

Targeting people

Key issues in the regulation of autonomous weapons systems

Convention on Conventional Weapons (CCW)

November 2019

Article 36 is a UK-based not-for-profit organisation working to promote public scrutiny over the development and use of weapons.*

www.article36.org
info@article36.org
[@Article36](https://twitter.com/Article36)

* This paper was written by Maya Brehm

*'We shape our technologies and our technologies shape us'*¹

Since states parties to the Convention on Certain Conventional Weapons (CCW) framed their deliberations in terms of 'emerging technologies in the area of lethal autonomous weapons systems', questions have been raised about whether the reference to 'lethality' indicated a particular or, perhaps, an exclusive concern with anti-personnel, as opposed to anti-matériel technologies or applications. Ethical considerations, definitions of autonomous weapons systems (AWS) put forth by certain actors and increasing concern about 'algorithmic bias', have recently reinforced the sense that stakeholders in the debate on AWS should pay focused attention to issues raised by the prospect of targeting people.

This briefing note:

- x identifies a number of concerns raised by the targeting of people through AWS
- x argues for a prohibition on the targeting of people through AWS
- x stresses that an anti-personnel prohibition is not sufficient to address the wide range of urgent concerns raised by AWS

Targeting people in war

In keeping with the CCW's mandate and given the absence of a widely shared understanding of what falls within the scope of 'emerging technologies in the area of lethal autonomous weapons systems', respectively, '(fully) autonomous weapons' or 'killer robots', this paper is broadly concerned with the use of armed force *in the conduct of hostilities during an armed conflict* where those who plan or decide upon an attack do not stipulate specific times and places of force applications. Instead, the human decision to attack is implemented through sensors that acquire data in the world, algorithms that process and classify sensor data according to pre-encoded profiles of targets, and actuators that apply kinetic or other force to targets – *without* human evaluation of the sensor data prior to force application.² For the sake of simplicity, we call such human-machine configurations 'AWS'. In relation to targeting people, AWS could detect and track objects using cameras, ultra-sonic sensors, etc. that capture certain contours, features, gestures or similar 'proxy indicators' used to represent human beings.

Who may legally be attacked in military combat is primarily determined with reference to the rules of international humanitarian law (IHL) governing the conduct of hostilities.³ In principle, the only people that may legally be attacked are members of the armed forces of a party to the conflict, that is, members of a state's armed forces (combatants) and members of a non-state organised armed group who have a continuous combat function. All other people are civilians and protected against direct attack, *unless* and *for such time* that they take a direct part in hostilities.⁴

Parties to conflict must take 'constant care' in the conduct of military operations to spare civilians. They must also safeguard persons placed *hors de combat*, that is, anyone who is in the power of an adverse party, is defenseless because of unconsciousness, shipwreck, wounds or sickness, or who has clearly expressed an intention to surrender. Conducting hostilities on the basis that there shall be no survivors (denial of quarter) is prohibited. Attackers must do 'everything feasible' to verify that targets are not protected under the law

against direct attack, and to choose means and methods of attack and targets so as to avoid, and at any rate, minimise civilian harm. An attacker must be in a position to assess the expected civilian harm resulting from an attack and suspend or cancel it to prevent disproportionate civilian harm. If there is doubt about whether a person is a legitimate target of attack, the person must be presumed to be protected against direct attack.

Although these rules of IHL are generally accepted, there are many open theoretical questions and disputes about their application in practice. The legal intricacies need not detain us here, but any discussion about AWS and the targeting of people must have regard to the fact that key legal questions are as yet unresolved.⁵

Concerns raised by the algorithmic targeting of people

The prospect of using AWS to direct force against human beings raises a set of inter-related moral and legal issues, including:

- x the risk that the ‘wrong people’ (e.g. civilians or combatants *hors de combat*) may be targeted
- x questions about how measures adopted to manage the risk of undesired consequences might affect the normative protection of people in the long-run
- x procedural concerns about the *process* of targeting, that is, how and why a person is made the object of attack or harmed

Without aiming to be exhaustive, these issues are sketched out below.

