

KILLING BY MACHINE

| | |

KEY ISSUES FOR
UNDERSTANDING
MEANINGFUL HUMAN
CONTROL



UNDERSTANDING HUMAN CONTROL OVER WEAPONS

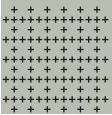
In current practice, there is an expectation that human control is exercised over the use of weapons. This means when, where and how weapons are used; what or whom they are used against; and the effects of their use. Increasingly autonomous weapon systems threaten to erode what we have come to expect in terms of human control over weapons. Weapon systems that operate outside of the parameters of meaningful human control are neither ethically acceptable nor legally permissible.

Article 36 argues that meaningful human control over weapon systems is required in every individual attack. States should develop new international law to make this requirement explicit. Understanding how we exercise control over existing weapons provides critical guidance for developing an adequate response to increasing autonomy in weapon systems.

In the case of weapon systems that can detect and attack target objects without direct human intervention, critical aspects of human control broadly relate to:

- ✘ The pre-programmed target parameters, the weapon's sensor-mechanism and the algorithms used to match sensor-input to target parameters.
- ✘ The geographic area within which and the time during which the weapon system operates independently of human control.

The charts on the reverse illustrate these two key themes, and how they relate to each other.

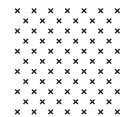


THE CAMPAIGN TO STOP KILLER ROBOTS

Article 36 is a founding member of the *Campaign to Stop Killer Robots*.

The Campaign to Stop Killer Robots calls for a pre-emptive and comprehensive ban on the development, production, and use of fully autonomous weapons, also known as lethal autonomous robots. This should be achieved through new international law (a treaty), as well as through national laws and other measures.

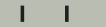
www.stopkillerrobots.org



ARTICLE 36

Article 36 is a UK-based not-for-profit organisation working to prevent the unintended, unnecessary or unacceptable harm caused by certain weapons. Article 36 undertakes research, policy and advocacy and promotes civil society partnerships to respond to harm caused by existing weapons and to build a stronger framework to prevent harm as weapons are used or developed in the future. The name refers to article 36 to the 1977 Additional Protocol I to the Geneva Conventions that requires states to review new weapons, means and methods of warfare.

