Autonomous weapon systems – “Killer robots” in the popular culture – are weapon systems that can select and attack targets without human intervention. The first informal experts’ meeting on lethal autonomous weapons systems (LAWS) was organized in 2014 at the UN Convention on Certain Conventional Weapons (CCW) in Geneva, on the initiative of France, which also presided over the meeting. The last of these annual meetings took place April 11–15, 2016, under German presidency for the second consecutive year. It confirmed the growing interest in the subject from states and civil society: 95 states participated in the debates, alongside several UN institutions, the International Committee of the Red Cross (ICRC), numerous NGOs from around the world, and 34 international experts (compared to 90 states and 30 experts in April 2015, and 87 state and 18 experts in May 2014).

At this meeting states for the first time unanimously adopted general recommendations for the fifth CCW Review Conference, which will take place December 12–16, 2016. The recommendations included the creation of a Governmental Group of Experts (GGE) beginning in 2017, cementing the transition from informal meetings (2014–2016) to a formal process.
Three years of meetings have left civil society and several states increasingly impatient, and so this year there was a clear desire to achieve results. Already anti-LAWS NGOs had indicated that they would no longer be fooled by certain states’ delaying strategies, which have been attempting to drag out the subject of LAWS at the CCW, where the consensus-based decision-making allows them to exert a level of control, rather than meaningfully advancing the discussion.

In particular, there was a growing desire to move forward without resolving all the conceptual problems, against the wishes of states like China, which suggested that “first, we must know what autonomy means” before being able to discuss LAWS. Canada’s statement that it “does not find it useful to debate the meaning of autonomy out of context” captured the wishes of many states to move toward “more concrete” matters.

The April 2016 meeting confirmed that the main topics of discussion have remained unchanged over the last three years, namely, issues of definition, human control, responsibility, and legal review. Rather than going to the root of these questions, this paper will address the procedure, negotiations, the balance of power, and generally the diplomatic dimension of the last round of Geneva debates.

**What Are We Talking About?**

Unlike previous years, this time the question was less the definition of LAWS or autonomy but rather of whether it was even necessary to define them. The delegations initially divided themselves into two camps: those who made the definition a prerequisite condition for any discussion (which was interpreted by others, particularly civil society, as a way of blocking discussion), and those who agreed to leave the definition until later. A consensus was reached that an exhaustive and final definition at this stage was impossible, but that nevertheless a provisional *working definition* should be used to move discussions forward.

Analyzing definitions in disarmament treaties supports this gradual approach: the definitions, which are not always based on the same criteria (effects, functions, usage, or possible targets of the weapons), are generally among the very last issues to be settled. However, the case of LAWS is special, which limits the relevance of such precedents. As Brazil usefully recalled, this time the weapon being defined *does not yet exist*. The weapons previously prohibited had existed for decades and their effects were perfectly known. They could therefore be banned without a precise definition. The absence of shared experience or understanding make LAWS different.

France declared that “it is necessary for States parties to the CCW to work towards a common characterization of a LAWS” so as to ensure that everyone is discussing the same thing, but without making it an absolute precondition for all discussion. France even provided a very precise, and therefore restrictive, definition: **LAWS are mobile systems, capable of adapting to their terrestrial, maritime, or aerial environment, and of autonomously selecting a target and opening fire with lethal munitions, meaning without any human intervention or validation.** For France, **LAWS are fully autonomous weapons systems**, in the sense that there is absolutely no link (of communication or control) with the chain of command. The capacity for self-learning in an evolving environment makes their behavior somewhat unpredictable. However, **LAWS do not exist** and should not be
confused with either remotely operated or supervised weapons systems, which always involve a human operator, or automatic weapons systems, corresponding to the distinction between autonomy and automation highlighted by the experts.

Without a preliminary definition of LAWS, or with an imprecise definition, there is a risk of conflating them with existing weapons. The delegations mostly agreed that LAWS do not yet exist, and hence the first version of the recommendations proposed by the German presidency spoke of a “common understanding”. However, this was contested by Pakistan, as well as the International Committee for Robot Arms Control (ICRAC), which claims that “a number of States are developing them.”

The ICRC is also interested in existing automated defensive systems, not because they are problematic in themselves, but out of concern that they may form the basis for all future development. LAWS will not appear out of nowhere, but as an evolution from pre-existing technologies. However, the subtlety of this evolution means that identifying it may be difficult. For example, the difference between a piloted MQ-9 Reaper drone and an autonomous one will be in the software rather than the casing, and so not immediately visible. In reality, “It would be hard for a country to even know if an adversary had used an autonomous system, as opposed to a remotely piloted system.”

