Autonomous Weapons
An ethical overview

Alessandro Dicosola
alessandro.dicosola@studio.unibo.it

Pierpasquale Colagrande
pierpasqu.colagrande@studio.unibo.it

Sami Osman
sami.osman@studio.unibo.it

1 Introduction

The improvements in Artificial Intelligence and robotics research unleashed the possibility to develop weapons that are more intelligent and that automatize some functions. These weapons are called Autonomous Weapons Systems (AWS for short) and some of these are currently used in war context and, in certain cases, even in law enforcement situation. However, technology is moving fast in this area and the progressive development and demployment of new autonomous weapons raise some questions about wether they should be regulated or banned, especially looking at the possibility that, in the near future, these systems can become more and more autonomous and, potentially, more and more dangerous. Activists and human rights experts conclude that they sould be banned or regulated otherwise. Moreover, the spread of these weapons may cause other important problems like the proliferation, which could result in the usage of such weapons by criminal groups, non-state armed groups and terrorists. For these reasons, activists, human rights experts and political personalities encourage the start and continuation of discussions about these weapons in the various international forums. For example, mexico cartels used drones to drop explosive on police, hurting two officiers. A group of artists "armed" BostonDynamics' Spot robot with a paintball gun that can be remotely controlled to raise awareness about the future of autonomous weapons and, more generally, autonomy.

2 Autonomy in Weapons Systems

Before talking about the problems and concerns about the use and development of Autonomous Weapons Systems is, however, necessary to define what autonomy is and how it is included in AWS. Autonomy is already used in various tasks like aiming, tracking and identifying targets, deciding the timing to fire a weapon and detonation time. However, there is still no definition of what an autonomous weapon is and this lack of clarity on basic terminology is a recipe for disaster, mostly because concerns about autonomous weapons are usually aimed at potential future weapons, not at weapons that already exist, therefore giving to "autonomous weapons" a broad meaning is wrong. The rapid development of these kind of weapons underscores a urgency about understanding more clearly the challenges posed by potential future autonomous weapons. We will try to define "autonomous weapons" by answering to three questions:

- What is autonomy?
- How is autonomy used today in weapons?
- What is an "autonomous weapon"?

2.1 What is autonomy?

An autonomous system is "a machine that, once activated, performs some tasks or functions on its own. Autonomous systems are included in a lot of equipment

that people use everyday, like autonomous lane keeping and intelligent cruise control. In fact, increased autonomy can have many advantages such as increased safety and reliability. Defining the level of autonomy that a system has is a crucial aspect to understanding the ethical challenges with the increasing autonomy of such systems. Researchers define levels of autonomy, so we can think of autonomy having three main dimensions along which it can vary:

- human-machine command-and-control relationship
- complexity of the machine
- type of decision being automated

2.1.1 The human-machine command-and-control relationship

The first dimension of autonomy is the relationship between the machine and the human. It is possible to define three levels of autonomy related to this dimension:

- semiautonomous, or human in the loop: machines that perform a function for some period of time, then stop and wait for human input
- human-supervised autonomous, or human on the loop: machines that can perform a function entirely on their own but have a human in a monitoring role, with the ability to intervene
- fully autonomous, or human out of the loop: machines that can perform a function entirely on their own with humans unable to intervene

2.1.2 The complexity of the machine

The second dimension of autonomy is the complexity of the system. It is possible to define three levels of autonomy related to this dimension:

- automatic systems: systems that have very simple mechanical response to environmental input
- automated systems: more complex, rule-based systems
- autonomous systems: machines that execute some kind of self-direction, self-learning or emergent behavior that is not directly predictable from an inspection of its code

2.1.3 Type of function being automated

The third dimension of autonomy is the function being automated. Different decisions have different levels of complexity and risk and any machine would have some tasks that are automated while others that are completely in human control. So, defining which task is automated also helps to define the level of autonomy that the machine has, more risky and complex tasks being automated

means a machine that is more autonomous. For example, a self-driving car can bring you from point A to point B, but it is still the human that decides the destination and, in some cases, the path to follow. In military, for example, acquiring, tracking and engaging targets may be autonomous, but the human is still deciding the target. An increased intelligence of a systems doesn't mean transferring more control from the human to the machine, so, when talking about autonomous weapons, we must clarify autonomous with respect to which functions or tasks.

