LOSING HUMANITY

The Case against Killer Robots

HUMAN RIGHTS WATCH

IHRC | INTERNATIONAL HUMAN RIGHTS CLINIC
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL

Losing Humanity

The Case against Killer Robots

Human Rights Watch is dedicated to protecting the human rights of people around the world. We stand with victims and activists to prevent discrimination, to uphold political freedom, to protect people from inhumane conduct in wartime, and to bring offenders to justice. We investigate and expose human rights violations and hold abusers accountable. We challenge governments and those who hold power to end abusive practices and respect international human rights law. We enlist the public and the international community to support the cause of human rights for all.

Human Rights Watch is an international organization with staff in more than 40 countries, and offices in Amsterdam, Beirut, Berlin, Brussels, Chicago, Geneva, Goma, Johannesburg, London, Los Angeles, Moscow, Nairobi, New York, Paris, San Francisco, Tokyo, Toronto, Tunis, Washington DC, and Zurich.

For more information, please visit our website: http://www.hrw.org

The International Human Rights Clinic (IHRC) at Harvard Law School seeks to protect and promote human rights and international humanitarian law through documentation; legal, factual, and strategic analysis; litigation before national, regional, and international bodies; treaty negotiations; and policy and advocacy initiatives. IHRC also critically examines the human rights movement and engages in innovative clinical education to develop advanced practice techniques and approaches to human rights advocacy. IHRC collaborates with leading international and local human rights organizations and bridges theory with practice at the law school while also advancing the interests of clients and affected communities around the world.

For more information, please visit IHRC's website:
http://www.law.harvard.edu/programs/hrp/ihrc.html

Losing Humanity
The Case against Killer Robots

Summary

With the rapid development and proliferation of robotic weapons, machines are starting to take the place of humans on the battlefield. Some military and robotics experts have predicted that "killer robots"—fully autonomous weapons that could select and engage targets without human intervention—could be developed within 20 to 30 years. At present, military officials generally say that humans will retain some level of supervision over decisions to use lethal force, but their statements often leave open the possibility that robots could one day have the ability to make such choices on their own power. Human Rights Watch and Harvard Law School's International Human Rights Clinic (IHRC) believe that such revolutionary weapons would not be consistent with international humanitarian law and would increase the risk of death or injury to civilians during armed conflict. A preemptive prohibition on their development and use is needed.

A relatively small community of specialists has hotly debated the benefits and dangers of fully autonomous weapons. Military personnel, scientists, ethicists, philosophers, and lawyers have contributed to the discussion. They have evaluated autonomous weapons from a range of perspectives, including military utility, cost, politics, and the ethics of delegating life-and-death decisions to a machine. According to Philip Alston, then UN special rapporteur on extrajudicial, summary or arbitrary executions, however, "the rapid growth of these technologies, especially those with lethal capacities and those with decreased levels of human control, raise serious concerns that have been almost entirely unexamined by human rights or humanitarian actors."[1] It is time for the broader public to consider the potential advantages and threats of fully autonomous weapons.

The primary concern of Human Rights Watch and IHRC is the impact fully autonomous weapons would have on the protection of civilians during times of war. This report analyzes whether the technology would comply with international humanitarian law and preserve other checks on the killing of civilians. It finds that fully autonomous weapons would not only be unable to meet legal standards but would also undermine essential non-legal

---

[1] Interim Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Philip Alston, UN Doc. A/65/321, August 23, 2010, http://documents.un.org/mother.asp (accessed September 30, 2012), p.10.

safeguards for civilians. Our research and analysis strongly conclude that fully autonomous weapons should be banned and that governments should urgently pursue that end.

Definitions and Technology

Although experts debate the precise definition, robots are essentially machines that have the power to sense and act based on how they are programmed.[2] They all possess some degree of autonomy, which means the ability of a machine to operate without human supervision. The exact level of autonomy can vary greatly. Robotic weapons, which are unmanned, are often divided into three categories based on the amount of human involvement in their actions:

- **Human-*in*-the-Loop Weapons**: Robots that can select targets and deliver force only with a human command;
- **Human-*on*-the-Loop Weapons**: Robots that can select targets and deliver force under the oversight of a human operator who can override the robots' actions; and
- **Human-*out-of*-the-Loop Weapons**: Robots that are capable of selecting targets and delivering force without any human input or interaction.