The ICRC therefore seeks firstly the transparency of states, and asks them to explain how human control is exerted over the systems that they already deploy. This is a highly sensitive question, and it is unlikely that states—at least those most advanced in this domain—would agree to share precise information on the functioning of, for example, their anti-missile defenses. The question also risks contaminating the debate at the CCW with strategic issues that stray outside the topic of conventional weapons, to the irritation of the nuclear powers. However, these powers can also use the potential link between LAWS and nuclear weapons to silence the CCW debate when they see fit, as China proved when it claimed that “LAWS are not conventional weapons because they can be used for the full spectrum of weapons, including nuclear, and having an international humanitarian law-centric approach to such a complex issue is not enough.”

Moreover, several states (Egypt, Switzerland, Ireland) claimed—as the ICRC has for a long time—that lethality is contingent and that the CCW should consider autonomous weapons systems in general, even nonlethal ones, so as to include systems that could potentially be used for law enforcement (and whose use does not concern international humanitarian law (IHL) but rather international human rights law). Christof Heyns, UN Special Reporter on Extrajudicial, Summary or Arbitrary Executions, favors a preventive ban of LAWS and has also called for the debate to cover “autonomous weapons” in general.

This insistence on qualifying the criteria of lethality is highly calculated: it is part of a strategy aiming to take the topic outside the CCW, which is purely IHL territory, into other areas potentially more amenable to a preventive ban, like the Human Rights Council. For the opponents, the debate on nonlethality, like that on existing defensive weapons, is a way to make the subject less futuristic and more real. Aware of the slim chances of winning a preventive ban at Geneva, the abolitionist camp is determined to take the subject outside of the CCW. Many states (Algeria, Austria, Brazil, Cuba, Ecuador, Mexico, New Zealand) have recommended changing the first version of Germany’s
proposed recommendations, which called the CCW “the appropriate forum for dealing with issues of LAWS,” to read “an appropriate forum” so as not to exclude others. France, the state that originally brought the subject to the CCW, will have to stay vigilant to this risk of “leakage.”

**What Level of Human Involvement?**

The last UN meeting highlighted an implicit consensus that a certain level of human involvement is desirable, but also clear differences in the terms employed and whether to make human involvement a fundamental condition for lawfulness. Most opponents seek to preventively ban any LAWS that lack “meaningful human control,” an expression that was introduced in previous years and that many presented as a condition for legality and legitimacy of all weapons systems.\(^\text{7}\) The problem is that no one knows exactly what “meaningful” means. France disputes the relevance of the concept, which it finds both vague and in contradiction of full autonomy, which, in its opinion, characterizes LAWS. By definition, a weapons system over which “meaningful”—or “appropriate”, “effective”, “significant”—human control can be exerted would not be autonomous.

The United States, which is also critical of the concept of “meaningful human control,” prefers the concept of an “appropriate level of human judgment,” which it introduced in 2012 in its famous directive 3000.09 (which makes the United States, for four years now, the only country to have a written policy on the subject), according to which “Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force.”\(^\text{8}\) However, an “appropriate level” is still fairly vague and subjective. No more precisely, at the UN meeting in 2016 the United Kingdom spoke of “an evolving intelligent partnership between operators and computers.”

In order to bring everyone to agreement, the German presidency introduced the expression “appropriate human involvement” in its recommendations, which has the benefit of having never been used and thus being more neutral. The final adopted text took equal care not to settle the dispute, making it acceptable to all positions. It simply underlines that “there was a general understanding that ... views on appropriate human involvement with regard to lethal force and the issue of delegation of its use are of critical importance ... and should be the subject of further consideration.”

**Will LAWS Respect the Laws of War?**

All states recognize the applicability of IHL and, with two exceptions (China and India), the obligation and interest in verifying that the envisaged weapons conform with IHL before developing or employing them, as required by Article 36 of the first Additional Protocol to the Geneva Conventions. Like the others, France is therefore committed to developing or employing LAWS only if “these systems demonstrated their full compliance with international law.” Such a determination “is to be made on the basis of normal use of the weapon as anticipated at the time of evaluation”; and “A State is not required to foresee or analyze all possible misuses of a weapon, for almost any weapon can be misused in ways that would be prohibited.”\(^\text{10}\)

This answers the common argument of the risk of LAWS being used by a crazed dictator to massacre his people, or by terrorists to carry out attacks. This is a security risk to take into account (as discussed later), but it does not provide a legal basis for a preventive ban of LAWS. As Hugo Grotius wrote in 1625, “The use that wicked men make of a thing does not always hinder it from
The legal debate has been dominated by the question of weapons reviews. The ICRC submitted a questionnaire to states, inviting them to share their legal review procedure, but not necessarily the results of this monitoring, whose confidential nature the ICRC understands. In a shared concern for transparency, many states joined in by sharing their procedures for the first time (Belgium, Canada, Germany, Israel, Japan, Netherlands, Russia, Sweden, Switzerland, United Kingdom, and United States). The NGOs rejoiced.