2.2 How is autonomy used today in weapons?

Automated military systems with human-supervised modes have been in existance from decades and are used to day by very different military forces. So, when it comes to weapons, we can divide autonomous weapons systems into three main categories:

- **human in the loop**, or semiautonomous: systems that use autonomy to engage individual targets or group of targets that a human as decided are to be engaged
- **human on the loop**, or human-supervised: systems that use autonomy to select and engage targets where a human has not decided to those specific targets are to be engaged, but the human can intervene and halts the system operations
- **human out of the loop**, or fully autonomous: systems that use autonomy to select and engage targets where a human has not decided to those specific targets are to be engaged and human cannot intervene and halt the system operations

2.3 Human in the loop

These are weapons systems that use autonomy to engage individual targets or group of targets that a human as decided are to be engaged. Weapons that are currently in use in this category are:

- guided munitions: projectiles, bombs, missiles and other weapons that can actively correct the initial-aiming or subsequent errors after being fired in order to precisely hit a target chosen by a human operator, they are divided into go-onto-target, munitions using sensors to hit a target based on its signature, and go-onto-location-in-space, munitions designed to hit a particular location in geographic space, via GPS or laser guidance. Some can be aborted, controlled or re-targeted in flight, while others cannot even be recalled.
- **identifying, acquiring, tracking, cueing and prioritizing targets**: radars and other sensors that help to acquire, track, identify and cue potential targets to human operators; used in combination with guided

weapons, use of force can therefore involve a high degree of autonomy, however the decision of which target(s) to hit is still made by a person.

These weapons are designed to strike a specific target chosen by a human, not to search over a wide area for a target and then hit it, so the decision related to the use of force is still a human decision. They have also give human a greater control over warfare, because they are more precise, rather than carpet-bombing whole cities. We can so deduce that increased autonomy of weapons used in warfare is not a bad thing since it can help to significantly reduce the number of civilian and non-combatant victims. These kind of weapons do not raise new issues, instead they are a continuation of the way in which autonomy has been used for years.

2.4 Human on the loop

These weapons are used in defensive situations where requiring a human to remain "in the loop" would be physically impossible because of the required reaction times. Autonomy is used to to complete the engagement over incoming threats that meet certain criteria according to pre-programmed rules. Humans can intervene to halt the weapon, but do not make a positive decision to engage specific targets. Current existing systems of this type are the Aegis, Patriot and Phalanx systems used to defend warships and military bases. These systems only target objects, not humans, and the human can still intervene if necessary or terminate further functioning of the system. Moreover, human also have a physical access and can exercise hardware-level overrides. These weapons are not a threat to human rights since they do not select humans but only objects.

2.5 Human out of the loop

Currently there are very limited number of existing weapons that have a human fully out of the loop. These weapons use autonomy to engage general classes of targets according to pre-programmed rrules abd humans controllers are not aware of the specific tarrgets being engaged. These weapons include:

- **loitering munitions**: munitions launched in a general area where they will loiter, searching for a pattern of targets withing a general class (enemy radars, ships or tanks). Once they find a target that meets the requirements, the weapon flies into it and destroys it.
- **encapsulated torpedo mines**: they are a special case of sea mine that, once activated by a passing ship, release a guided torpedo that engages the target.

Thus, the human operator does not know which target are going to be engaged, only knows that the weapon will engage a specific lass of target in a broad geographic area. So, differently from guided munitions, the human operator does not select any target. so the mine is given a greater freedom of maneuver

and release the torpedo missile, so these mines are more similar to loitering munitions and the human does not know the targets that will be hit, but only the geographic location.

2.6 What is an autonomous weapon?

There is no shared definition of autonomous weapon, however it is possible to define autonomous weapons as "weapons that, once activated, can select and targets on their own". From a technical perspective, we can define three classes of autonomous weapons:

- **autonomous weapon system**, or human out the loop: weapon system that, once activated, can select and engage targets where a human has not decided those specific targets are to be engaged
- **human-supervised autonomous weapon system**, or human on the loop: weapon systems that have the same characteristics of an autonomous weapon system but with the additional possibility of a human operator to intervene, monitor and halt the system operations if necessary
- **semi-autonomous weapons**, or human in the loop: weapon systems that, once activated, are intended to only engage individual or group targets that a human has decided to be engaged

Another way to define autonomous weapons is to focus on the functions being automated, so, for example, autonomous weapons can be called "self-targeted weapons" while semi-autonomous weapons can be called "human-targeted" weapons. All these definition help understanding the human and the machine role in decisions about the use of force, which is crucial when talking about the ethical outcomes of the use and development of autonomous weapons.