Targeting the wrong people

A key objection to the targeting of people is the fear that a human being may be classified as a target in spite of there being no moral justification for targeting them or them being protected from attack by law.

So-called ‘signature strikes’, drone attacks that target ‘groups of men who bear certain signatures, or defining characteristics associated with terrorist activity, but whose identities aren’t known’,⁶ are sometimes drawn upon to assert that valid targets of attack can be identified solely on the basis of sensor data and computations. In present practice, human agents intervene in the drone targeting process before and during every force application, but ‘signature strikes’ have been challenged on grounds that point to concerns with the algorithmically mediated targeting of people more generally:⁷

- x **The categories used to identify targets of attack do not map sufficiently onto the legal definitions of persons that may be the object of direct attack.** Whether abstract legal concepts, such as ‘membership in an organised armed group’, ‘direct participation in hostilities’ or ‘hors de combat’ can be adequately encoded into algorithmic calculations is highly questionable. These are relatively fluid categories whose morally and legally salient features are not easily modelled or captured by sensor technologies. Interpreting and applying such concepts calls for evaluations against legal and moral standards that do not predict the right outcome in any given situation, but instead demand context-specific value-judgments, in light of concrete circumstances.

Detection of a firearm, for example, would not be sufficient to assign the person handling it to the category of valid targets – not even in conjunction with data pertaining to their attire, gestures, height, location or similar.

It has also been pointed out that the categories underpinning ‘signature strikes’ are gendered and racialised with the result of adversely affecting certain people’s protection from direct attack, notably, civilian men of ‘military age’.⁸

- x **Force applications are based on insufficient evidence that the targeted persons in fact exhibited characteristics or behaviors that would justify attack.** The detection and tracking of humans with autonomous systems is technically challenging in real world situations. What would be the evidentiary requirements to reverse the presumption of civilian immunity from direct attack? Could the necessary level of certainty obtain in real situations (‘the fog of war’ and ‘heat of battle’), equipment malfunctions and adversary interference?

Wrongs resulting from risk management

The risk that AWS use may result in the wrong people being targeted is often cited as a ground for legal constraints or a comprehensive ban on AWS. Such propositions are vulnerable to two types of counter arguments, however, that point to further structural concerns about the normative protection of people from algorithm-based targeting:

- x **‘Work-arounds’:** if technical limitations make the use of AWS problematic in the real world, the world is altered so as to reduce the likelihood of challenging circumstances arising. For example, if it cannot be guaranteed that force will only be directed against persons who may legally be targeted, one could remove all people from the space where an AWS is used and tightly control access to the zone.⁹

This type of risk mitigation is common in weapons control, for example, legal requirements on the marking and fencing of minefields. But making space for AWS in this manner bears the risk that persons finding themselves in an AWS ‘danger zone’ will be expected to demonstrate that they are not targetable, rather than the AWS-user having to demonstrate that the people are in fact legal targets of attack.¹⁰ Work-arounds can thus lead to subtle but significant shifts in protective presumptions with detrimental consequences for the protection of all victims of war in the long-run. The adoption of work-arounds also tends to obscure that, in contrast to measures aimed at mitigating risks from existing weapons, in the context of emerging weapon technologies, the expectation ought to be that new weapons enhance our capacity for moral and legal conduct.¹¹

- x **‘Imaginary techno-fixes’:** an, as yet unspecified but more sophisticated technological solution situated somewhere in the future holds the promise of alleviating challenges. A prime example of this is Arkin’s ‘Ethical Governor’.¹² More generally, the propositions that highly advanced AWS would be acceptable if (read: when) they could reliably distinguish between civilians and combatants, assess proportionality or recognise ‘hors de combat’ situations also smack of techno-solutionism.¹³

Such suggestions imply that moral conduct and the interpretation

and application of law are equivalent to algorithmic computations. This provides both a poor representation of how norms govern behaviour and anthropomorphises machine processes. Applying technological solutions to solve social or political problems suggests that we can change the way things function without questioning the power structures embedded in technologies. But if disarmament is to realise a more just and peaceful world, it is critical that we understand weapons systems in a broader social and ethical context, and question how they reify and extend the root causes of political problems and ethical predicaments.¹⁴ Today's disarmament policy-making should therefore not be guided by future techno-fixes with quasi-magical properties.