However, Gilles Giacca of the ICRC noted that weapon reviews do not solve the issue, and that the context of deployment (when and where) is decisive. Canada equally insisted on the importance of the operational environment (terrestrial, aerial, or maritime) and geopolitical context. Generally, the ICRC presented weapon reviews as necessary but insufficient: they do not supersede the multilateral debate at the CCW.

The debate on LAWS’s capacity to respect IHL inevitably leads to a deadlock as it concerns systems that do not yet exist. Neither those who claim that they will be able to do so nor those who claim the contrary are able to prove anything, and so the discussion peters out. It certainly seems difficult to program principles such as the distinction between civilians and combatants or proportionality, but in this fast-evolving sector what seems difficult today might well become possible tomorrow. For this reason, France refuses to dismiss the possibility that “in certain circumstances, autonomous systems might better respect IHL principles than humans.” As the United States also recalled, “IHL does not prohibit autonomy; on the contrary, in many instances the use of autonomy could enhance respect of IHL.” France concluded that “the development and use of lethal autonomous weapons systems cannot be regarded as intrinsically contrary to IHL. Any preventive prohibition of the development of any potential LAWS would therefore appear premature.”

Moreover, as humans do not respect IHL perfectly (if they did, there would be no war crimes), the question is not so much whether LAWS will be capable of respecting IHL or not, but whether they will be capable of doing so better or worse than humans under the same circumstances. This argument was clearly expressed by law professor Marco Sassòli at the 2014 meeting, where he explained that the system must be required to “be able to recognize wounded, not like God would do, but like a human being would.”

Another way of limiting the risk of violations of IHL is to restrict the context of deployment. LAWS are by definition autonomous after activation, but activating and deploying them in a certain zone remains a human decision. France deduced from this that “commanders and those using the weapon will continue to exercise their judgment about a number of factors, such as the likely presence of civilians and likelihood of their suffering unintentional damage; the expected military advantage; the specific characteristics and conditions of the environment where the system will be deployed; and the weapon’s safety characteristics, capabilities, and limitations”. In other words, the military targeting process (planning, execution and assessment) is one way for humans to remain in relative control of LAWS.

What Ethical Debate?

With the legal debate on the capacity of not-yet-existing weapons to respect IHL quickly coming to
an impasse, the opponents’ motivations are shown to be mostly ethical. They invoke “human dignity,” the “conscience of humanity” in the Martens clause,\textsuperscript{14} and the “morally unacceptable” character of LAWS even when they satisfy the test of legality. Even the ICRC relies on the ethical principles of IHL, such as the “principle of humanity” and “the dictates of public conscience.” The problem, of course, is that these are what Raymond Aron called “vague and grandiose words”\textsuperscript{15}: everyone is impressed when they are uttered, but no one knows exactly what they mean. They are an argumentative dead end, where reason crosses over into belief.

In the “Ethics” panel at the last Geneva conference, which curiously lacked a single philosopher, let alone an ethicist, what was supposed to be an ethical debate merged with the legal one. Ethics is paradoxically both the foundation of the debate and is neglected in the debate. As long as notions like “human dignity” and “conscience of humanity” are invoked without definition or explanation, it will be hard for the discussion to progress. For example, Christof Heyns, who opposes LAWS in the name of the right to life and the right of human dignity (“it is reducing humans to objects”) could have been asked if an indiscriminate bombardment by humans is always preferable—always a lesser violation of human dignity—than a targeted strike by a LAWS, independently of the consequences, meaning even when the former would cause far more collateral damage than the latter.

The deontologist position, which condemns LAWS as \textit{mala in se} because they are disrespectful, (independent of their potentially positive consequences, such as becoming better than humans at discriminating) is widespread and—from my consequentialist perspective at least—seems grounded in nothing more than a belief without reason.\textsuperscript{16} That makes the ethical debate difficult in diplomatic forums such as the CCW, where there is neither time nor competency to discuss at length the ethical postulates of each state and non-state actor.

