Robot Weapons: What's the Harm?

By Jerry Kaplan, *The New York Times*, August 17, 2015

LAST month over a thousand scientists and tech-world luminaries, including Elon Musk, Stephen Hawking and Steve Wozniak, released an open letter calling for a global ban on offensive "autonomous" weapons like drones, which can identify and attack targets without having to rely on a human to make a decision.

5 The letter, which warned that such weapons could set off a destabilizing global arms race, taps into a growing fear among experts and the public that artificial intelligence could easily slip out of humanity's control—much of the subsequent coverage online was illustrated with screen shots from the *Terminator* films.

The specter of autonomous weapons may evoke images of killer robots, but most applications are 10 likely to be decidedly more pedestrian. Indeed, while there are certainly risks involved, the potential benefits of artificial intelligence on the battlefield—to soldiers, civilians and global stability—are also significant.

The authors of the letter liken A.I.-based weapons to chemical and biological munitions, space-based nuclear missiles and blinding lasers. But this comparison doesn't stand up under scrutiny. 15 However high-tech those systems are in design, in their application they are "dumb"—and, particularly in the case of chemical and biological weapons, impossible to control once deployed.

A.I.-based weapons, in contrast, offer the possibility of selectively sparing the lives of noncombatants, limiting their use to precise geographical boundaries or times, or ceasing operation upon command (or the lack of a command to continue).

20 Consider the lowly land mine. Those horrific and indiscriminate weapons detonate when stepped on, causing injury, death or damage to anyone or anything that happens upon them. They make a simple-minded "decision" whether to detonate by sensing their environment—and often continue to do so, long after the fighting has stopped.

Now imagine such a weapon enhanced by an A.I. technology less sophisticated than what is 25 found in most smartphones. An inexpensive camera, in conjunction with other sensors, could discriminate among adults, children and animals; observe whether a person in its vicinity is wearing a uniform or carrying a weapon; or target only military vehicles, instead of civilian cars.

This would be a substantial improvement over the current state of the art, yet such a device would qualify as an offensive autonomous weapon of the sort the open letter proposes to ban.

30 Then there's the question of whether a machine—say, an A.I.-enabled helicopter drone—might be more effective than a human at making targeting decisions. In the heat of battle, a soldier may be tempted to return fire indiscriminately, in part to save his or her own life. By contrast, a machine won't grow impatient or scared, be swayed by prejudice or hate, willfully ignore orders or be motivated by an instinct for self-preservation.

35 Indeed, many A.I. researchers argue for speedy deployment of self-driving cars on similar grounds: vigilant electronics may save lives currently lost because of poor split-second decisions made by humans. How many soldiers in the field might die waiting for the person exercising "meaningful human control" to approve an action that a computer could initiate instantly?

Neither human nor machine is perfect, but as the philosopher B. J. Strawser has recently argued, 40 leaders who send soldiers into war "have a duty to protect an agent engaged in a justified act from harm to the greatest extent possible, so long as that protection does not interfere with the agent's ability to act justly." In other words, if an A.I. weapons system can get a dangerous job done in the place of a human, we have a moral obligation to use it.

Jerry Kaplan, who teaches about the ethics and impact of artificial intelligence at Stanford, is the author of Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence.