

'Autonomous weapons' based on artificial intelligence could change warfare—and why that's worrisome

In the 1970 science fiction film “Colossus: The Forbin Project,” the United States decides to turn over control of its strategic arsenal to Colossus, a massive supercomputer. Big mistake. Almost immediately it becomes clear that, as its creator Dr. Charles Forbin says, “Colossus is built even better than we thought.” In fact, it’s a self-aware artificial intelligence — quickly discovering that the Soviets have also activated an almost identical system and joining up with it to take over the planet. Along the way, Colossus nukes a Russian oil complex and a U.S missile base to enforce its control. Now, instead of two human superpowers threatening nuclear Armageddon, humanity’s continued survival is at the mercy (or mercy’s AI equivalent) of a supercomputer.

8-13-2018 6:00 cover1

“Army of None:

Autonomous Weapons
and the Future of War,”

by Paul Scharre (W.W.
Norton, 448 pages).

“The object in constructing me was to prevent war,” Colossus announces. “This object is attained. I will not permit war. It is wasteful and pointless. ... Man is his own worst enemy. ... I will restrain man.” To its cool machine reasoning, it’s all perfectly rational. But its definition of rationality differs tragically from that of human beings.

We’re in no danger of a Colossus taking over the planet, at least not yet. But the prospect of lethal autonomous weapons (AWs) under nonhuman control is all too real and immediate. As Paul Scharre points out in “Army of None: Autonomous Weapons and the Future of War,” we already have robots doing everything from cleaning the living room to driving cars to tracking down (and sometimes taking out) terrorists. The step from armed drones controlled remotely by humans to fully autonomous machines that can find, target, and kill all on their own is less a matter of technology than our own choice: Do we turn on Colossus or not?

Such questions used to be strictly the province of science fiction, fantasy, and legend, from the golems of Jewish culture to Mary Shelley’s Frankenstein to the robots of Karel Čapek, Isaac Asimov, and the “Terminator” series. “I wonder if James Cameron had not made the ‘Terminator’ movies how debates on autonomous weapons would be different,” notes Scharre. “If science fiction had not primed us with visions of killer robots set to extinguish humanity, would we fear autonomous lethal machines?” Possibly not. But in the 21st century, contemplating the morality and advisability of creating artificial agents capable of independent deadly action has swiftly moved from an intellectual diversion to an imminent concern.

Scharre himself is a front-line veteran not only of the halls of Washington and the Pentagon as a consultant and policymaker, but of four combat tours in Iraq and Afghanistan as a U.S. Army Ranger. When he speaks of military ethics, decision-making, and killing, it’s with the authority of a man who’s been

there, someone who's zeroed other human beings in the sights of a rifle and faced the decision whether or not to pull the trigger — the same decision some now want to delegate to machines.

"Humanity is at the threshold of a new technology that could fundamentally change our relationship with war," he writes. Ever since the invention of the bow and arrow, technology has dictated the direction of warfare, making it possible to destroy and kill more efficiently and at greater distances. Automation first became a factor in the American Civil War with the invention of the Gatling gun, followed by the devastating machine guns of World War I. But even if the weapons operated more or less automatically, human beings were still pulling the triggers.

Now we're approaching a new era in which human control, agency, and ethical decision-making could be superfluous. As Scharre demonstrates, the technology of fully autonomous weapons is advancing rapidly. But while the evolution of the technology is inevitable, using it is not.

Whether we call them robots, drones, semiautonomous weapons systems, or some other fancy Pentagonese term, it was inevitable that once the technology existed, it would be adapted for military use. "No one planned on a robotics revolution, but the U.S. military stumbled into one as it deployed thousands of air and ground robots to meet urgent needs in Iraq and Afghanistan," Scharre says. Such devices certainly weren't new; primitive drones existed [as far back as World War I](#), and guided missiles had achieved extreme sophistication and accuracy by the 1960s. But as they became ever cheaper and more versatile, their use also became easier and more reasonable: Why risk human soldiers on risky recon patrols or bomb disposal when remotely controlled robots could do the same? "Unshackled from the physiological limits of humans," Scharre points out, such machines "can be made smaller, lighter, faster, and more maneuverable. They can stay out on the battlefield far beyond the limits of human endurance, for weeks, months, or even years at a time without rest. They can take more risk, opening up tactical opportunities for dangerous or even suicidal missions without risking human lives." Computerized systems can also handle multiple threats in chaotic combat situations moving too fast for humans to handle.