Targeting people wrongly

Still more fundamental objections to the targeting of people concern how and why a person is made the object of attack or harmed.¹⁵ Objections of this type concern the process of killing through AWS in principle, irrespective of its results, and raise questions that are not easily amenable to techno-fixes or work-arounds.¹⁶ Arguments of this type take many forms, some aspects of which are highlighted below.

Unreasoned

'For the killing of a human to be meaningful, it must be intentional. That is, it must be done for reason and purpose' and 'intentionality requires understanding the meaning and significance of an act'.¹⁷ An unspoken assumption that 'the decision to use lethal force must be reasonable and taken by a human'¹⁸ is reflected in legal and moral codes, including the IHL requirement that 'those who plan or decide upon an attack' take certain decisions about who to target and how.¹⁹

Such decisions cannot therefore be 'delegated' to an artefact. The concern with algorithmically mediated violence is, however, that the connection between the intention of human decision-makers, morally responsible agents, their actions and the consequences experienced by other human beings would be 'eroded' when a potentially harmed individual is no longer considered in real time, but 'factored' into a pre-determined process.²⁰

To be morally acceptable, a decision-making system needs to 'track (relevant) human moral reasons', that is, it must 'demonstrably and verifiably be responsive to the human moral reasons relevant in the circumstances'.²¹ IHL demands particular attention to the reasons of those who plan or decide upon an attack (military commanders, rather than sensor technicians or program designers). IHL provides for doubt to operate and requires ongoing adjustments based on what is deemed 'reasonable' and 'feasible' in the light of prevailing, evolving and complex circumstances. Those human beings who decide upon an attack have a choice to attack or not to attack, even in cases where a person may fit the description of a valid target,²² and they have 'a duty to constantly exercise discretion'.²³

Weapons systems must be designed so as to allow human agents to engage in reflexive, 'deliberative reasoning'²⁴ and enable them to 'make meaning and take meaningful actions'.²⁵ Fettering one's discretion through the pre-binding of decisions in the form of algorithms is unacceptable.²⁶

Unaccountable

Closely related to the above is the concern that it may not be possible to provide adequate reasons or ethical justification for the targeting of or harm done to an individual, resulting in an accountability gap.²⁷

The 'process of reason-giving and justification also establishes moral responsibility, and makes people feel they are treated justly'.²⁸ Especially where algorithm-based decisions have potentially irreversible consequences, it is essential that human agents can provide a concise, intelligible account of how input features relate to predictions and categorisations. This is a prerequisite for the legal review and moral justification of the use of force in concrete circumstances, to identify and punish those responsible, and to assist victims and society at large in their quest for the truth. Present-day algorithms present important challenges in this regard.²⁹

To be morally acceptable, a decision-making system's actions need to be 'traceable to a proper moral understanding on the part of one or more relevant human persons who design or interact with the system'.³⁰ Attributing formal (legal) responsibility to a person who cannot properly understand and exercise sufficient control, is morally unfair.

