AI researchers should not retreat from battlefield robots, they should engage them head-on

August 3 2015, by Sean Welsh

There are now over [2,400](#) artificial intelligence (AI) and robotics researchers who have signed an open letter calling for autonomous weapons – often dubbed "killer robots" – to be banned.

They cite a number of concerns about [autonomous weapons](#), particularly advancing a version of the proliferation argument, which states that military robots will proliferate and lead to destabilising arms races and more conflict around the world.

However, the open letter not only misses out on a number of other concerns raised about autonomous weapons, but it effectively argues that AI and robotics researchers ought to retreat from work on autonomous weapons entirely.

Rather, I think there are very good reasons why AI and robotics researchers ought to engage with autonomous weapons head-on, not least to help make them behave more ethically and to improve our understanding of moral cognition.

In the letter

The open letter argues that the development of autonomous weapons could lead to a serious proliferation problem:

If any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow.

It re-affirms the standard policy conclusion of the Campaign to Stop Killer Robots: autonomous weapons should be banned.

There is an interesting difference in the open letter's position, though: the ban in the letter is only on *offensive* autonomous weapons. *Defensive*autonomous weapons are, apparently, acceptable. It seems the campaign has done the numbers and accepted that NATO, the Gulf Emirates, Israel, Saudi Arabia, Japan and South Korea have spent billions on defensive autonomous weapons like Patriot, Phalanx and C-RAM and will not support these being banned.

The letter also "airbrushes" AI history to support a second policy conclusion: that AI researchers should not sully their hands with the blood and guts of military AI but keep their discipline pristine lest it be "tarnished" by war.

To be blunt, if military applications "tarnish" a field, then AI was born tarnished. It was sired by Alan Turing in Bletchley Park to serve signals intelligence. Turing built his celebrated machine for military ends.

Perhaps TCP/IP – which underpins the entire internet today – is equally "tarnished" because it was born to facilitate continuity of command and control of nuclear missiles.

The "global AI arms race" started in the 1940s and has never stopped. And yet society reaps rich benefits from civilian applications of technologies originally designed for military purposes.

Martin Seligman's work on positive psychology, [Flourish](#), has been [funded by the US Army](#), yet we don't generally consider it to be "tarnished". Likewise, Daniel Kahneman's [research](#) into "System 1" (fast) and "System 2" (slow) was partly funded and inspired by the [Israeli Defence Force](#), yet it has made a terrific contribution to helping us understand how we think.

Not in the letter

Yet there are also other reasons why we might be concerned about autonomous weapons which are not covered in the open letter:

- Robots cannot discriminate targets (i.e. distinguish between friend, foe and civilian)
- Robots cannot calculate proportionality (e.g. figure out if killing a high value terrorist target is worth risking the deaths of innocent children)
- Robots cannot be held responsible for their actions (because they are not genuine moral agents with free will that can choose their actions and be praised, blamed or punished for them) so there is an "accountability gap" or "responsibility gap"
- Robots will lower the enter cost of war and make wars and conflicts more likely
- Robots will exacerbate the decline of martial valour already started with the use of remotely piloted vehicles
- Robots should not make the decision to kill people.

These are serious issues. However, they are the kinds of issues that AI and robotics researchers are perfectly placed to tackle. They are precisely the ones with the expertise to develop solutions that might make autonomous weapons able to reduce human casualties on future battlefields by [making them more ethical](#).

It is for this reason that I did not sign the open letter.

AI should pay more attention to the ADF manual

Using Kahneman's terminology, rather succumb to the "fast" intuitive appeal to the System 1 of their moral cognition, AI and robotics researchers should engage their "slow" System 2 and work on the deep problem of moral cognition in combat.

The Australian Defence Force manual is a good place to start. It is better than the American manuals because Australia has actually signed the Additional Protocols to the Geneva Conventions and the Convention on Certain Conventional Weapons, and the various other major treaties of International Humanitarian Law. The United States, alas, has not.

AI researchers should read this manual. There is far more to military life than the decision to shoot the enemy when necessary. There are obligations to mitigate and minimise the calamities of war, obligations to uphold human rights and to protect cultural property. Sometimes cognitive agents with weapons must decide to fire to defend these rights and obligations.

Having read the manual, AI (and related fields) should research what functions required of service personnel can be automated. They should research the risks of automation, and recognise the failures of "meaningful human control".

We need them to research the ethics of risk transfer, risk imposition and risk elimination, and figure out how to represent and process such normative data in a machine.

Or, if they prove it cannot be done, then design a better machine that *can* do it. They should decide whether humans or robots are better

options for the battlefield as technology advances, and contemplate the psychological damage done by combat even to "cubicle warriors".

We can then make policy decisions on the basis of hard data derived from experiments.

AI should thus not withdraw, but engage in the challenge of ethical cognition in the military robotics front line.

This story is published courtesy of [The Conversation](link) *(under Creative Commons-Attribution/No derivatives).*

Source: The Conversation

Citation: AI researchers should not retreat from battlefield robots, they should engage them head-on (2015, August 3) retrieved 7 May 2024 from https://phys.org/news/2015-08-ai-retreat-battlefield-robots-engage.html