

World split on how to regulate 'killer robots'

April 18 2016, by Sean Welsh, University Of Canterbury



DARPA is developing an autonomous anti-submarine warfare vessel, ACTUV.
Credit: DARPA

Diplomats from around the world met in Geneva last week for the United Nations' third [Informal Expert Meeting](#) on lethal autonomous weapons systems (LAWS), commonly dubbed "killer robots".

Their aim was to make progress on deciding how, or if, LAWS should be regulated under [international humanitarian law](#).

A range of views were expressed at the meeting, from Pakistan being in

favour of a full ban, to the UK favouring no new regulation for LAWS, and several positions in between.

Despite the range of views on offer, there was some common ground.

It is generally agreed that LAWS are governed by international humanitarian law. For example, robots cannot ignore the principles of [distinction between civilians and combatants](#), or [proportionality in the scale of attack](#).

Human commanders would also have command responsibility for their robots, just as they do for their service men and women. Robots cannot be lawfully used to perpetrate genocide, massacres and war crimes.

Beyond that, there are broadly four positions that the various nations took.

Position 1: Rely on existing laws

The UK's position is that existing international humanitarian law is sufficient to regulate emerging technologies in artificial intelligence (AI) and robotics.



The Northrop Grumman X-47A Pegasus drone is being trialed by the US Navy.
Credit: DARPA

The argument is that international humanitarian law was sufficient to regulate aeroplanes and submarines when they emerged, and it will also cope with many kinds of LAWS too. This would include [Predator drones](#) with an "ethical governor" – which is software designed to determine whether a strike conforms with the specified rules of engagement and international humanitarian law – or autonomous anti-submarine warfare ships, such as the US Navy's experimental autonomous [Sea Hunter](#).

Position 2: Ban machine learning

The French delegation said a ban would be "premature" and that they are open to accepting the legality of an "off the loop" LAWS with a "human in the wider loop". This means the machine can select targets and fire autonomously, but humans still set the rules of engagement.

However, they were open to regulating [machine learning](#) in "off the loop" LAWS (which do not yet exist). Thus, they might support a future ban on any self-learning AI – similar to AlphaGo, which recently beat the human world Go champion – in direct control of missiles without humans in the wider loop. The main concern is that such AIs might be unpredictable.

Position 3: Ban 'off the loop' with a 'human in the wider loop'

The Dutch and Swiss delegations suggested "off the loop" systems with a "human in the wider loop" could comply with international humanitarian law, exhibit sufficiently meaningful human control and meet the dictates of the public conscience.

The UK, France and Canada spoke against a ban on such systems.

Advocates of such robotic weapons claim they could be morally superior to human soldiers because they would be more accurate, more precise and less prone to bad decisions caused by panic or revenge.

Opponents argue they could mistarget in cluttered or occluded environments and are morally unacceptable.

For example, the [Holy See and 13 other nations](#) think a real-time human intervention in the decision to take life is morally required, so there must always be a human in the loop.

This position requires exceptions for already fielded "defensive" weapons such as the [Phalanx Close-In Weapon System](#), and long-accepted "off the loop" weapons such as naval mines, which have existed since the 1860s.

Position 4: Ban 'in the loop' weapons

Pakistan and Palestine will support any measure broad enough to ban telepiloted drones. However, most nations see this as beyond the scope of the LAWS debate, as humans make the decisions to select and engage targets, even though many agree drones are a human rights disaster.

Defining lines in terms of Turing

Formally, an AI is a Turing machine that mechanically applies rules to symbolic inputs to generate outputs.

A ban on machine learning LAWS is a ban on AIs that update their own rule book for making lethal decisions. A ban on "wider loop" LAWS is a ban on AIs with a human-written rule book making lethal decisions. A ban on "in the loop" LAWS is a [ban](#) on robots being piloted by humans being used as weapons at all.

Opinions also differ as to whether control of decisions by Turing computation qualify as meaningful or human.

Next steps

The Geneva meeting was an informal expert meeting to clarify definitions and gain consensus on what (if anything) might be banned or regulated in a treaty. As such, there were no votes on treaty wording.

The most likely outcome is the setup of a panel of government experts to continue discussions. AI, robotics and LAWS are still being developed. As things stand, the world is at Position 1: relying on existing international humanitarian law.

Provided an AlphaGo in charge of missiles complied with principles like discrimination and proportionality, it would not be clearly illegal, just arguably so.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: World split on how to regulate 'killer robots' (2016, April 18) retrieved 19 April 2024 from <https://techxplore.com/news/2016-04-world-killer-robots.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.