Unjust

Targeting people on the basis of pre-programmed profiles can produce unjust outcomes, adversely and disproportionately affecting certain individuals. Paralleling the prohibition of discrimination in human rights law, adverse distinction in the application of IHL based on race, colour, sex, language, religion or belief, political or other opinion, national or social origin, wealth, birth or other status, or on any other similar criteria is prohibited. This applies to all persons protected by IHL, including in relation to the conduct of hostilities.³¹

Algorithm-based targeting of people is problematic from this perspective on several accounts. As mentioned above in relation to 'signature strikes', adverse distinction among civilians based on characteristics such as gender or age can result in military-aged civilian men being denied protection against direct attack equal to that afforded to non-male civilians. It is also arguable that differentiations among people not protected against direct attack amount to adverse distinction in that the prohibition against denial of quarter, the prohibition on weapons that render death inevitable, and the privileges owed to persons *hors de combat* are not applied equally to all combatants.

Finally, joining wider procedural concerns touched upon above, differentiations – whether intended or the unexpected result of patterns reflecting structural biases – demand a reasonable and objective justification. Differentiations based on characteristics that the targeted individual cannot easily change, such as their sex or gender (as opposed to the wearing of certain insignia) are especially hard to justify. Our difficulty in providing a reasoned explanation for algorithm-based decisions in concrete circumstances and the prevailing corporate and military secrecy resulting in opaque 'black-box' algorithms cast serious doubt on the capacity of AWS-users' capacity to justify differential treatment.³²

Inhumane

Whether based on the considerations set out above or on different grounds, the algorithmic targeting of people as envisaged by proponents of AWS has been criticised as being inhumane, an affront to

human dignity and repugnant to human conscience. Lack of respect for human dignity tends to be cited among the main reasons for objecting to AWS, both, by members of the military and the general public.³³

Targeting people through an AWS fails to take account of potentially harmed individuals as fellow human beings, equal in worth and dignity, is dehumanising and objectifies human beings: 'They become zeros and ones in the digital scopes of weapons which are programmed in advance to release force without the ability to consider whether there is no other way out, without a sufficient level of deliberate human choice about the matter.'³⁴

Towards an anti-personnel prohibition...

At least some of the concerns about the wrongful targeting of people are 'overriding considerations'.³⁵ In light of the concerns about human agency, human dignity and moral responsibility being especially acute in relation to AWS that target humans directly, the ICRC has suggested that they 'perhaps preclud[e] the development and use of anti-personnel autonomous weapon systems'.³⁶

There is ample precedent for articulating legal constraints on weapons in terms of their intended or actual use against human beings as opposed to objects: the prohibition of exploding bullets, the ban on blinding laser weapons, the differential treatment of anti-personnel compared to other landmines, and the prohibition on the anti-personnel use of incendiary weapons under customary IHL illustrate this approach.

Support for an anti-personnel prohibition can also be found in certain national policies. A position paper by the German army published in October 2019, for example, raises particular concerns regarding autonomous weapon systems that are primarily intended for use against persons and explicitly excludes from this notion, among others, weapon systems that are primarily intended for use against objects, e.g. projectiles, aerial vehicles, tanks or ships.³⁷

...and beyond

In the debate on AWS, objections to anti-personnel weapons or applications are sometimes coupled with arguments aimed at excluding other systems and applications from the purview of potential future regulation. The aforementioned German army paper, for example, suggests that a prohibition of 'lethal' anti-personnel systems (a notion qualified by a number of other attributes) would be sufficient to prevent undesirable systems from being developed. Excluded from the notion are not only weapons systems that are primarily intended for use against objects, but also those that serve multiple purposes, or that apply force to people by 'non-' or 'less-lethal' means.

Such a narrow conception of what is problematic about AWS is unwarranted, however:

- x AWS raise more diverse issues than an anti-personnel ban can address. In addition to separate ethical, legal, military and security concerns, many of the objections set out above equally apply to anti-material weapons and applications, though, in that context, they may be justifiable in limited circumstances – which are yet to be defined.