It's still true that with few exceptions, machines such as drones must be controlled from afar by human operators, requiring communication links that can be disrupted or jammed, rendering the drone or robot essentially useless. Hence the next step in robotic evolution: full autonomy — not merely to enable passive observation over enemy territory, but to find the enemy and destroy it. To do that, it's necessary (as the strategists say with cool detachment) to "delegate lethal authority."

Follow the latest news and policy debates on sustainable agriculture, biomedicine, and other 'disruptive' innovations. Subscribe to our newsletter.

[SIGN UP](#)

Scharre describes a few such systems that already exist, such as the Israeli Harpy drone, designed to loiter over hostile territory and destroy any enemy radar it detects. As far back as the 1980s, the U.S. Navy developed the Tomahawk Anti-Ship Missile (TASM), which could be fired from a ship to a remote area to automatically seek and destroy enemy vessels. Though never actually used in such a way, it was technically the world's first fully operational AW.

Like most of the other questions now facing us regarding AI technologies (Do you want to trust your life to a self-driving car? How much of your personal life do you want Alexa to overhear?), the issues surrounding AWs are fraught with complexity and complications, but on a much more profound level, penetrating to the heart of human morality and ethics. If we give machines the power of life and death, who's responsible for their victims? How can we ensure that they'll make the same distinctions between hostile insurgents and innocent civilians, judgment calls that are sometimes impossible to make rationally—and what if they're wrong?

Scharre gives examples from his own combat experience dealing with such challenges, and how his own humanity and judgment affected his decisions. But we can't be sure that autonomous weapons will behave similarly. As one AI researcher tells Scharre, "It's almost certain that as AI becomes more complicated, we'll understand it less and less." What seems like an eminently reasonable decision to the inscrutable algorithms controlling an AW may be morally abhorrent to human beings. Colossus, Alexa, and your Roomba don't think the way we do — their intelligence is different from that of humans.

[ai 8 14 18 3](#)

Image not found or type unknown

And it doesn't take fully autonomous weapons to create such problems. Scharre recounts various examples of automated and semi-automated weapons, all with human controllers ostensibly "on or in the loop" (a vital distinction that he also explains), that have nonetheless caused tragedies, including the Patriot missile systems that downed friendly aircraft in the Iraq war.

An obvious measure is to negotiate and decide upon some kind of mutually accepted restrictions on the development and deployment of AWs under the authority and guidance of international law, much as has

been attempted with other weapons in the past. But that's far from a perfect solution, Scharre makes clear. Even when the nations of the world decide that a particular technology is simply too horrible to use — for example, poison gas or germ warfare — it takes just one terrorist or ruthless dictator to upset the applecart. The U.S. and other nations can set their noble standards and refrain from building “inhumane” weapons, but what if other countries don't go along? Scharre provides a lengthy table of successful and unsuccessful international weapons bans, from poisoned arrows and crossbows to aerial bombardment and submarines to land mines and cluster bombs. It's not an encouraging record.

Scharre provides possibilities but no firm solutions to the issues surrounding autonomous weapons, because as he admits, “there are no easy answers.” Yet he also offers some hope. There's still time to consider, to question, and to decide on restraint and caution, whatever form it may take and however imperfect it may be. “The technology to enable machines that can take life on their own, without human judgment or decision-making, is upon us,” he says. “What we do with that technology is up to us.”

In the meantime, those who are contemplating, designing, or dreaming of autonomous weapons would do well to heed some advice from Dr. Forbin, the creator of Colossus: “I think ‘Frankenstein’ ought to be required reading.”

Mark Wolverton is a science writer, author, and playwright whose articles have appeared in Undark, Wired, Scientific America, Popular Science, Air & Space Smithsonian, and American Heritage, among other publications. His forthcoming book “Burning the Sky: Operation Argus and the Untold Story of the Cold War Nuclear Tests in Outer Space” will be published in November. In 2016-17, he was a Knight Science Journalism fellow at MIT.

A version of this article was originally published on Undark's website as [“The General Is a Robot: Artificial Intelligence Goes to War”](#) and has been republished here with permission.