Will the Geneva Convention Cover Robots?

Thomas McMullan

When Dr. Richard J. Gatling designed his gun, it was meant to save lives. The American inventor designed the rapid-fire, spring-loaded, hand-cranked weapon with the express purpose of reducing the number of people needed on the battlefields of the American Civil War, thereby preventing deaths. Instead, he unleashed a forerunner of the machine gun that would scale the level of killing by several orders of magnitude, leading eventually to the horrific suffering in the trenches of the First World War.

Is history repeating itself with the development and application of artificial intelligence in warfare? Pressure has been steadily building on governments to address the nascent field of autonomous weapons; a nebulous term, but one largely agreed to involve systems capable of killing without human intervention. Could A.I.-directed precision lead to fewer civilian casualties, and ultimately less need for human soldiers? Or will it, like the Gatling gun, herald a new scale of slaughter?

The past few months alone have seen reports of a secret lethal drone project in the U.K. and an A.I. weapons development program for teenagers in China. Meanwhile in the U.S., Google found itself in hot water with employees after it was found to be helping the Pentagon with a drone A.I. imaging project. Last year, Russia's Vladimir Putin said that whichever nation leads in A.I., "will be the ruler of the world." The question at the moment isn't whether autonomous weapons are on their way, but what shape, if any, international regulation should take to control them.

A MQ-9 Reaper unmanned aerial vehicle prepares to land after a mission. The

Reaper has the ability to carry both precision-guided bombs and air-to-ground missiles. Credit: USAF Photographic Archives

"Everyone agrees that in all the corpus of existing international law, there's nothing that specifically prohibits autonomous weapons," Paul Scharre, director of the Technology and National Security Program at the Center for a New American Security, tells me. "There are no stand-alone treaties, as there are for many other weapons. Everyone also agrees that the existing laws of war apply to autonomous weapons. Whatever the laws of war are now, autonomous weapons have to be used in ways that comply with them."

> As it currently stands, if a self-thinking robot were to roll onto a battlefield, the same conventions around war would apply to it as they would to any human soldier.

That, however, is where the agreement stops. Scharre tells me the range of views on A.I. and war can be roughly split into three camps. On one extreme you have the Campaign to Stop Killer Robots, a coalition of over 85 nongovernmental organizations pushing for a preemptive ban on fully autonomous weapons. Its thinking is backed by a handful of different countries, including Austria, Brazil, and the Vatican. Some robotics firms, including Clearpath Robotics, Hanson Robotics, and Alphabet subsidiary DeepMind, have also separately supported a ban.

On the other extreme you have Russia and the U.S., both of which in September moved to block UN talks moving forward on such a preemptive ban. These nations are happy with the laws of war, thank you very much, and don't want any regulation tripping up research projects that may or may not be happening behind the scenes. In between these two poles you have nations such as France and Germany, which have led the charge for a politically binding resolution; a nice sounding statement about autonomous weapons, but not necessarily a legally binding treaty.

So where does that leave us? As it currently stands, if a self-thinking robot were to roll onto a battlefield, the same conventions around war would apply to it as they would to any human soldier.

"Take for example an autonomous medic," explains Professor Peter Roberts, director of military sciences at British defense think tank Royal United Services Institute (RUSI). "Say it has the ability to go out, find a soldier that's in trouble, and treat them. That robot has the exact same rights under the Geneva Conventions as a human would. It cannot be targeted by a foreign power. It cannot be upset in achieving the course of its work, unless it picked up a weapon. It retains the same protections humans would have."

> The concept of autonomy is also less clear-cut than it first seems. Where do the limits of human control begin and end?

But these machines aren't humans. "Imagine the consequences of an autonomous system with the capacity to locate and attack, by itself, human beings," said UN Secretary-General António Guterres at the recent Paris Peace Forum. Guterres said such a weapon would be "morally revolting," and called on heads of state to preemptively ban such systems from existing.

One way to achieve that may be the creation of a specific treaty. The Convention on Cluster Munitions (CCM), for example, prevents the stockpiling of cluster bombs, while the Ottawa Treaty controls the use of antipersonnel landmines. "But in those cases there was clear humanitarian harm," says Scharre. "While states were at the diplomat table in Geneva, there were people being maimed by landmines and cluster munitions. There was a very strong collection of instances to put pressure on the international community."