3 Automous Weapons in warfare

3.1 Examples

Autonomous Weapons and, more generally, robots are currently being tested for usage in warfare context. The use of such weapons can have different effects both on the civilians and the soldiers. The use of these weapons over battlefield have some PROs, in term of human dignity, but also some CONs, for this reason researchers suggest to discuss over the deployment of these systems in order to avoid a preventively ban only on the basis of fear and prejudice. However, experts and researchers also suggest that this regulation should be addressed very carefully in order to propely adapt the law to these new weapons and to avoid the use of these systems to further violate human rights. When it comes to warfare, the general "code of conduct" in humanitarian terms that the soldiers and war parties in general should adhere is the International Humanitarian Law (or IHL for short). Soldiers must respect the guidelines described in this

law in order to mantain a dutiful behaviour to human rights even in a warfare environment. However, when it comes to practice, soldiers have the tendency to break these laws for various factor, mostly related to human emotions. For example, French army is testing Boston Dynamics' Spot robot in combat scenarios, using it for reconnaisance, while Azerbaijan deployed Israel's kamikaze drones in fighting with Armenia.

3.2 PROs of AWS in warfare

Artificial intelligence and robotics applied to weapons helped increasing the accuracy and precision of weapons systems. This means that the more weapons get precise thanks to these technologies, the more it is possible to avoid civilian and non-combatant causalities, property damage and friendly causalities.

3.2.1 Problems with human soldiers

Humans have difficulties in adhering laws, especially in situations like war. This happens because humans feel emotions like fear and usually have the tendency to self-preservation. This could lead to a tendency of seeking revenge, to dehumanize enemies, to feel pleasure from hurting or killing and, when it comes to saving their lives or their colleagues lives, it is very difficult for soldiers to respect such legal and ethical limt in modern warfare. Furthermore, battlefield atrocities persist, despite the introduction of the IHL and in certain cases there are well documented examples of abuse of unmanned robotic systems. Moreover, human soldiers may suffer of the psychological problem of "scenario fulfillment" which is a war scenario in which humans, due to stressful situation, tend to use new incoming information to only fit their pre-existing belief patterns, ignoring informations contradicting this patterns.

3.2.2 Why technology can lead to a reduction of causalities on the battlefield?

Autonomous weapons can help addressing all those issues. If it is achievable to make autonomous weapons conforming IHL as well as, or even better than, our soldiers, this would result in a reduction of collateral damages, translating into saving innocent lives. It is reasonable to believe that, in the future, robots will be able to ultimately treat us more humanely with respect to human warfighters. Weapons have no instinct to protect themselves and thus they can be used in a self-sacrificing manner, meaning that future weapons system will be able to act with a "first-do-not-arm" approach rather that a "shoot first, ask questions later" one. This is something that soldiers are taught in but, in practice, tend to not apply because of the self-preservation instinct. Moreover, the development of more sofisticated weapons may help in percieving better the fog of war, meaning that weapons will know better the battlefield, resulting in more precision and less non-combatant victims. In addition, machines can help

with the "scenario fulfillment" problem because they do not need to be vulnerable to the humans' pre-existing belief patterns. Machines can also monitor ethical behaviour from all parties on the battlefield, reporting any violation, thus reducing human ethical infractions.

3.2.3 Addressing counter-arguments

Of course, counter-arguments have been raised, especially about if autonomous weapons will be able to adhere IHL framework. However, simply stating that these systems will never be able to comply such ethical and legal requirements is wrong. If that was so, we would not have other technologies like self-driving cars. Some researchers see no scientific barrier in the creation of autonomous weapons. Of course, the design, development and deployment of such systems is not a short-term goal, but indeed will take considerable time and effort.

3.3 How can we reduce human atrocities?

To eliminate human atrocities we must look to other forms of intelligent autonomous decision-making in the conduct of war, which can be able to compy with the laws of war better. We should not let fear and ignorance rule our decision regarding autonomous weapons, but we must also proceed very cautiously and not rush the design, development and deployment of these systems without crucially examining their consequences on all parties, especially civilians and society in general and this can only be done through a discussion over the issues of these weapons. If these systems can be properly developed to outperform human soldiers in adhering the IHL, the end product can then be a reduction of non-combatant deaths. To do this, we must continue exhamine the development and deployment of autonomous weaponry. These systems can counterintuetively make warfare safer in the long run for the innocents in the battlespace, if coupled with morality, narrow situational use and careful graded introduction.

