

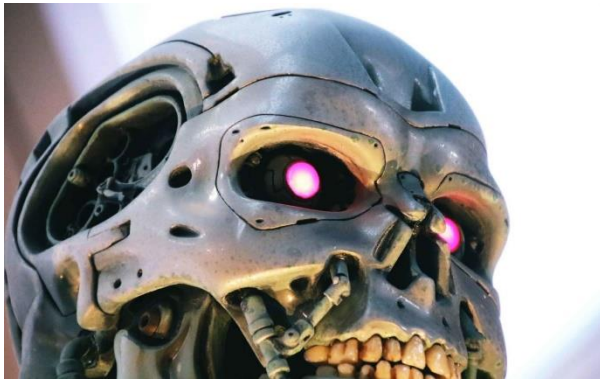
# REACTION

---

## How do you make killer robots act ethically?

[Nayef Al-Rodhan](#)

[June 23, 2022](#)



Credit: Usa-Pyon via Shutterstock

Trenches, artillery barrages, lumbering tanks: at first glance the [war in Ukraine](#) looks like a throwback to a time-yellowed history book. But there is more to it than meets the eye and the conflict could be a preview of future warfare. Both sides are using artificial intelligence-enabled lethal autonomous weapons systems and so-called suicide drones that independently detect, stalk and shoot targets. These autonomous weapon systems – killer robots, if you like – are increasingly becoming a staple of modern warfare.

In recent months, UN officials in Geneva have tried unsuccessfully to chisel out a treaty that sets clear-cut legal limits for autonomous weapons. As the US Defence Department updates its guidance on autonomous weapons to incorporate advances in artificial intelligence, the international community needs to ask – and answer – some important existential questions: how can we regulate killing machines before it is too late? And how do we guarantee vital moral safeguards in the face of a global AI arms race?

The United Kingdom, Australia and the United States – all big investors in weapons with autonomous capabilities – are quick to claim that increased robotisation will help sanitise the battlefield by keeping soldiers out of harm's way. But this brushes over serious ethical and legal concerns about the harmful consequences of these lethal weapons.

Chief among them is the “responsibility gap”. Machines, no matter how sophisticated, will never be able to conform with the legal and moral requirements of the laws of war. When a soldier makes a fatal error in war – for example by confusing civilians for combatants – the incident can often be traced back to individuals who can be held accountable. This is not the case with robots: if a machine commits a war crime on its own volition, who do you hold responsible? The commander who sent the robot into battle? The programmer? Or the government that invested in the technology in the first place?

These debates are still largely being held on a theoretical level, as in most cases a human controller still makes the final call on whether to authorise the attack. But this might soon change as weapons acquire ever more autonomous capabilities. One quivers at the thought of killer robots being used by state and non-state actors, who could use facial recognition and other AI technologies to target individuals or groups. We are still a far cry from the world depicted in Isaac Asimov’s science fiction novels, where robots make life and death decisions, but developments in quantum computing and neuromorphic chip technology have made the prospect of fully autonomous robots tangible. Through inbuilt chips that emulate the human neuronal networks, robots will, in theory, be able to develop their own moral codes as they engage with their surroundings. This raises another thorny question: whose moral codes would robots inherit?

The international community has so far failed to regulate lethal autonomous weapons systems (LAWS), despite a decade of on-off UN talks. Countries at the forefront of LAWS development have held their ground, hiding behind existing international humanitarian laws. But the tide is changing. A growing number of countries have called for restrictions on developing and using LAWS. An international coalition of nongovernmental organisations is gaining traction and has managed to enlist vocal supporters in the tech industry.

A widely accepted regulatory framework governing LAWS is still a distant prospect, but optimists can find glimmers of hope in the [Treaty on the Non-Proliferation of Nuclear Weapons](#) (NPT), which has been signed by almost 200 states. A similar non-proliferation regime governing LAWS could help control their development, procurement and use, although it would be tough to implement, not least because of the dual-use nature of technology.

Legal tools alone will not be enough to shield the world from the dangers of autonomous weapons. A broader debate about the ethics of these robotics is needed. This debate needs to take place at all levels of society. Why is it that so few technology academies or Big Tech companies offer classes in ethics? Is it unthinkable to introduce an equivalent of the Hippocratic oath to the robotics field? There is an urgent need to reconcile new innovations with norms and ethics of war, including international humanitarian law and the Geneva Conventions.

Above all, we need to dig deep and re-evaluate what makes us human. Neuroscientific research shows that most of us, most of the time, are neither innately moral, nor immoral, but rather amoral. This means that our moral compass will largely depend on our “Perceived Emotional Self-Interest” and will be heavily influenced by our personal circumstances as well as our inborn predilections, among them a predisposition to choose actions that maximize our chances of survival. This should sound alarm bells for global leaders as the global AI arms race heats up and advanced technologies make their way from research labs to the military to the open market. We should stay alert to the likely exponential increase in violence that will result from killer machines. This enhanced and personalised brutality will most likely complicate post-conflict reconciliation and reconstruction.

As Bertrand Russell said in the early 1920s: “Without more kindness in the world, technological power will serve to increase men’s capacity to inflict harm on one another.” Almost a hundred years

on, Russell's words are a painful reminder that humanity is becoming more creative at inflicting harm. Only strict regulation and diplomatic strong-arming will protect us from our worst selves.

*Professor Nayef Al-Rodhan is a neuroscientist and philosopher. He is an Honorary Fellow at St Antony's College, University of Oxford, and the Head of the Geopolitics & Global Futures Programme at the Geneva Centre for Security Policy (GCSP).*