- x The notion of 'lethality' should not be drawn upon to foster a regulatory response. As the ICRC pointed out, 'some anti-materiel weapons can also result in the death of humans either directly (humans inside objects, such as buildings, vehicles, ships and aircraft) or indirectly (humans in proximity to objects)'.³⁸ In the same vein, Switzerland has explained from the perspective of IHL-compliance that AWS should include 'means and methods of warfare that do not necessarily inflict physical death, but the effects of which may be restricted to causing, for example: (1) physical injury short of death, (2) physical destruction of objects, or (3) non-kinetic effects such as through cyber operations'.³⁹

When the prospect of a new weapons technology raises acute humanitarian concerns, the ethically appropriate course of action is to elaborate new rules and social and institutional practices to prevent moral and legal wrongs and ensure that humans retain meaningful control over the use of force. As this paper and work by others illustrates, AWS raise a range of fundamental ethical challenges. In light of the considerable social costs involved in making space for AWS, it is difficult to argue that their adoption is in the best interest of humankind. Thus far, moral and social considerations have had a difficult standing in the CCW. If the multilateral response to AWS is to align with our values, ideals and moral codes, however, ethical and related human rights and social considerations must take centre stage.⁴⁰

NOTES

- 1 S. Hankey and M. Tuszynski, Efficiency and Madness: Using Data and Technology to Solve Social, Environmental and Political Problems, 2017, <https://www.boell.de/en/2017/11/28/efficiency-and-madness-using-data-and-technology-to-solve-social-environmental-and-political-problems>.
- 2 'Target profiles', Article 36, 2019, <http://www.article36.org/wp-content/uploads/2019/08/Target-profiles.pdf>.
- 3 Attacks are defined under IHL as 'acts of violence against the adversary, whether in offence or in defence'. Most commentators consider that an attack 'as a whole' can encompass a series of 'force applications', multiple targets and weapons.
- 4 N. Melzer 'Interpretive Guidance on the Notion of Direct Participation in Hostilities Under International Humanitarian Law', ICRC, 2009, <https://www.icrc.org/en/doc/assets/files/other/icrc-002-0990.pdf>.
- 5 For example, there is ongoing debate about the meaning of 'direct participation in hostilities', the criteria used to determine a person's status and the evidence necessary to temporarily lift the presumption of civilian 'immunity from attack'. Additionally, the practical difficulty of distinguishing among members of the armed forces, civilians taking a direct part in hostilities and other civilians in contemporary wars cannot be overstated. Debate also persists on the interpretation of the prohibition to attack and the requirement to safeguard persons placed *hors de combat*, and controversy persists regarding the question whether a member of the armed forces (who is not *hors de combat*) may be attacked at any time (the so-called 'kill or capture' controversy).
- 6 K. J. Heller, 'One Hell of a Killing Machine': Signature Strikes and International Law', in 11(1) Journal of International Criminal Justice (2013) 89-119.
- 7 Ibid.
- 8 Sex and Drone Strikes: Gender and Identity in Targeting and Casualty Analysis, Reaching Critical Will and Article 36, 2014, <http://www.reachingcriticalwill.org/resources/publications-and-research/publications/9367-sex-and-drone-strikes-gender-and-identity-in-targeting-and-casualty-analysis>.
- 9 On such 'boxed autonomy', see e.g. Focus on Technology and Application of Autonomous Weapons, iPRAW, 2017, https://www.ipraw.org/wp-content/uploads/2017/08/2017-08-17_iPRAW_Focus-On-Report-1.pdf.
- 10 IHL rules aimed at preventing that 'evacuations' of civilians undermine civilian protection against forcible transfer are reflective of the risk of reversing protective presumptions. To give another example, if it cannot be guaranteed that all medical personnel will be correctly identified as non-targetable, the use of AWS could be facilitated by rendering these persons 'machine-readable'. In present practice, persons – whether military or civilian – who are exclusively assigned (permanently or temporally) to medical duties must be respected and protected in all circumstances. They should be identifiable and in areas where fighting is taking place or is likely to take place, civilian medical personnel should be recognizable by the distinctive emblem. If they commit, outside their humanitarian function, acts harmful to the enemy, they lose their protection. But neither the wearing of an enemy military uniform, carrying of light individual weapons, nor the failure to wear a distinctive emblem deprives medical personnel of protection. The emblem, therefore, is the visible manifestation of, but is not constitutive of the protection afforded to medical personnel under IHL. (Study on the Use of the Emblems, ICRC, (undated), p 30, <https://www.icrc.org/en/doc/assets/files/publications/icrc-001-4057.pdf>). The introduction of identification amenable to be captured by sensors bears the risk that medical personnel failing to present in machine-readable terms will be expected to demonstrate that they are non-targetable, rather than the AWS-user having to demonstrate that these people are legal targets at the time of a specific attack.