In this report, the terms "robot" and "robotic weapons" encompass all three types of unmanned weapons, in other words everything from remote-controlled drones to weapons with complete autonomy. The term "fully autonomous weapon" refers to both out-of-the-loop weapons and those that allow a human on the loop, but that are effectively out-of-the-loop weapons because the supervision is so limited.[3] A range of

---

[2] Human Rights Watch telephone interview with Noel Sharkey, professor of artificial intelligence and robotics, University of Sheffield, Sheffield (UK), September 6, 2012. Others have described robots as machines that can sense, think, and act. See, for example, Gianmarco Veruggio and Keith Abney, "Roboethics: The Applied Ethics for a New Science," in Patrick Lin, Keith Abney, and George A. Bekey, eds., *Robot Ethics: The Ethical and Social Implications of Robotics* (Cambridge, MA: MIT Press, 2012), p.349; P.W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century* (New York: The Penguin Press, 2009), p. 67. "Think" does not mean to imply processing information in the same sophisticated way as humans. Instead "think" refers to processing "if…, then…" commands. Human Rights Watch telephone interview with Noel Sharkey, September 6, 2012.

[3] Major Jeffrey Thurnher, a US Army lawyer, notes the importance of a meaningful override. While a proponent of what he refers to as "lethal autonomous robots," he writes that such robots "should be required to have some version of a human override" and that "[t]his oversight would not be effective if the human operator were merely a rubber stamp to approve an engagement." Jeffrey S. Thurnher, "No One at the Controls: Legal Implications of Fully Autonomous Targeting," *Joint Forces Quarterly*, issue 67 (Fourth Quarter 2012), p. 83.

other terms have been used to describe fully autonomous weapons, including "lethal autonomous robots" and "killer robots."[4]

Fully autonomous weapons, which are the focus of this report, do not yet exist, but technology is moving in the direction of their development and precursors are already in use. Many countries employ weapons defense systems that are programmed to respond automatically to threats from incoming munitions. Other precursors to fully autonomous weapons, either deployed or in development, have antipersonnel functions and are in some cases designed to be mobile and offensive weapons. Militaries value these weapons because they require less manpower, reduce the risks to their own soldiers, and can expedite response time. The examples described in this report show that a number of countries, most notably the United States, are coming close to producing the technology to make complete autonomy for robots a reality and have a strong interest in achieving this goal.

Safeguards for Civilian Protection

According to international law and best practices, states should evaluate new or modified weapons to ensure they do not violate the provisions of international humanitarian law, also called the laws of war.[5] States should conduct weapons reviews at the earliest stages of development and continue them up through any production decision. Given military plans to move toward increasing autonomy for robots, states should now undertake formal assessments of the impacts of proposed fully autonomous weapons and technology that could lead to them even if not yet weaponized.

As this report shows, robots with complete autonomy would be incapable of meeting international humanitarian law standards. The rules of distinction, proportionality, and military necessity are especially important tools for protecting civilians from the effects of war, and fully autonomous weapons would not be able to abide by those rules. Roboticists have proposed different mechanisms to promote autonomous weapons' compliance with these rules; options include developing an ability to process quantitative algorithms to

---

[4] Human Rights Watch telephone interview with Noel Sharkey, September 6, 2012. See also generally Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Surrey, UK: Ashgate Publishing Limited, 2009). Due to different definitions and understandings, these terms do not necessarily mean the exact same thing to various experts.
[5] Article 36 of Additional Protocol I to the Geneva Conventions codifies this rule. Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflicts (Protocol I), adopted June 8, 1977, 1125 U.N.T.S. 3, entered into force December 7, 1978, art. 36.

analyze combat situations and "strong artificial intelligence (AI)," which would try to mimic human thought. But even with such compliance mechanisms, fully autonomous weapons would lack the human qualities necessary to meet the rules of international humanitarian law. These rules can be complex and entail subjective decision making, and their observance often requires human judgment. For example, distinguishing between a fearful civilian and a threatening enemy combatant requires a soldier to understand the intentions behind a human's actions, something a robot could not do. In addition, fully autonomous weapons would likely contravene the Martens Clause, which prohibits weapons that run counter to the "dictates of public conscience."