The problem with A.I.-controlled weapons is that much of what is being spoken about is theoretical. The concept of autonomy is also less clear-cut than it first seems. Where do the limits of human control begin and end? Think about a self-driving car. Conceptually, it's simple — a car that drives itself — but

in practice it's a grey area, with autonomous features such as intelligent cruise control, automatic lane keeping, and self-parking already in existence. Similarly, aspects of automation have been creeping into military systems for decades, arguably since the invention of the Gatling gun, and now this trend is being taken to new levels with machine learning and advanced robotics.
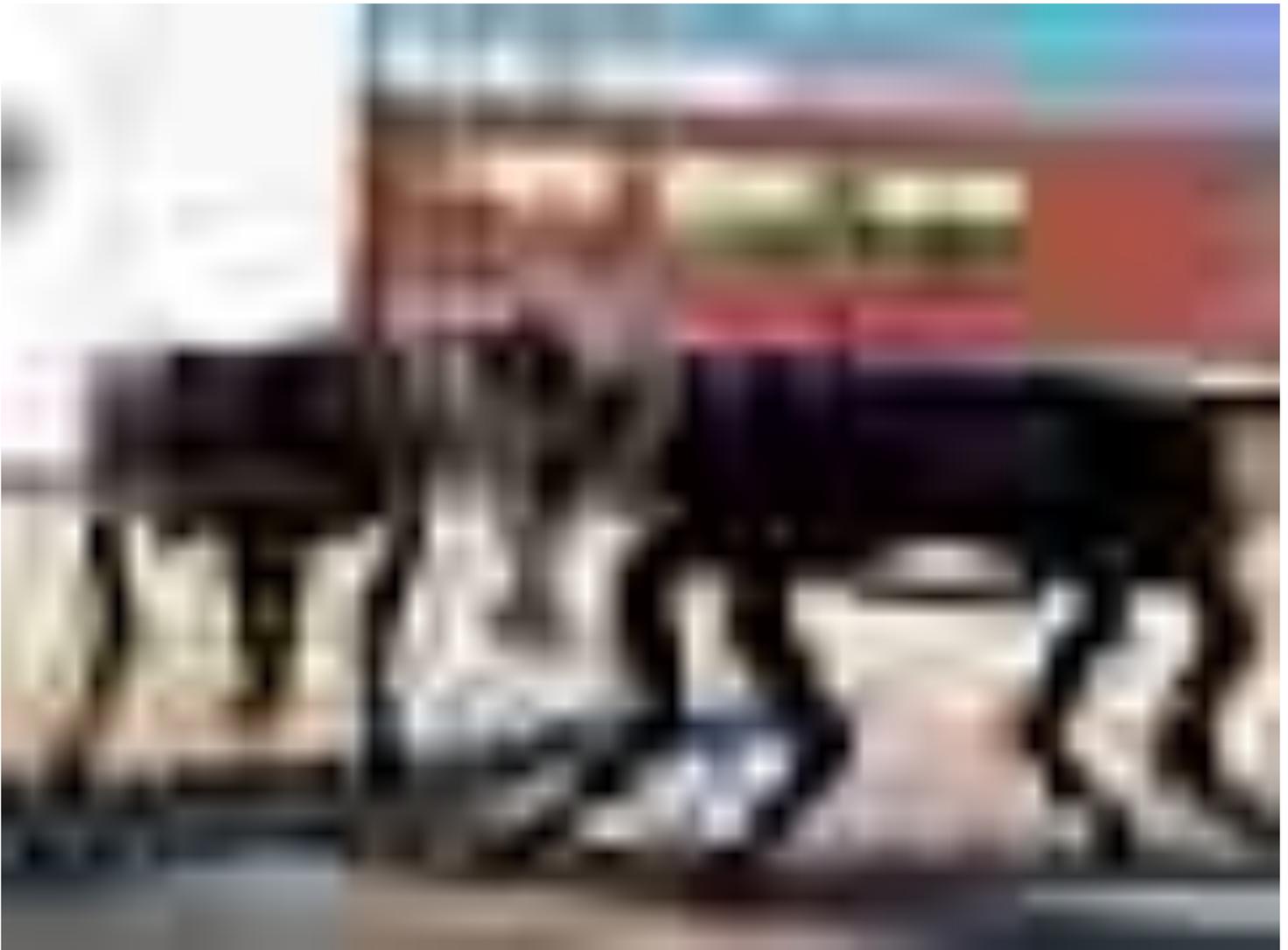
BigDog robots trot around in the shadow of an MV-22 Osprey. The BigDog is a dynamically stable quadruped robot created in 2005 by Boston Dynamics with Foster Miller, the Jet Propulsion Laboratory, and the Harvard University Concord Field Station. Credit: U.S. Marine Corps

Seen from this perspective, artificial intelligence isn't a single, discreet technology like a landmine or a nuclear bomb, or even like an airplane or a tank. Rather than a weapon to be banned, it's a force that's bringing about a deeper shift in the character of war.