- 11 P. Asaro, 'On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making', in 94(886) IRRC (2012) 703; UN doc A/65/321, §48 (recommending proactive steps to ensure that new technologies are optimized in terms of their capacity to promote more effective compliance with IHL and IHRL); AfCommHPR, General Comment no 3, §35 ('The use during hostilities of new weapons technologies such as remote controlled aircraft should only be envisaged if they strengthen the protection of the right to life of those affected' [emphasis added]).
- 12 R. C. Arkin et al., 'An Ethical Governor for Constraining Lethal Action in an Autonomous System', CSE Technical reports. 163. <http://digitalcommons.unl.edu/csetechreports/163>.
- 13 On this concern, see E. Rosert and F. Sauer, 'Prohibiting Autonomous Weapons: Put Human Dignity First', in 10(3) Global Policy (2019), <https://onlinelibrary.wiley.com/doi/full/10.1111/1758-5899.12691>.
- 14 S. Hankey and M. Tuszynski, Efficiency and Madness: Using Data and Technology to Solve Social, Environmental and Political Problems, 2017, https://www.boell.de/sites/default/files/efficiency_and_madness.pdf. For a more general argument to this effect, see M. Brehm, Envisioning Sustainable Security: The evolving story of Science and Technology in the context of disarmament, Article 36, 2019, <http://www.article36.org/wp-content/uploads/2019/10/Envisioning-Sustainable-Security.pdf>.
- 15 'It matters why something is done, especially when decision have grave implications for an individual's life. The notion of 'meaningful human control' in this connection draws attention to human activities of meaning-making. Meaning-making is a social practice by which human beings interact with each other to make common sense of the world. Meaning includes moral understandings of right and wrong, cognitive understandings of true and false, perceptual understandings of like and unlike.' (M. Brehm, Presentation to the informal meeting of experts on lethal autonomous weapons systems of the Convention on Certain Conventional Weapons (CCW), Geneva, 14 April 2015, [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/897D-1C5358C307BDC1257E280028024B/\\$file/BREHM_Presentation+on+M-HC_14.04.2015.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/897D-1C5358C307BDC1257E280028024B/$file/BREHM_Presentation+on+M-HC_14.04.2015.pdf)) 'The central argument here is that it matters not just if a person is killed and injured but how they are killed and injured' (Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>).
- 16 The ICRC has remarked that 'the enduring ethical arguments against these weapons [which 'focus on autonomous weapon systems that are designed to kill or injure humans'] are those that transcend context – whether during armed conflict or in peacetime – and transcend technology – whether simple or sophisticated: (Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>).
- 17 P. Asaro, 'Jus nascendi, Robotic Weapons and the Martens Clause', <https://pdfs.semanticscholar.org/5706/1ce20febb4c58ab7e41c7e1463d352ba496b.pdf>.
- 18 C. Heyns, 'Human Rights and the Use of Autonomous Weapons Systems During Domestic Law Enforcement', 38 Human Rights Quarterly (2016) 362, fn 53.
- 19 US DoD, Law of War Manual, s 6.5.9.3, p 330, according to which 'The law of war rules on conducting attacks ... impose obligations on persons. These rules do not impose obligations on the weapons themselves; of course, an inanimate object could not assume an "obligation" in any event.' See also 'Towards a "Compliance-Based" Approach to LAWS', Informal Working Paper submitted by Switzerland, CCW Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Geneva, 11-15 April 2016, 30 March 2016, §16, [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/D2D66A9C-427958D6C1257F8700415473/\\$file/2016_LAWS+MX_CountryPaper+Switzerland.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/D2D66A9C-427958D6C1257F8700415473/$file/2016_LAWS+MX_CountryPaper+Switzerland.pdf) (noting that 'a manifest presumption of human agency' is reflected in a range of IHL provisions).
- 20 E. Lieblisch and E. Benvenisti, 'The Obligation to Exercise Discretion: Why Autonomous Weapons Systems are Unlawful', in N. Bhuta et al (eds), Autonomous Weapons Systems: Law, Ethics, Policy, 2016.
- 21 F. Santoni de Sio and J. van den Hoven, 'Meaningful Human Control Over Autonomous Systems: A Philosophical Account', in 5 Frontiers in Robotics and AI (2018), <https://www.frontiersin.org/articles/10.3389/frobt.2018.00015/full>. The authors provide the example of a system misclassifying an incoming rocket as not-a-target despite it being fired by the adversary, because the type of rocket was also used by friendly forces.
- 22 Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>.
- 23 E. Lieblisch and E. Benvenisti, 'The Obligation to Exercise Discretion: Why Autonomous Weapons Systems are Unlawful', in N. Bhuta et al (eds), Autonomous Weapons Systems: Law, Ethics, Policy, 2016.

