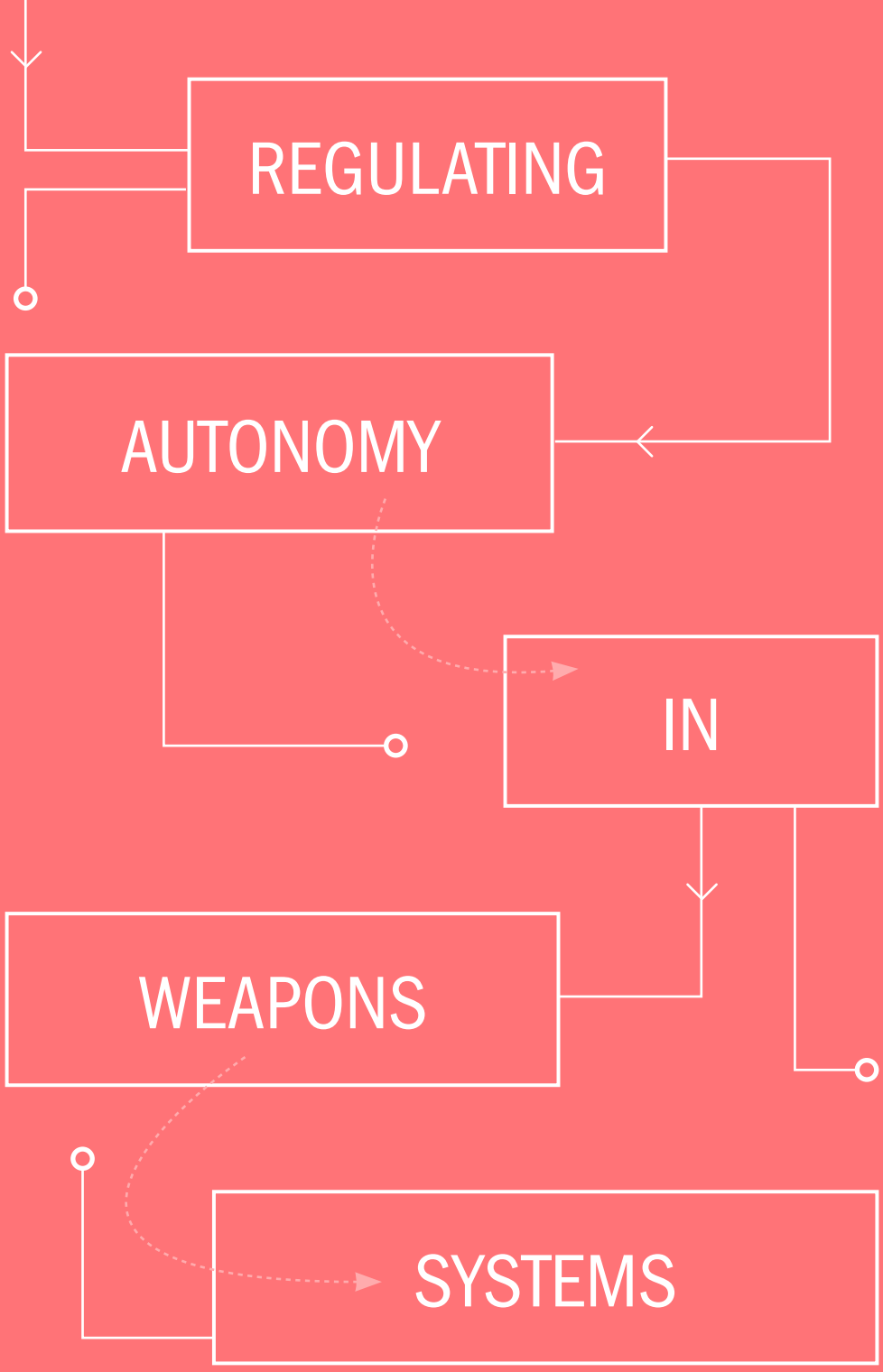

REGULATING
AUTONOMY
IN
WEAPONS
SYSTEMS

We are working to avoid a dehumanised future, where machines can be tasked to kill and apply force without people understanding or being fully responsible for the consequences.