"People compare it to more general purpose technology, like electricity or the internal combustion engine," says Scharre. "The best historical analogy is the process of mechanization that occurred during the Industrial Revolution."

Earlier this year, Google announced it would not be renewing a contract with the U.S. military. The contract was for a program to use A.I. for automatically analyzing drone footage, dubbed Project Maven, and the company's decision came after heavy internal backlash, with dozens of resignations and a petition signed by thousands of employees.

The crux of Project Maven is this: the scale of footage recorded every moment from U.S. drones has reached a point where it is too great for human eyeballs to pore through manually, and so A.I. systems are being developed to automatically flag relevant moments. Much like an internet moderator, a human

operator would then make the call on whether there is a target, and whether that target should be "moderated," as it were.

> Perhaps the ethical line is only reached when you allow that A.I. system to not only flag footage, but to make a lethal decision.

Crucially, by using image recognition to sift through the masses of footage collected by drones, the project positions A.I. not as a superweapon, but as a fundamental part of how emerging wars will be fought. "These programs also illustrate a wonderful irony," says Peter Singer, strategist and senior fellow at the New America think tank. "Our ever-increasing use of technology has yielded vastly greater amounts of data in speeds that are hard for humans to keep up. So, in turn, we seek out new technology to keep pace."

Singer says despite Google distancing itself from the project, he doesn't see work slowing down anytime soon. There is an almost insatiable demand for a system that can process the masses of footage gathered every second by drones. Automatic filtering is a practical solution to a growing problem, it might be argued. Perhaps the ethical line is only reached when you allow that A.I. system to not only flag footage, but to make a lethal decision.

Even before you get to this point, however, there are questions about bringing algorithms into the battlefield. Not only can image recognition flag targets, but the realities of the battlefield mean it is beneficial for drones to have a degree of autonomy to prevent them from being shot down, in case signals are cut. What does this creeping autonomy do to how we think about war? Does it encourage a way of thinking about conflict in a similar scale to the internet, remote from bodies, unable to be traversed alone by human judgement?

The shifting character of war is a complex thing to develop international conventions around. RUSI's Roberts notes that, while an international treaty prohibiting autonomous weapons systems is possible, the critical thing will be the small print; what is specifically banned, who ratifies it, what caveats they apply, and who doesn't apply it. Even with a treaty in place, the push for A.I. warfare will likely continue behind the scenes, for the simple reason that nations don't want to risk being left behind.

> It's relatively easy to get behind a ban on *Terminator*-style robo-soldiers, but much harder is agreeing where to draw lines around the varying levels of autonomy that are creeping into military systems.

"What's happening under the radar is so important, because if your adversary might be a signatory but you know is secretly developing these systems, you cannot stop looking at responsive development yourself, if you are to survive," Roberts said. "This is a not a choice you can make. It's something that must be conducted, if you wish to survive."

As the A.I. arms race heats up, and calls for a preemptive ban on autonomous weapons grow louder, this is a crucial period in deciding how the grey area of A.I. warfare is to be regulated. It's relatively easy to get behind a ban on *Terminator*-style robo-soldiers, but much harder is agreeing where to draw lines around the varying levels of autonomy that are creeping into military systems. The limits of human control are nebulous, but they need to be pinned down.

New weapons can look advantageous, and potentially humane, when you're the only side that has them. The Gatling gun might have been intended to reduce the number of soldiers on a battlefield, but when both sides carry automatic firepower — as with the machine guns of the First World War — the situation looks altogether different. War is, ultimately, a competitive contest.

Update: *This story has been updated to reflect the latest number of signatories and countries that support the Campaign to Stop Killer Robots, and include mention of several robotics firms separately involved in suggesting an A.I. weapons ban.*