4 Concerns about the use of AWS in warfare situation

Researchers and Human Rights experts have also raise some concerns related to the use of AWS in warfare situations, mainly related to the fact that weapons that have high levels of autonomy and that are capable to decide the target to engage and to release force may be against human dignity, as well as they may not be able to comply with IHL and that their use would produce an accountability gap.

5 AWS create an accountability gap

Many researchers have raised concerns about the lack of accountability in case an autonomous weapon acts unlawfully. In fact, researchers have objected that these weapons may never be able to distinguish between enemies and civilians or, more generally, combatants and non-combatants. In such case, the deployment and use of these weapons may produce unethical behaviour and innocent victims, also creating an accountability gap. A fully autonomous weapon could commit a criminal act (like a war crime), but it would lack the mental state to make these wrongful actions prosecutable crimes. These robots would not fall under the natural person jurisdiction of international courts and, even if such courts were amended to encompass a machine, a judgment would not fulfill the purposes of punishment for society or the victim because the robot could not perceive or appreciate being "punished." At this point, we must try to find a person accountable for the unlawful acts of a fully autnomous weapon. It is very difficult, however, to assign to a specific person direct responsibility for the unlawful acts of an autonomous weapon. In fact:

- programmers and manufacturers cannot be held accountable because they might lack the military understanding of the circumstances or they might not know variables the robotwould encounter and respond to and this diminishes the likelihood that it could be proved they developed the weapon to be unlawful
- the autonomous nature of killer robots would make them legally analogous to human soldiers in some ways, and thus it could trigger the doctrine of indirect responsibility, or command responsibility, however, given that the weapons are designed to operate independently, a commander would not always have sufficient reason or technological knowledge to anticipate the robot would commit a specific unlawful act and it is very unlikely that he will be held accountable; even if he or she knews of a possible unlawful act, the commander would often be unable to prevent the act, for example, if communications had broken down, the robot acted too fast to be stopped, or reprogramming was too difficult for all but specialists. In addition, "punishing" the robot after the fact would not make sense.
- it is difficult to held the operator accountable because he or she cannot predict perfecly the behaviour of the AWS in certain scenarios

An alternative approach would be to hold a commander, an operator, a state, a programmer or a manufacturer liable for negligence if, for example, the unlawful acts brought about by robots were reasonably predictable, even if not intended, but regardless of the nature of the penalties, but this would also be very unlikely since the victim must start a civil or criminal lawsuit and this could be very difficult because:

- helding accountable the manufacturer for any harm caused by its autonomous weapon is hard from a legal point of view because the victim

would have to do a civil lawsuit which will be very hard to arry on due to jurisdiction (the court must have jurisdiction over the manufacturer) and national laws could mitigate the responsabilities of the manufacter

- the nature of fully autonomous weapons would make very complicate to hold a lawsuit in court because the opinion of developers and engineers will be needed to help the court and the other legal figures to fully understand the fact
- helding lawsuits against manufacturer could be very expensive and time-consuming for civilians, while manufacturers and military parties possess far more financial capabilities
- military contractors often enjoy immunity
- products liability law, the most common method of imposing civil liability on manufacturers,cannot adequately accommodate claims regarding autonomous devices

However, it is possible to held accountable programmers or manifacturers when they negligently or intentionally contribute to delivering a device that either would not achieve the intended function or would achieve the intended function, but this function necessarily entails unacceptable consequences. Similarly, the commanders deploying a well-functioning automated device may be blameworthy for the harm caused by the device, when either the device is used outside the circumstances where it should have been deployed or the device is used for its designed function, but this function necessarily involves unacceptable consequences. The same thing is valid for operators and states. In the end, fully autonomous weapons would not fit well into the scheme of criminal liability designed for humans, and their use would create the risk of unlawful acts and significant civilian harm for which no one could be held criminally responsible. If a system fails to possess behavioural competence, epistemical competence, practical reasoning or moral skills, then those who designed and deployed the system may be blameworthy and even criminally liable. It is needed to fill this accountability gap in order to deter future violations, provide retributions to victims and provide compensatory justice. These concerns are strictly related to the use of fully autonomous weapons because semi-autonomous and human-supervised autonomous weapons are still a tool in the hands of a human and so the human controlling these weapons must be held accountable.