International discussions on ‘autonomy’ in weapons systems are building a common understanding of the prohibitions and other obligations needed to preserve human dignity and ensure meaningful human control. It is now necessary to focus in detail on the specific components of a solution.

This pamphlet provides a basic model of how a treaty to address autonomous weapons could be structured – and illustrates how that structure responds to the problems that increased autonomy in weapons systems raise.

For us, there are two key problems that we need to work together to solve:

- firstly, which systems within the scope of discussion are fundamentally unacceptable; and
- secondly, how human control can be maintained over the remaining systems in this area, in order to adequately uphold both legal obligations and more profound moral and ethical principles.

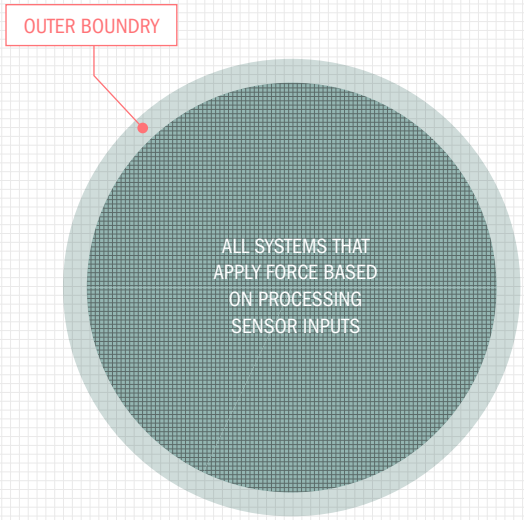
AUTONOMOUS WEAPONS CHALLENGE OUR VALUES

1. See analysis in Article 36 (2019) 'Targeting people,' https://bit.ly/2Zuy4cb

A STRUCTURE TO REGULATE AUTONOMY IN WEAPONS SYSTEMS

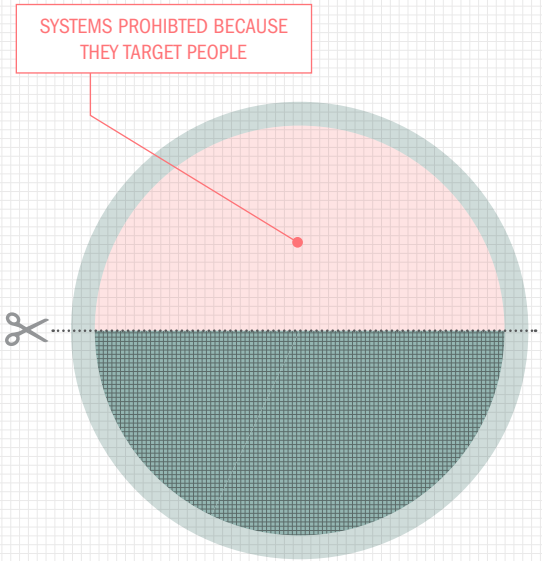
1. A BROAD SCOPE OF TECHNOLOGIES

We are regulating systems that use sensors to determine where and when force will occur, without this being set specifically by a person. So those systems all fall within the outer boundary here. Although there are many different notions of 'an autonomous weapon' all of them are based on this foundation. Within this broad category, our structure of regulation is going to prohibit certain ways of functioning and apply broad rules for the use of others.



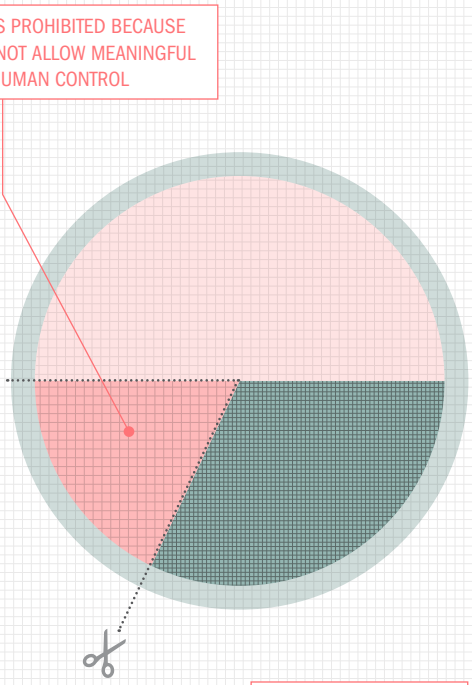
2. NOT KILLING PEOPLE WITH SENSORS

We then divide our category of sensor-based systems into two – those that use target profiles that represent people, and those that don't. Systems that target people should be prohibited because they undermine human dignity – we are not allowing machines to identify people to be subject to harm (whether these are 'lethal' systems or not).