\textbf{Will LAWS Be “Secure”?}

Is the relative unpredictability of LAWS, which all states recognize since it is intrinsic to autonomy, not a reason in itself to prohibit them? This is the argument of the abolitionists: for them, the risk of a LAWS becoming uncontrollable or committing fatal errors is too great. This is a legitimate source of concern, but it must be put in perspective. On one hand, as some states recalled (Canada, Poland), this risk is not unique to LAWS; it is part of all human action. As Neha Jain from the University of Minnesota raised, “Every human is a wild card, and we accept it each time we deploy a soldier on the battlefield.” Why would the acceptable risk of LAWS be different to that which is accepted of humans?

On the other hand, if it turns out that LAWS are too unpredictable, states can be trusted not to use them—not because they would violate IHL, but simply because they would serve no military purpose. France and the UK share the belief that LAWS might never exist, as \textit{fully} autonomous weapons would have no military utility. As France explained, “For the armed forces, total autonomy and the absence of a link with a human operator which it entails goes against the military command’s need of situation awareness and operational control.” If LAWS are defined as fully autonomous, they will thus be unpredictable and, because “predictability is the very basis for the usefulness and effectiveness of a machine,” “such a weapons system would be militarily useless.” Thus, states will only use autonomous weapons to the extent that they consider themselves to be able to predict their behavior with a sufficient degree of certainty.
The debate has also covered classic security risks, such as of lowering the threshold of the use of force (mentioned by Afghanistan and Pakistan), which remains to be proved as while the physical risk is lower, the economic and strategic risks remain; of arms races (India, ICRAC); and of reinforcing asymmetry, which would compel weak states lacking LAWS to revert to nonconventional means. However, none of these concerns invoke a precise scenario, nor are they specific to LAWS.

Delegations also cited more specific risks, like the system becoming uncontrollable in a chain reaction. Paul Scharre of the Center for a New American Security gave the example of the 1983 nuclear close call, where human judgment avoided escalation, to illustrate the risk of escalation and hence strategic instability that could be caused by automation. The hijacking of autonomous systems, even nonlethal ones, to commit attacks was also mentioned. Daesh (the Arabic term for the so-called “Islamic State”) is already working on remote-controlled car bombs, but it is a long way from developing autonomous technologies.

The most serious risk is undoubtedly that of “swarming”: the possible use of a great number of coordinated LAWS to saturate an adversary. “There cannot be meaningful human control in this kind of swarm,” claimed the ICRAC. “How can we face tens of thousands of LAWS on the battlefield?” asked China. To which Dan Saxon of Leiden University responded, “We cannot. No human or machine interface would be able to face such a threat.”

The “swarm” is undoubtedly the greatest security challenge that can be conceived, but no doubt there are still others that have not yet been thought of. As John Borrie of the United Nations Institute for Disarmament Research confirmed, “We are discussing known unknowns, but what I am more worried about are unknown unknowns, because they are the products of interactions we cannot foresee.”

**Toward a Code of Conduct**

If the upcoming December Review Conference applies the recommendations adopted in April, a formal process will begin in 2017. Russia has already advised that it does not want anything beyond a “discussion” and that if the process becomes too formal (meaning that if it would negotiate a somewhat binding instrument), it would not exclude withdrawing. The reluctance of states should not be reduced to a Machiavellian maneuver to continue secretly developing LAWS on the back of unending multilateral discussions: it also expresses the sincere belief that it is too early to take a definitive decision such as a moratorium on weapons that do not yet exist and whose potentially positive consequences are ignored. This is just prudence.

Numerous states oppose the potential deployment of fully autonomous weapons without going so far as to advocate their preventive ban. As the Netherlands explains, a moratorium could have unintended negative effects, slowing technological progress that could potentially benefit the civilian domain without effectively impeding the development of such weapons. On the other hand, civilian progress in the field of artificial intelligence and self-learning may anyway put these technologies on the market and therefore into military hands, making such a ban ineffective.

For these reasons, the pool of states susceptible to rallying the abolitionist camp will shrink. The “Campaign to Stop Killer Robots” was delighted at the conclusion of the meeting to count 14 states favorable to the preventive ban of LAWS (Algeria, Bolivia, Chile, Costa Rica, Cuba, Ecuador, Egypt,
Ghana, The Holy See, Mexico, Nicaragua, Pakistan, Palestine, and Zimbabwe). However, none of them has or will foreseeably have the capacity to produce LAWS by themselves, and the most powerful states that have this capacity are precisely those that oppose a moratorium. Because of technological democratization there will be more and more “capable states.” In the absence of deployment, and thus proof of the problematic nature of these weapons, this means more and more states likely to become hostile to the idea of a pure and simple ban.