5.1 AWS may not comply with IHL

When it comes to war, military parties must respect IHL limits. Of course, this is valid also for autonomous weapons: when a military party decides to deploy an autonomous weapon on the battlefield, this weapon must have been developed to comply the four principles of war given by IHL. Therefore, AWS must be limited in order to respect such regulamentations. Arguments are made over the fact that machines do not have and will probably never have human

judgment, human interpretation of laws and other human skills that are needed on the battlefield. The four principles of war are:

- **military necessity**: the release of force must be done in order to accomplish a military objective and based on the situation; in this case, AWS must be allowed to use a limited amout of force in ottder to avoid unnecessary killings or limit the collateral damages; in order to do this, the weapon must be fully aware of the military situation and thus it could be very difficult to respect such a principle; moreover, the need of a state to win the war deploying AWS is not considered a military necessity
- **distinction**: soldiers must be able to distinguish between combatants and non-combatants; soldiers usually do not respect this principle because they're put in stressful situations, however also AWS may not be able to correctly distinguish between combatants and non-combatants because it is very difficult to translate in rules and codes what a civilian or a non-combatant is, mostly because civilian are usually described using what a civilian is not (Protocol 1 of Geneva Convention); moreover, some systems are able to distinguish between friends and enemies, while other systems may not be able to do so and may not be able to identify other immune actors such as wounded combatants or those who have surrendered
- **proportionality**: the collateral damage of the use of force should be proportional with the military advantage obtained though them; this principle is context-dependant and AWS nowadays doesn't seem to have a metric to assess the proportionality correctly; this is something that is also very difficult for humans to do; humans are also unable to predict any warfare situations and "code" them in the AWS; AWS may be able to comply this principle only if deployed in situations were the collateral damage is known to be small or absent
- **humanity**: this principle forbids the infliction of unnecessary injuries, suffering or destruction; as said before, thanks to the precision of AWS systems, these weapons may be able to comply with this principle, however some experts are still concerned about this

5.2 AWS may violate human dignity

There also concerns that AWS, even if able to comply with IHL, will be against human dignity. Most researchers share the idea that AWS will be against human dignity; they claim that machines do not understand or comprehend the value of life and the significance of its loss and, for this reason, machines should not decide wether to take someone life. Moreover, they violate human dignity because the person who is being engaged by these weapons may not be able to appeal to the humanity of the one targeting. Dignity of the ones using these weapons is also compromised because the opportunity to make moral decisions is lost when using these weapons. Human dignity is defined by researchers in

different ways but all those definitions share the fact that the right to human dignity contains some rights like the right to not be severely humiliated, the right of a minimum of freedom of acitons and decisions, the right to recieve support in situations of severe need, the right to a normal life worth of living and the right to not be treated merely as a means to other people's end. All these rights are affected by the development and deployment of AWS, posing threats to civilians and preventing a minimal quality of life. AWS could affect all of the set of basic human rights listed before. It can be claimed that war and killing in general are against human dignity and so that AWS would be contrary to human dignity as well as any other lethal weapon, even those entirely operated by humans. However, death by algorithm crosses a moral line and is against human dignity because AWS will be unable to understand the value of life.

5.3 Other concerns

Researchers and experts are also concerned about the fact that these machines are inevitabily going to fail in hardware, software and reasoning processes.

Moreover, the deployment of these weapons will make war not likely since states can instigate conflixts without risking their nation's human soldiers.

AWS can also create psychological stress over the population that is currently threatened by the possibility of the deployment of such weapons.

These concerns are stricly related to the use of fully autonomous weapons because with semi-autonomous and human-supervised autonomous weapons they are still a tool in the hands of a human and, for this reason, the human controlling these weapons must be held accountable in case of a committed war crime.

6 Autonomous weapons in law enforcement

Since military technology often is transferred to the law enforcement environment, it is very likely that, in the future, military forces will adopt AWS also for policing operations. States are currently testing the deployment of such weapons in law enforcement operations while various companies around the world are developing AWS that can be used in law enforcement situations, advertising them as tools for police and internal security. The population, the researchers and the human right experts are complaining about this possibility. These weapons differ from the ones used in war scenarios because they're called "non-lethal", since they might be used to break up riots or, more generally, inhibit subjects without the risk of severely hurting or killing them. However, researchers argue that these kind of weapons should be called "less-lethal" instead, because they can still cause accidental death. They also argue that this weapons are also against human rights and human dignity. In fact, it seems that the use of such weapons in law enforcement operations may have very few advantages with respect to the high number of concerns that they may raise. Moreover, a lot of issues raised from the usage of AWS in warfare are also present when using

AWS in internal security operations and, in some cases, they are even more alarming. When it comes to law enforcement, the framework regulating the use of force in these situations is the International Human Rights Law (or IHRL for short) and such weapons must be able to fully comply the limits imposed by this framework in order to be used in these scenarios.