By eliminating human involvement in the decision to use lethal force in armed conflict, fully autonomous weapons would undermine other, non-legal protections for civilians. First, robots would not be restrained by human emotions and the capacity for compassion, which can provide an important check on the killing of civilians. Emotionless robots could, therefore, serve as tools of repressive dictators seeking to crack down on their own people without fear their troops would turn on them. While proponents argue robots would be less apt to harm civilians as a result of fear or anger, emotions do not always lead to irrational killing. In fact, a person who identifies and empathizes with another human being, something a robot cannot do, will be more reluctant to harm that individual. Second, although relying on machines to fight war would reduce military casualties—a laudable goal—it would also make it easier for political leaders to resort to force since their own troops would not face death or injury. The likelihood of armed conflict could thus increase, while the burden of war would shift from combatants to civilians caught in the crossfire.

Finally, the use of fully autonomous weapons raises serious questions of accountability, which would erode another established tool for civilian protection. Given that such a robot could identify a target and launch an attack on its own power, it is unclear who should be held responsible for any unlawful actions it commits. Options include the military commander that deployed it, the programmer, the manufacturer, and the robot itself, but all are unsatisfactory. It would be difficult and arguably unfair to hold the first three actors liable, and the actor that actually committed the crime—the robot—would not be punishable. As a result, these options for accountability would fail to deter violations of international humanitarian law and to provide victims meaningful retributive justice.

Recommendations

Based on the threats fully autonomous weapons would pose to civilians, Human Rights Watch and IHRC make the following recommendations, which are expanded on at the end of this report:

To All States

- Prohibit the development, production, and use of fully autonomous weapons through an international legally binding instrument.
- Adopt national laws and policies to prohibit the development, production, and use of fully autonomous weapons.
- Commence reviews of technologies and components that could lead to fully autonomous weapons. These reviews should take place at the very beginning of the development process and continue throughout the development and testing phases.

To Roboticists and Others Involved in the Development of Robotic Weapons

- Establish a professional code of conduct governing the research and development of autonomous robotic weapons, especially those capable of becoming fully autonomous, in order to ensure that legal and ethical concerns about their use in armed conflict are adequately considered at all stages of technological development.

LOSING HUMANITY

The Case against Killer Robots

Machines have begun to take the place of humans on the battlefield. Some military and robotics experts have predicted that fully autonomous weapons could be developed within 20 to 30 years. Such weapons, also known as "killer robots," would be able to select and engage targets without human intervention. Their use could undermine the protection of civilians in armed conflict.

After describing the trend toward increased autonomy, this report argues that fully autonomous weapons would be incapable of meeting international humanitarian law standards, including the rules of distinction, proportionality, and military necessity. These robots would lack human qualities, such as the ability to relate to other humans and to apply human judgment, that are necessary to comply with the law.

Fully autonomous weapons would also threaten essential non-legal safeguards for civilians. They would not be constrained by the capacity for compassion, which can provide a key check on the killing of civilians. Furthermore, while their use could reduce military casualties, it might make going to war easier and thus shift the burden of armed conflict onto civilians.

Finally, the use of fully autonomous weapons raises serious questions of accountability because it is unclear who should be held responsible for any unlawful actions they commit.

This report calls for an absolute ban on the development, production, and use of fully autonomous weapons. While domestic measures can serve as a starting point, states should adopt an international treaty prohibiting the weapons before they show up in national arsenals.
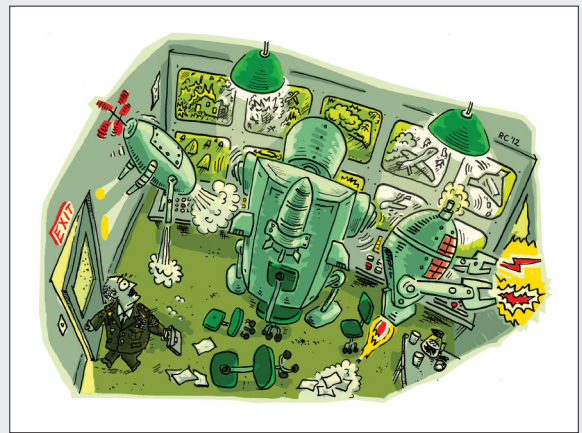
© 2012 Russell Christian for Human Rights Watch

HUMAN RIGHTS WATCH

IHRC | INTERNATIONAL HUMAN RIGHTS CLINIC
HUMAN RIGHTS PROGRAM AT HARVARD LAW SCHOOL

hrw.org

www.law.harvard.edu/programs/hrp/ihrc.html