies in conflict: in what are already some of the most dangerous environments on the planet. The people in these places, who will live with the consequences, must be front of mind for states when they make these choices.

There is still time to decide what role we want machines and algorithms to play in war.

Related articles

- Eberechukwu Owuamanam SJ, [The machines must not choose who dies: a humanitarian case for disarming AI](#), August 13, 2026
- Laura Bruun and Marta Bo, [‘Constant care’ must be taken to address bias in military AI](#), August 28, 2025
- Anna M. Gielas, [Warfare at the speed of thought: can brain-computer interfaces comply with IHL?](#), August 21, 2025
- Ruben Stewart, [The shifting battlefield: technology, tactics, and the risk of blurring lines in warfare](#), May 22, 2025
- Joanna L D Wilson, [AI, war and \(in\)humanity: the role of human emotions in military decision-making](#), February 20, 2025

Tags: Artificial Intelligence, Autonomous Weapons, AWS, IHL, international humanitarian law



🕒 15 mins read

Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Autonomous Weapons / IHL / New Technologies
Vanessa Murphy, Siobhan Sparkes McNamara & Tilman Rodenhäuser

As social media and gaming platforms create new interfaces between tech-savvy children and armed actors, ...

🕒 14 mins read

Analysis / Artificial Intelligence and Armed Conflict / Artificial intelligence in military decision-making / Autonomous Weapons / IHL / New Technologies
Vanessa Murphy, Siobhan Sparkes McNamara & Tilman Rodenhäuser

As social media and gaming platforms create new interfaces between tech-savvy children and armed actors, ...