For example, NYPD has tested the deployment of a BostonDynamics' Spot robot in their policing operations, raising not few complains. https://www.nytimes.com/2021/04/14/nyregio dog-nypd.html

A texan company has developed a drone with a 80000 volt stun gun that can be used in policing operations, whileaA mining company has ordered 25 unmanned aerial vehicles for breaking up riots with pepper spray and "blinding lasers." https://www.cnet.com/news/useful-the-drone-with-an-80000-volt-stun-gun/ https://www.cnet.com/news/pepper-spray-and-laser-armed-drone-receives-first-orders/

6.1 Autonomous weapons in law enforcement: advantages

The advantages related to the usage of AWS in law enforcement are very few and these usually only advantage the ones using such weapons and not both parts. These advantages are, like the warfare usage of AWS advantages, related to the increased precision of these weapons and to the possibility of hurting less innocents and police officiers. Another advantage of this application of AWS in, for instance, a hostage situation, is that since the precision of these machines could help in hitting only the hostage takers without risking to hurt the hostages.

6.2 Autonomous weapons in law enforcement: concerns

The use of AWS in law enforcement raises even more concerns than ones raised when talking about warfare. In fact, in both cases, international human rights standard are not respected and Certain Conventional Weapon (CCW for short) convention and other groups urge a discussion. Amnesty International has identified 5 key points to address the issues related to the use of AWS in law enforcement operations and, more generaly, non-war contexts.

6.3 CCW doesn't cover law enforcement

The scope of CCW convention only covers warfare situations, avoiding situations of internal disturbances and tensions, excluding this way the potential law enforcement situations in which these weapons can be used, like riots. States and UN bodies started years ago the discussion on the usage of AWS. Experts like Christof Heyns submitted reports in order to develop a moratoria on each phase of the production and deployment of AWS, emphasizing that both lethal and non-lethat AWS systems would harm human rights both in warfare and law enforcement. However, the main focus of international discussions was the armed conflict, with few or no mentions to the use of these systems in law

enforcement. CCW members and states should so work together with human rights experts and arms control experts to also encourage the discussion over the usage of such weapons in internal security.

6.4 AWS will not comply IHRL

The international community has elaborated various standards to guide states in ensuring human rights compliant use of force in law enforcement. Any decision regarding the use of force in these scenarios must be inline with international standards. These principles for the use of force in law enforcement are:

- **legality**: the use of force needs to serve a legitimate objective established by law; the domestic legislation must be in line with IHRL and standards; domestic legislation regarding the use of force in law enforcement must be carried out without any discriminatory bias
- **necessity**: officials must prefer non-violence before resorting the use to the use of force and, if force is required, they must also decide which kind, how much and for how long to use force, in concomitance with the objective
- **proportionality**: whenever the lawful use of force is inevitable, officiers must act in proportion to the seriousness of the offence and the legitimate objective to be achieved, balacing the benefits of the use of force and the potential consequences and harm
- **accountability**: not only the ones responsible for the unlawful acts must be held accountable, also their superiors, supervisors or the whole law enforcement agency must be held accountable
- **distinction**: officiers must distinguish, when using force, between the civilan population and the combatants

AWS need to be technically capable of following and performing the Legal obligations under existing law. IHRL require officiers to try resolve any situation through other means that the use of force, such as persuation, negotiation and de-escalation. In order to do this, human empathy, negotiation skills, crowd behaviour understanding, methods of persuasion etc. are required. All these skills cannot, of course, be held by autonomous weapons, therefore the use of AWS in non-lethal contexts is incompatible with these principles. They can't deal with complex situation and manifestations or law enforcement operations because it is hard to code such skills. There are even situations in which it is lawful for police to use lethal force and the use of this kind of force may be decided using human skills like identifying correctly who is posing the threat, deploy different modes of communications to allow a graduated response and the ability to respond with a unique response to a unique situation. Therefore, for the same reasons described above, it is very unlikely that machines will be

able to deal with such complex situations mantaining a lawful behaviour. Moreover, lethal AWS can be used by states to suppress liberty of freedom and to suppress domestic enemies.