Therefore, the work of the Governmental Group of Experts (GGE) is not likely to deliver a moratorium. However, they cannot (and should not) unveil nothing, since it would call into question the value of having even discussed the subject for many years at the CCW. The most probable and desirable outcome is that the GGE finishes with what could be called a *Geneva Document on pertinent legal obligations and good practices for states related to lethal autonomous weapons systems*, a non-binding text comparable to the *Montreux Document* for private military and security companies. This “code of conduct” could develop the following conditions:

- Only use LAWS against “by nature” military objectives (meaning not against those that become so by their location, purpose, or use).
- Only use them in certain contexts (maritime, spatial, aerial, or desert, but certainly not urban).
- Make the activation of the autonomous mode reversible (humans should retain the ability to deactivate the LAWS’s firing function, to cancel or halt an attack).
- Delimit the duration and location of their operations.
- Program doubt: Faced with an unforeseen event, the LAWS would stop and consult its commander, applying the rule “if in doubt, do not fire.”
- Record LAWS’s actions.
- Train their “operators” (launchers, commanders) in IHL.
- Only use LAWS in situations where humans cannot take the decision themselves, for example, due to a lack of time or a communication breakdown (principle of subsidiarity).

Because of the fast-evolving civilian research in artificial intelligence, its connection with military application, and the reluctance of powerful states to give up on a technology potentially useful, a preventive ban, and more generally the implementation of an arms-control regime is illusory. The question is no longer about how to stop an AI arms race already under way, but about how to manage it. In such attempt, the coming debate on a code of conduct, while certainly insufficient, can help elaborate a normative framework.

*The views expressed in this article are the author’s own and do not represent those of any institution to which he is or was affiliated. A previous version of this essay was published in French as “Diplomatie des armes autonomes : les débats de Genève,” Politique Etrangère, 3/2016, p. 119-130. Translated with the help of Andreas Capstack. Reproduced with permission of the copyright owner.*

2. When no source is given, citations in this article are drawn from public exchanges at the meeting, in which the author participated.

3. The list of participants, certain states’ working papers and declarations, the experts’ presentations and the president’s report can all be found on the UN Geneva office’s website: www.unog.ch/80256EE600585943/ (http Pages)/37D51189AC4FB6E1C1257F4D004CAFB2. The citations from France in this article are all drawn from its official statements: www.delegfrance-cdgeneve.org/Meeting-of-CCW-experts-on-lethal-autonomous-weapons-systems-Geneva-11-15-of

4. The text can be downloaded from


7. V. Kozyulin (PIR Center, Moscow) also declared that LAWS “can be used for conventional, nuclear or even chemical munitions: therefore, it is especially dangerous. Autonomous weapons are more dangerous than nuclear.”


12. This argument is known in the literature as the “Arkin test,” named after the American roboticist Ronald C. Arkin. The principle is that a robot satisfies the legal and moral requirement- and can consequently be deployed- when it is demonstrated that it can respect the law of armed conflicts not perfectly but as well or better than a human in similar circumstances. For example see G. R. Lucas, “Automated Warfare”, Stanford Law & Policy Review, 25:2, 2014, p. 322, 326 and 336.


14. Preamble to the 1899 Second Hague Convention, from the name of the Russian delegate who issued the declaration at the conference.


17. On 26 September 1983, the Soviet radars reported the launch of five American ballistic missiles. Despite the fact it happened in a context of exacerbated tensions between the US and the USSR (the alarm coincided with the beginning of a massive NATO exercise and only three weeks after Moscow shot down a South Korean airliner), the officer in charge of the soviet early-warning satellite system that day, Lt. Col. Stanislav Petrov, had the feeling it was a false alarm, and reported it as such. It was later found that the soviet satellite mistakenly interpreted the sun’s reflection off the cloud as a missile launch.

18. Netherlands opening statement, 11 April 2016:


21. See the definition of military objectives in Article 52(2) of Additional Protocol I to the Geneva Conventions.

22. That implies a permanent connection between human and machine which is counter-intuitive, since one of the points of autonomy is precisely to be able to maintain activity in case of interference or a breakdown in communication. Therefore, I assume here that if LAWS exist, they will not be fully autonomous – because full autonomy entails unpredictability, so military uselessness. They should – and probably will – be so-called “human-on-the-loop” weapons, not “out of the loop.”