- 24 N. Sharkey, 'The human control of weapons: a humanitarian perspective', in N. Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, 2016, <https://www.law.upenn.edu/live/files/3948-sharkey---human-control-of-weapons-pf-draftpdf>.
- 25 P. Asaro, 'Jus nascendi, Robotic Weapons and the Martens Clause', <https://pdfs.semanticscholar.org/5706/1ce20febb4c58ab7e41c7e1463d-352ba496b.pdf>. Although the precise procedural demands of IHL assessments and determinations remain nebulous in many regards, they demand timely adjustments and appear to assume capabilities of sensing, as well as capabilities of making sense (G. Noll, 'Analogy at War: Proportionality, Equality and the Law of Targeting', 43 *Netherlands Yearbook of International Law* (2012) 223).
- 26 E. Lieblich and E. Benvenisti, 'The Obligation to Exercise Discretion: Why Autonomous Weapons Systems are Unlawful', in N. Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy*, 2016. See also J. D. Herbach, 'Into the Caves of Steel: Precaution, Cognition and Robotic Weapon Systems Under the International Law of Armed Conflict', in 4(3) *Amsterdam Law Forum* (2012) 19, <http://amsterdamlawforum.org/article/view/277>.
- 27 Mind the Gap: The Lack of Accountability for Killer Robots, HRW, 2015, <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots>.
- 28 Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>.
- 29 Using force in a manner that renders an investigation into resulting deaths a priori incapable of determining whether force was justified in particular circumstances is, arguably, illegal. (M. Brehm, *Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems Under International Humanitarian and Human Rights Law*, Geneva Academy, 2017, p 67.)
- 30 F. Santoni de Sio and J. van den Hoven, 'Meaningful Human Control Over Autonomous Systems: A Philosophical Account', in 5 *Frontiers in Robotics and AI* (2018), <https://www.frontiersin.org/articles/10.3389/frobt.2018.00015/full>.
- 31 On the topic of adverse distinction and non-discrimination under IHL and IHRL, see in particular, G. Dvaladze, 'Non-discrimination under IHL and IHRL' in R. Kolb et al (eds), *Research Handbook on Human Rights and Humanitarian Law: Further Reflections and Perspectives* (forthcoming).
- 32 'Many of the fundamental building blocks required to understand AI systems and to ensure certain forms of accountability – from training data, to data models, to the code dictating algorithmic functions, to implementation guidelines and software, to the business decisions that directed design and development – are rarely accessible to review' (M. Whittaker et al, *AI Now Report 2018*, p 11, https://ainowinstitute.org/AI_Now_2018_Report.pdf).
- 33 Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>.