6.5 Existing semi-autonomous weapons pose challenges for the IHRL

Various companies around the world are developing robotic weapons that are not yet fully autonomous for law enforcement purposes. These weapons can be equipped with stun guns or fire chemical irritants and rubber or plastic projectiles. These kind of existing weapons, even if not fully autonomous, could still be used to violate many human rights and it is more likely that, in the future, AWS systems that are used in warfarer will be translated to law enforcement situations. Even if these weapons are not yet fully autonomous, it is only a matter of time before this happens, and this raises serious concerns over the level of autonomy that these weapons should incorporate.

6.6 AWS must be subject to independent reviews

Article 36 of the Geneva Conventions demand to High Contracting Parties (the states that signed the treat) to review new warfare systems and methods.

The Article 36 doesn't neither specify how to review nor define a standard for such review moreover states have no obbligation to show how review are done nor the results of such reviews.

The Article 36 involve weapons in armed conflict and not in law enforcement operation therefore could happen that weapons lawful in war are applied in law enforcement operations without being examined for policing tasks.

Therefore if AWS after being evaluated as lawful for war standards, are adopted in policing operation such system should comply with all the standards applied in this kind of situations.

Moreover non-lethal systems should be classified as less-lethal due to the injuries that such weapons could cause to the human body (teasers, rubber bullets,...)

In conclusion AWS lethal and less-lethal should pass independent reviews as well as comply with international laws in order to be used.

But the sentiment in general should be to not use this weapons in policing operation d due to uncertainity, unpredictability and complexity that this situations pose on AWS.

6.7 AWS erode accountability mechanisms

The accountability of AWS must be discussed since allow to have lawfull systems: if such feature is not present the right of life is violated and therefore the system doesn't comply anymore with any legal and ethical framework.

In particular, The Principle 22 declare that individual officers, commanders and the entire police organization should be accountable for human rights

violations as well as law violations by independent, impartial and effective investigations in order to fullfil the obligation of states to respect the right to life and dignity prohibiting such violations (if this is not respected the rights are violated) with investigations, punishments and redressing (e.g. compensation to victims) it.

Due to impossibility to account a machine of unlawful actions, actors involved, such as developers, officers ,superior officers should be accountable for unlawful behaviour of AWS.

Moreover due to the complexity and uncertinity of law enforcement operation and the impossibility to predict the behavior of AWS for such actors it's hard to forse and then stop AWS.

This produce the gap in accountability described by Human Rights Watch (¡¡ the actions of fully autonomous weapons would likely fall within an accountability gap that would contravene the right to a remedy ¿¿) that will cause the impossibility to satisfy the right to a remedy (to bring to justice humans that violate law or human rights) therefore victims will not be repaired by the harm suffered with compensation, rehabilitation or with the guarantee that such violations will not be repeated.

Moreover due to non-transparency of states investigation of AWS killing will be very difficult.

6.8 Violations of AWS

The rights violated by AWS systems could be:

- the right of the bodily integrity which includes **right to life**, **right to security**, **right against cruel, inhuman and degrading treatment**
  - right to life is violated with the loss of life but also when accountability is not provided when this action is done.
  - right to security when life and non-life threats are done by AWS systems to human being
  - right against cruel, inhuman and degrading treatment
- the right of dignity is violated from different point of view
  - those injuried or killed by AWS have the right to life, security violated as well as their life is affected by an entity unable to do moral judgemnts therefore unable to be moral hence by an agent that doesn't have dignity which act in name of the law: in this sense these people are deprived by thei dignity since are an objective of the law.
  - those who decide to use the AWS will have their responsability taken off by a machine therefore they will be unable to act morally and hence exercise their dignity
  - those caught in the crossfire for the same reasons as before

Other violations are:

- Principle of necessity since machine will use lethal or less-lethal force. Therefore they can't apply this principle since they don't have any other instruments to achieve it.
- Principle of proportionality will be violated since AWS are going to apply the indication learned or imposed during the development unable to react accordingly the situation in balancing the use of force. Moreover AWS will be unable to act on very unusual situation due to inability to learn from few experiences as human can do.
- Principle of distinction is also violated due to inhability to distinguish offender versus non-offender, moreover they will be unable to distinguish a surrender or act accordingly using personal information of the subject (such as the personal history, education, criminal history, ...) and therefore use **empathy**: machine mimic data not the human reasoning.
- Principle of accountability is of course violated violating automatically the **right to dignity**, the **right to remedy**, **the right to justice**