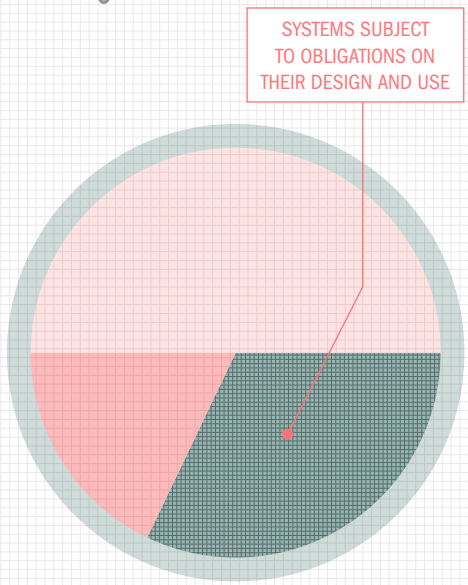
3. PROHIBITING SYSTEMS THAT CANNOT BE CONTROLLED

Next we cut out systems that cannot be effectively controlled. For example, although they are not targeting people, we still should not allow systems that 'set their own goals', or where the conditions under which they will apply force can change during use or where their functioning cannot be explained. Prohibitions and restrictive obligations on the development and review of systems will be needed to establish this line – preventing systems that cannot be used with meaningful human control.



4. ENSURING MEANINGFUL HUMAN CONTROL OVER WHAT IS LEFT

The systems that are left still use sensors to determine specifically when and where force will occur, which presents significant challenges. 'Positive obligations' – rules on the use of these systems – should require users to control location, duration and target specification, as well as other aspects of design and use. This is necessary to protect existing law from erosion.