- 34 'Autonomous Weapon Systems: Human rights and ethical issues', Talking points: Christof Heyns, UN Special Rapporteur on extrajudicial, summary or arbitrary executions, Meeting of High Contracting Parties to the Convention on Certain Conventional Weapons, Geneva, 14 April 2106, [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/205D5C0B0545853BC1257F9B-00489FA3/\\$file/heyns+CCW+2016+talking+points.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/205D5C0B0545853BC1257F9B-00489FA3/$file/heyns+CCW+2016+talking+points.pdf).
- 35 UN doc A/HRC/23/47, §§92–93.
- 36 'Towards limits on autonomy in weapon systems', Statement of the ICRC, Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts on Lethal Autonomous Weapons Systems, 9 April 2018, <https://www.icrc.org/en/document/towards-limits-autonomous-weapons>.
- 37 'Ein letales autonomes Waffensystem (LAWS) ist ein Waffensystem, welches in erster Linie dazu bestimmt ist, tödliche Gewalt allein gegen Personen zur Wirkung zu bringen, und welches, ohne jegliche menschliche Einflussnahme und Kontrolle sein Umfeld und seinen internen Zustand wahrnimmt, eine Beurteilung der Situation vornimmt, entscheidet, handelt, evaluiert und daraus lernt.' (Künstliche Intelligenz in den Landstreitkräften, Positionspapier des Amtes für Heeresentwicklung, 2019, p 9, https://www.deutschesheer.de/portal/a/heer/start/aktuell/nachrichten/jahr2019/oktober2019/!ut/p/z1/hVDRasJAEPyavN5uWk1i3xKOGIHbGtqYeyIn-cl4i5104z6SUfnxPBKFQcR8GdnZmGBYobIaq1jWC2UYrJt1e00AzGA3S6UO-Gy5cgneBbFryHOTJGXC8QH5PQt0Zb0yMkFUcCpcR3s7wlQMkdm869kV-abazklrDy3BGKmqIK8lddxhfCCStOSq2cxqHlyjYOhWFWm6u7PBnjLqS-poEB_nODw2sn_SabrWT6lwtE4TdYwByqk3l6-EavtYySAGr7jhhtyMo6ur-W2PTx562Pc9EVoleW5w8PA_S62PFjZ_IdAenqPVath9L3j-C1LcHhs!/dz/d5/L2dBISevZ0FBIS9nQSEh/#Z7_694IG2S0M06IE0QS6U78MA00L1).
- 38 Ethics and autonomous weapon systems: An ethical basis for human control?, ICRC, 2018, <https://www.icrc.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>.
- 39 'Towards a "Compliance-Based" Approach to LAWS', Informal Working Paper submitted by Switzerland, CCW Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Geneva, 11–15 April 2016, 30 March 2016, §5, [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/D2D66A9C-427958D6C1257F8700415473/\\$file/2016_LAWS+MX_CountryPaper+Switzerland.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/D2D66A9C-427958D6C1257F8700415473/$file/2016_LAWS+MX_CountryPaper+Switzerland.pdf).
- 40 In this vein, e.g. *The Weaponization of Increasingly Autonomous Technologies: Considering Ethics and Social Values*, UNIDIR, 2015, <https://unidir.org/files/publications/pdfs/considering-ethics-and-social-values-en-624.pdf>; E. Rosert and F. Sauer, 'Prohibiting Autonomous Weapons: Put Human Dignity First', in 10(3) *Global Policy* (2019), <https://onlinelibrary.wiley.com/doi/full/10.1111/1758-5899.12691>.