7 Regulation of AWS

There is no international consensus on laws regulating these weapons. The lack of legistlation poses real risks, since states can deploy untested or unsafe weapons that may harm civilians or, with the proliferation, AWS can be used unlawfully by criminals, terrorists or rogue states. The debate on regulating AWS has been centered around banning these weapons or not. However, there are not only binary options and some experts say that a full ban would foreclose the possibility to use AI to mitigate non-combatant arm. Moreover, there is a general desire to add a certain degree of human involvement in the use of these weapons. Arkin et al. propose a solution to this problem by considering five components.

7.1 Component 1: time-limited moratorioum on the development, deployment, transfer and use of anti-personnel lethal AWS

States should consider adopting a 5-year renewable moratorium on the development, deployment, transfer and use AWS able to select and engage human targets. This moratorium can exclude system as:

- fixed-point defensive systems to defend human-occupied bases
- automated counter-file systems that return fire in order to provide immediate, local defense of humans
- time-limited pursuit deterrent munitions or systems

- AWS that select as targets hand-held weapons or man-portable air defense systems, provided adequate non-combatants protection and ensuring IHL compliance
- anti-vehicle or anti-material weapons
- non-lethal anti-personnel weapons
- research on ways of improving autonomous weapon techonology to reduce non-combatant harm
- weapons that find, track and engage specific individuals whom a human has decided should be engaged

This moratorium would pause the development and deployment of anti-personnel lethal autonomous weapons systems to allow state to better understand their risk and improve their safety and effectivenes. Some objectives are:

- ensure that anti-personnel AWS can be usen in compliance with anti-vehicle
- lay he groundwork for a potentially legally binding diplomatic instrument
- decrease the geopolitical pressure on countries

To ensure compliance, states could adopt various approaches:

- developing an industry cooperation regime anologous to the one mandated under the Chemical Weapons Convention
- encouraging states to declare tehir inventory of AWS for transparency purposes
- facilitate scientific exchange and military-to-military contracts to increase mutual transparency and understanding on topics such as compliance verification and safety
- designing control systems to require operator identity authentication an records of operation, enabling checks in case of plausible non-compliant attacks
- designing weapons with firing authorization circuits that are connected to remote human operator
- avoid design of weapons that allow conversion from compliant to non-compliant behaviour through software updates
- designing weapons with formal encrypted proof of relevant properties (e.g. a weapon is unable to initiate an attack without human authorization)
- facilitate access to AI resources to all states that remain in compliance and stays transparent

7.2 Component 2: define universal guiding principles for human involvement in the use of force

- humans are legal and moral agents in military operations
- it is a human responsibility to ensure that any attack, including AWS attacks, complies with the laws of war
- humans responsible for initiating an attack must have sufficient understanding of the entire context and weapon to determine whether that particular attack is lawful
- the sttack must be bounded in space, time and target class in orrder for the determination about the lawfulness of that attack
- militaries must invest in training, education, policies, system design and human-machine interfaces to ensure that humans remain responsible for attacks

7.3 Component 3: develop protocols to mitigate the risk on unintentional escalation

- developing safe rules for AWS behavior in proximity of adversarial force to prevent unintentional escalation of force
  - no first-fire policy: AWS do not initiate jostilities without human authorization
  - a human must always be responsible
  - clearly distinguish peacetime military operations from attack in order to limit the possibility of reactions from adversary AWS
- deleop communication links to ensure recallability of autonomous systems in case of unauthorized behaviour; militaries should also refrain from jamming others' ability to recall their autonomous systems

7.4 Component 4: develop strategies to prevent proliferation to illicit uses

- multilateral control over large-scale sales and transfers of AWS and related components
- mesures to make AWS less harmful (hard-wired kill switch, non-reprogrammable hardware)

7.5 Component 5: conduct research to improve technologies to reduce non-combatant harm and ensure IHL compliance

- strategies to promote human moral engagement in decisions
- risk assesssment for AWS, including large-scale effects, geopolitical destabilization, and lowering thresholds to initiating conflict and for violence within conflict
- methodology to ensure reliability and security of AWS
- new techniques for verification, validation, explainability, characterization of failure conditions