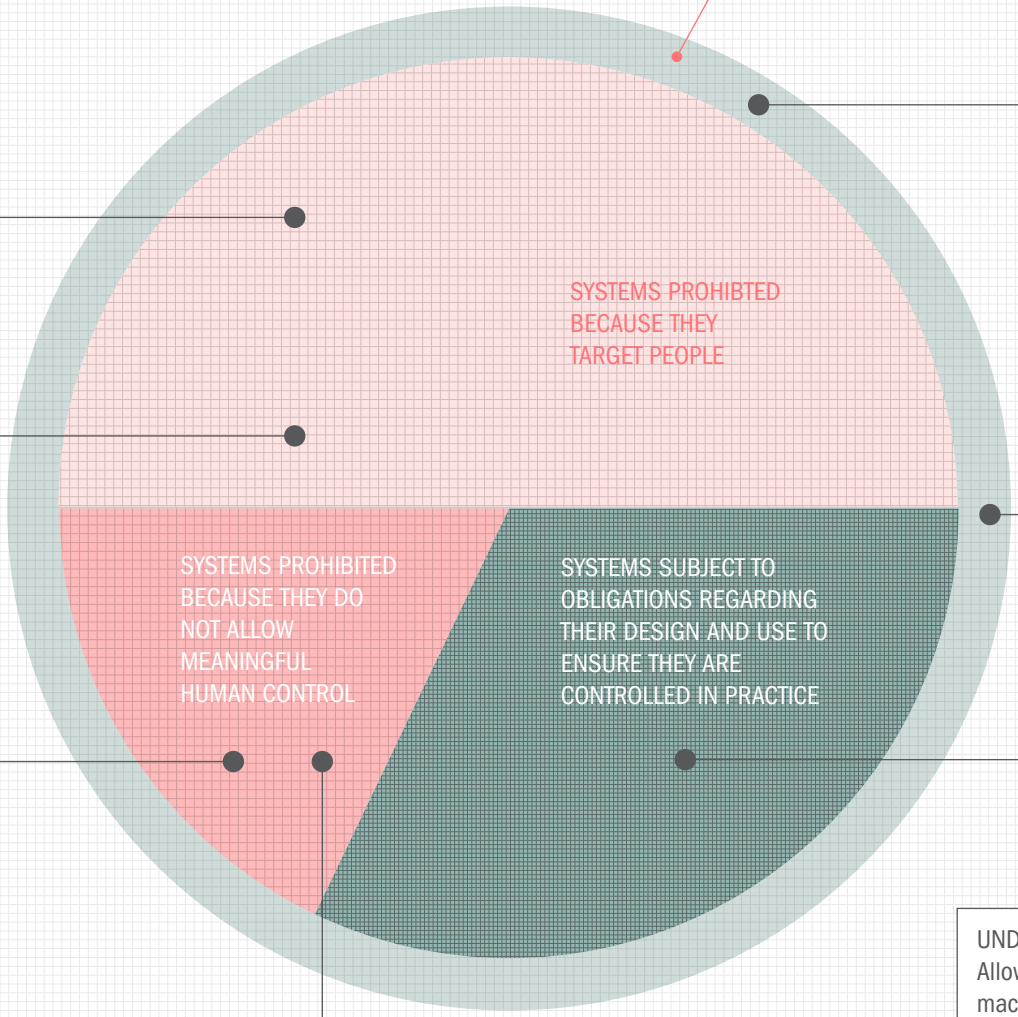


HOW THIS STRUCTURE ADDRESSES THE KEY PROBLEMS RAISED AROUND AUTONOMY IN WEAPONS SYSTEMS

OUTER BOUNDARY: SYSTEMS THAT APPLY FORCE BASED ON PROCESSING SENSOR INPUTS

DANGER TO CIVILIANS
Civilian protection is eroded by systems that target people, or that are not effectively controlled. All aspects of this regulation structure work to strengthen civilian protection.

DEHUMANISATION
Prohibiting all systems where sensors are used to target people, within a broad technological scope, would be a milestone for the protection of human dignity in the face of developing technologies.



RISKS TO PEACE AND SECURITY
An instrument with a broad scope, a logical structure and with clear normative lines – like the prohibition on targeting people – will set a compelling standard even for states that do not join it at first. An instrument structured along these lines will shape the development of technologies for the future.

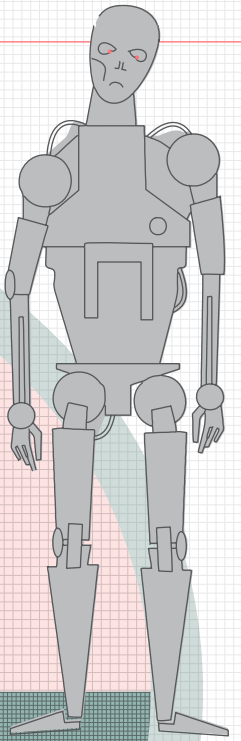
OPAQUE TECHNOLOGIES
Ensuring that systems can be effectively understood makes meaningful human control possible. Prohibiting the targeting of people altogether also removes the possibility of discimination against or between people on whatever grounds being reproduced in systems.

UNDERMINING THE LAW
Allowing unpredictable technologies, suggesting that machines are authorised to make legal decisions, or having people make legal decisions with no real understanding of the context of an attack would all erode fundamental aspects of the law. Prohibiting technologies that don't allow control, and placing obligations on how control is understood will protect the law for the future.

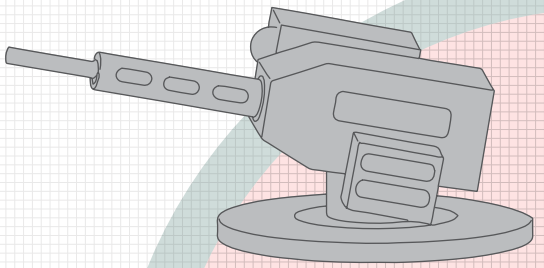
HOW THIS STRUCTURE ADDRESSES DIFFERENT TECHNOLOGIES*

*REAL OR IMAGINED!

TERMINATORS!
Targeting people, and not amenable to meaningful human control – the Terminator would be prohibited! This is lucky given the system's high media profile...