www.article36.org

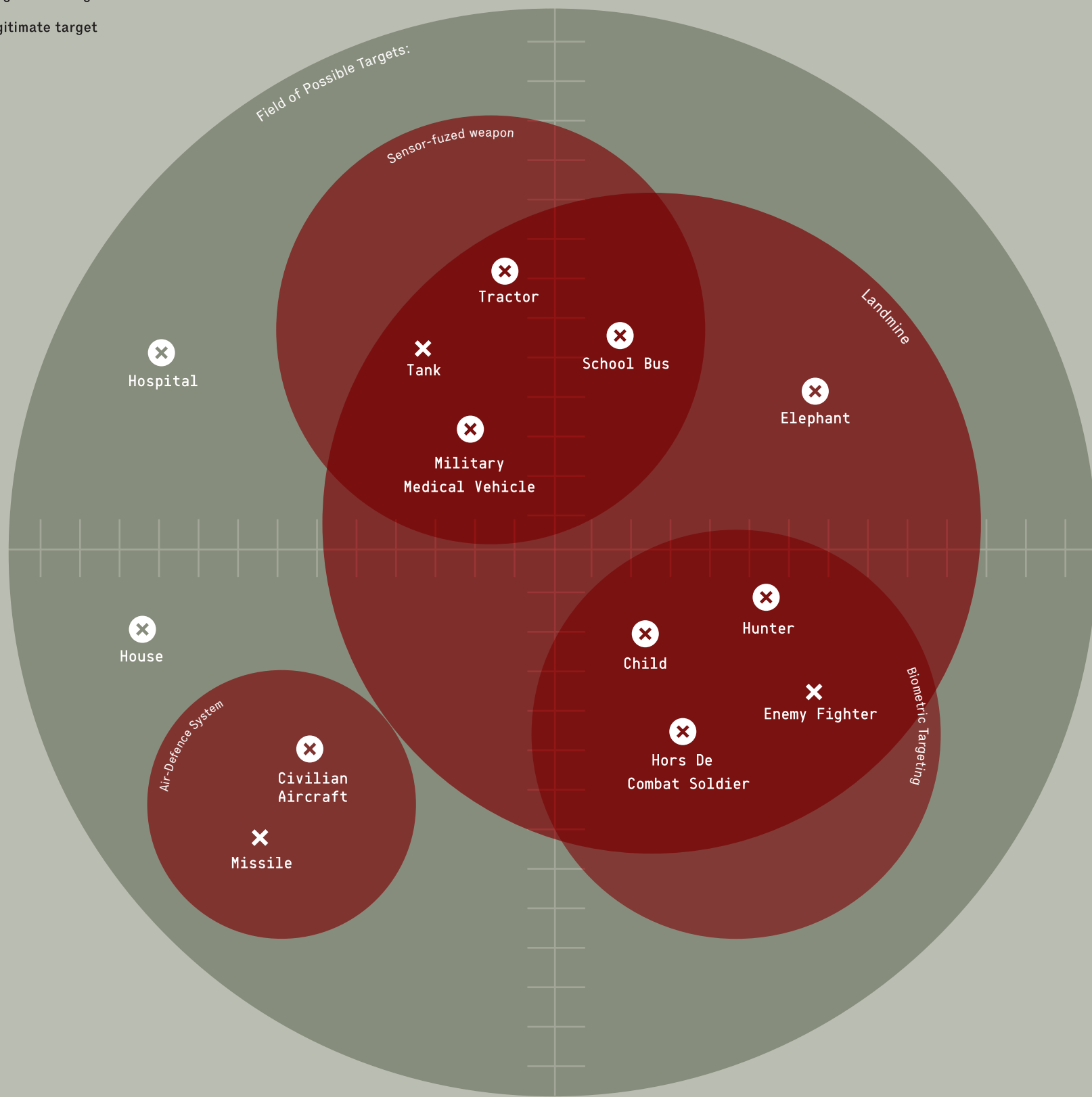


THE CONTRACTING OR
ACCEDING PARTIES RESERVE
TO THEMSELVES TO COME
HEREAFTER TO AN
UNDERSTANDING WHENEVER A
PRECISE PROPOSITION SHALL
BE DRAWN UP IN VIEW OF
FUTURE IMPROVEMENTS WHICH
SCIENCE MAY EFFECT IN THE
ARMAMENT OF TROOPS, IN
ORDER TO MAINTAIN THE
PRINCIPLES WHICH THEY HAVE
ESTABLISHED, AND TO
CONCILIATE THE NECESSITIES
OF WAR WITH THE LAWS OF
HUMANITY.[×]

× 1868 Saint Petersburg Declaration



⊗ illegitimate target
 ⊗ legitimate target



WHEN MACHINES MAKE PEOPLE THE OBJECT OF ATTACK

The technology behind an algorithmic 'decision' to detect, select and cause harm to a target is highly complex. Characteristics of objects that are amenable to being 'sensed' by a machine, including their infrared emissions or their shape, or biometric information about persons, are used as 'proxy indicators' of a target.

Rendering the world machine-sensible reduces people to objects. This is an affront to human dignity. As the UN Special Rapporteur on extrajudicial, summary or arbitrary executions has observed, grave consequences for individuals are at stake in any use of force, requiring deliberative, human decision-making that considers ethical implications. Without this, the use of force becomes dehumanized and there is a vacuum of moral responsibility. Social processes of moral and legal

reasoning on the use of force cannot be encoded into machines.

Related to the use of proxy indicators are concerns about probabilistic matching programs used for identifying and verifying appropriate objects of attack. Such programs have a specified 'tolerance of error' to offset their 'efficiency' in producing a match, with their 'accuracy' of producing a valid match (identifying an appropriate object of attack). Within the tolerance of error are objects that match the system's criteria (based on proxy indicators), but that are not appropriate objects of attack. The use of a weapon system that, by its very design, harms a certain number of people and objects that are not appropriate objects of attack (e.g. a civilian, a hospital, a civilian plane) is morally and legally highly problematic.

FIG. 1 (OPPOSITE)

Persons and objects other than legitimate targets may fall within the target parameters of a weapon and be made the object of attack.

A pressure-activated anti-personnel landmine can be triggered by any person or object weighing more than a specified amount. Consequently, children, other civilians, heavy animals, and civilian vehicles can be harmed by landmines, even though they are not appropriate targets of attack.

If a sensor-fuzed weapon is used in an area that contains civilian objects, it may identify civilian tractors, buses, or aircraft and military medical vehicles as valid targets on the basis that their shape or heat signatures match pre-programmed target parameters. Such vehicles are, however, a priori protected from attack under international humanitarian law.

Similarly, a weapon system identifying targets on the basis of biometric signatures or similar proxy indicators may identify children, adult civilians and soldiers hors de combat as targets, even though these persons are not legitimate targets of attack under international humanitarian law.

information about the locations in which harm can be caused by a weapon system, and what these locations contain. This is of particular importance in terms of people and objects that may be present in the location where force is applied, but should not be attacked or that must be protected against the effects of military operations..