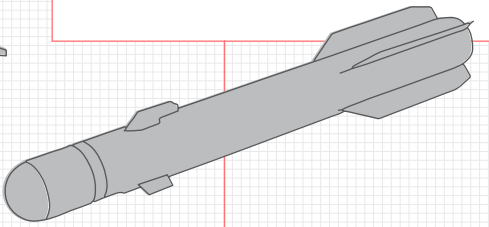


ANTI-PERSONNEL SENTRY ROBOTS
Would be prohibited where they would apply force automatically upon sensing a person (e.g. Super aEgis II in automatic mode).



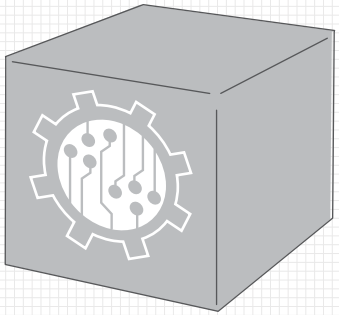
SYSTEMS PROHIBTED BECAUSE THEY TARGET PEOPLE

SYSTEMS RELEASED WITHIN A 'TARGETING AREA' to destroy objects with particular signatures (e.g. Brimstone anti-tank missile), would be subject to positive obligations. They must be sufficiently predictable, and their location and duration of operation must be sufficiently controlled to allow legal rules to be applied.

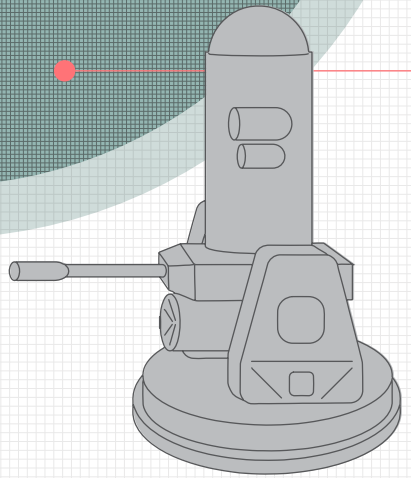


SYSTEMS PROHIBITED BECAUSE THEY DO NOT ALLOW MEANINGFUL HUMAN CONTROL

SYSTEMS SUBJECT TO OBLIGATIONS REGARDING THEIR DESIGN AND USE TO ENSURE THEY ARE CONTROLLED IN PRACTICE



'BLACK BOX' SYSTEMS - for example where target profiles are constructed through 'machine learning', or where target profiles might change during the course of use, without human approval. These would be prohibited because their implications in a specific use could not be sufficiently controlled.

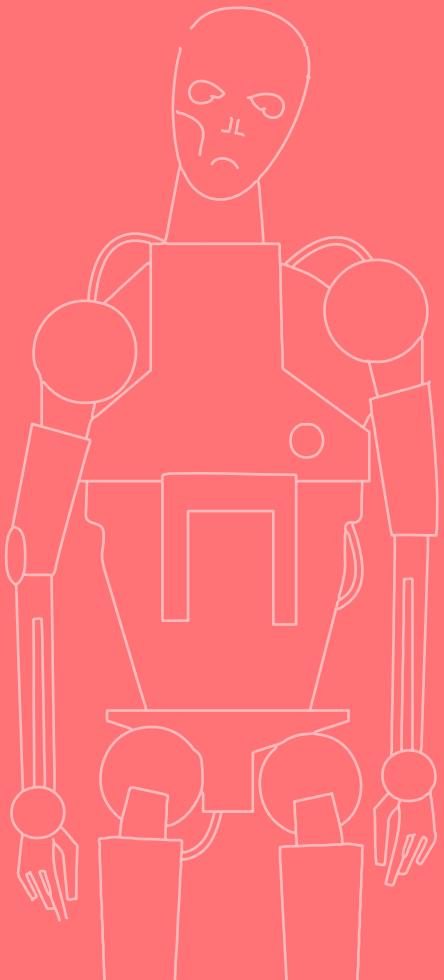


DEFENSIVE SYSTEMS operating at high speeds (e.g. missile defence systems like Phalanx CIWS) would be subject to positive obligations. These would promote the sorts of practices already used by certain militaries.

Article 36 is a specialist non-profit organisation focused on reducing harm from weapons.

www.article36.org
info@article36.org
@article36

Design: bb-studio.co.uk
Printed by Blackmore Ltd, Shaftesbury, Dorset



Article36