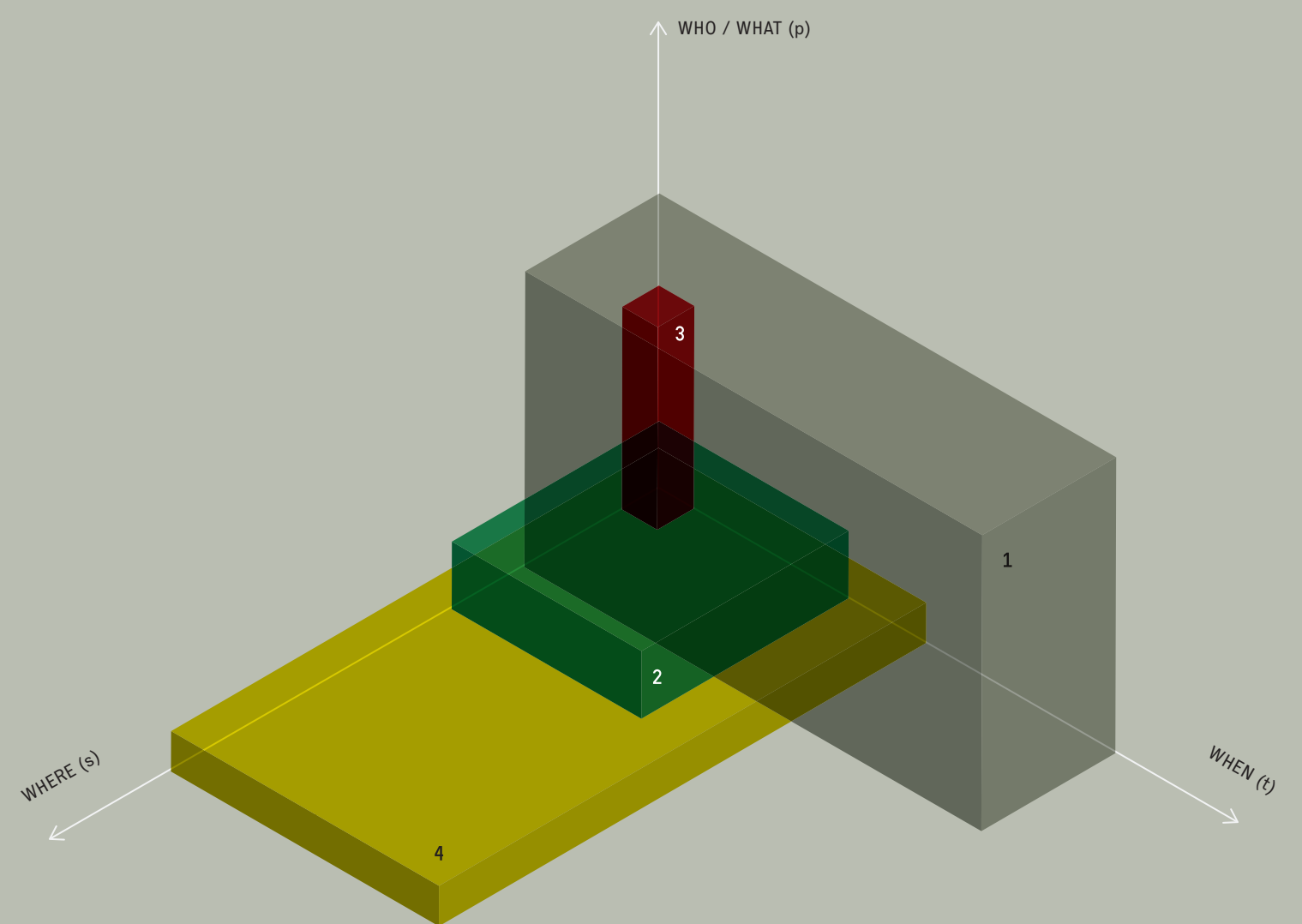
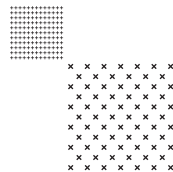
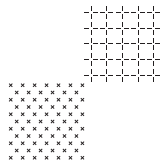


FIG. 2



JUSTIFICATION IS REQUIRED FOR ENGAGING IN WAR AND IN OTHER FORMS OF VIOLENCE, AND THEREFORE, IT SEEMS TO FOLLOW, JUSTIFICATION IS REQUIRED FOR ENGAGING IN THE SEARCH FOR THE KNOWLEDGE TO PROVIDE THE MEANS TO FIGHT WARS AND COMMIT OTHER FORMS OF VIOLENCE. ×

× John Forge



RECOMMENDATIONS

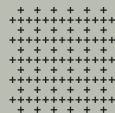
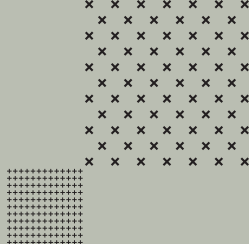
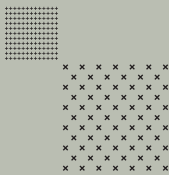
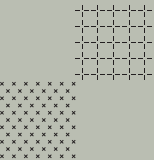
Human control over the use of weapons and their effects is essential to ensuring that the use of a weapon is morally justifiable and legal. Such control is also required for accountability over the consequences of the use of force. To demonstrate that such control can be exercised, states must show that they

- × understand the process by which a system identifies individual target objects, and
- × understand the context in space and time where an attack will take place.

Given the development of greater autonomy in weapons systems,

- + States should make it explicit that meaningful human control is required over individual attacks.
- + Weapon systems that operate without meaningful human control should be prohibited.
- + States should explain how control is applied over existing weapon systems, especially those with autonomous or automatic functions, and why they consider such systems to be acceptable and permissible.





Article 36
81 Rivington Street
London EC2A 3AY
United Kingdom
info@article36.org

www.article36.org

