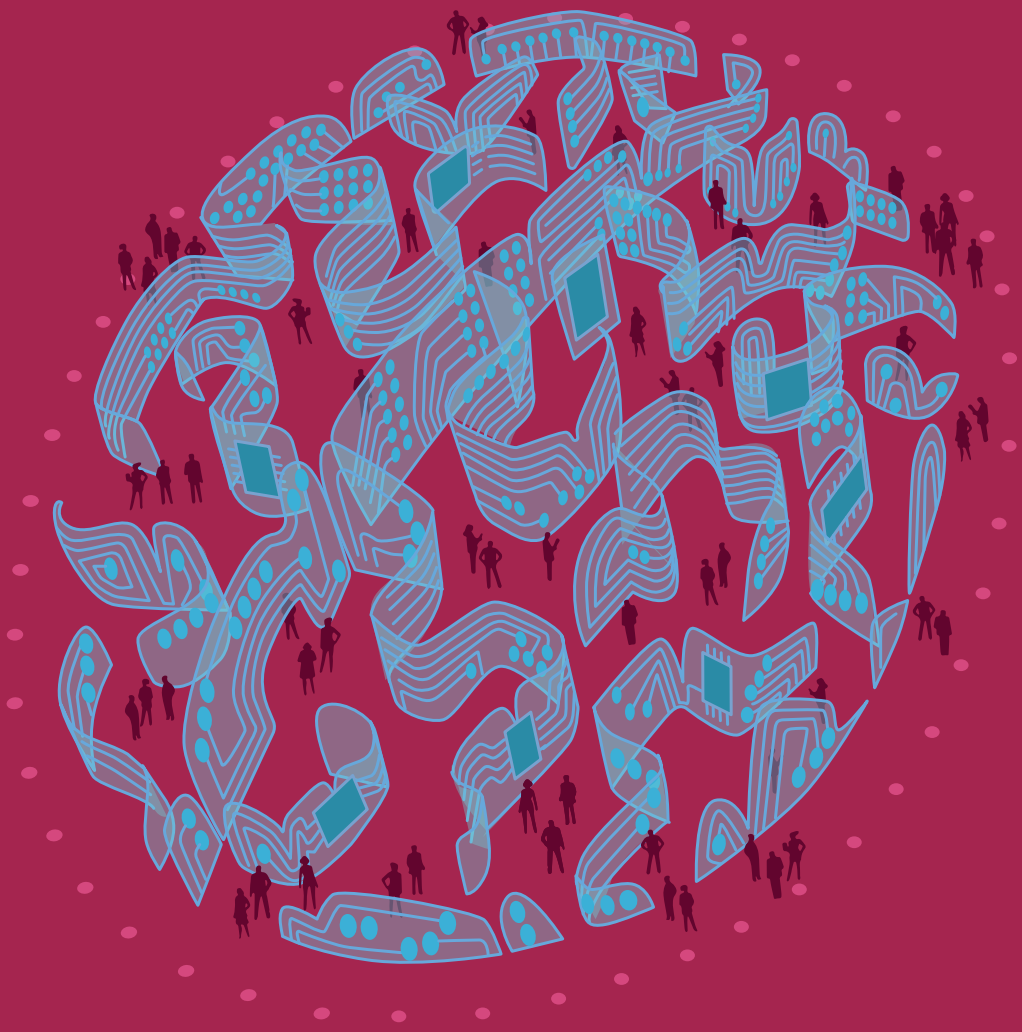
GLOBAL INFORMATION SOCIETY WATCH 2019

Artificial intelligence:
Human rights, social justice and development



Association for Progressive Communications (APC),
Article 19, and Swedish International Development Cooperation Agency (Sida)

Global Information Society Watch

2019

Global Information Society Watch 2019
Artificial intelligence: Human rights, social justice and development

Operational team
Valeria Betancourt (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Maja Romano (APC)

Project coordination team
Valeria Betancourt (APC)
Cathy Chen (APC)
Flavia Fascendini (APC)
Alan Finlay (APC)
Mallory Knodel (ARTICLE 19)
Vidushi Marda (ARTICLE 19)
Leila Nachawati (APC)
Lori Nordstrom (APC)
Maja Romano (APC)

GISWatch 2019 advisory committee
Namita Aavriti (APC)
Rasha Abdul Rahim (Amnesty International)
Alex Comninos (Research ICT Africa)
Malavika Jayaram (Digital Asia Hub)
J. Carlos Lara (Derechos Digitales - América Latina)
Joy Liddicoat (Centre for Law and Emerging Technologies, University of Otago)
Andrew Lowenthal (EngageMedia)
Micaela Mantegna (Geekylegal/Machine Intelligence Lab, Center for Technology and Society, San Andres University)
Valeria Milanes (Asociación por los Derechos Civiles)

Project coordinator
Maja Romano (APC)

Editor
Alan Finlay (APC)

Assistant editor and proofreading
Lori Nordstrom (APC)

Publication production support
Cathy Chen (APC)

Graphic design
Monocromo

Cover illustration
Matías Bervejillo

We would like to extend a special note of thanks to a number of authors who have made ad honorem contributions to this edition of GISWatch.
We gratefully acknowledge the following:

Philip Dawson and Grace Abuhamad (Element AI)
Anita Gurumurthy and Nandini Chami (IT for Change)
Rasha Abdul Rahim (Amnesty International)

APC would like to thank the Swedish International Development Cooperation Agency (Sida) and ARTICLE 19
for their support for Global Information Society Watch 2019.

Published by APC
2019
Printed in USA

Global Information Society Watch 2019 web and e-book
ISBN 978-92-95113-13-8
APC Serial: APC-201910-CIPP-R-EN-DIGITAL-302

The weaponisation of AI: An existential threat to human rights and dignity

Rasha Abdul Rahim
Amnesty International
www.amnesty.org

Introduction

Over the past decade, there have been extensive advances in artificial intelligence (AI) and other technologies. AI is being incorporated in nearly all aspects of our lives, in sectors as diverse as health care, finance, travel and employment. Another sphere where AI innovation is occurring at a rapid pace is in the military and law enforcement spheres, making possible the development and deployment of fully autonomous weapons systems which, once activated, can select, attack, kill and wound human targets without meaningful human control. These weapons systems are often referred to as Lethal Autonomous Weapons Systems (LAWS) and, more comprehensively, "Autonomous Weapons Systems" (AWS), which encompass both lethal and less-lethal systems.

The rapid development of these weapons systems could not only change the entire nature of warfare, it could also dramatically alter the conduct of law enforcement operations and pose extremely serious human rights risks.[1]

With continuous advances in technology and states such as China, France, Israel, Russia, South Korea, the United States (US) and United Kingdom (UK) heavily investing in and developing weapons with increasing autonomy in the critical functions of selecting and using force on targets, other states are considering how to respond to the automation of warfare and policing. What is clear is that the development and use of AWS raises serious legal, ethical, technological, accountability and security concerns, which is why the Campaign to Stop Killer Robots,[2] of which Amnesty International is a member, is calling for a prohibition on AWS in order to ensure meaningful human control over weapons systems.

Under the auspices of the Convention on Certain Conventional Weapons (CCW), the Campaign has since 2014 been advocating for states to urgently begin negotiations on a legally binding instrument to ensure that meaningful human control is retained over the use of force by prohibiting the development, production, transfer and use of AWS. But while AI weapons technologies race ahead, legal and policy responses to this issue lag woefully behind.

Human rights risks of AWS

AWS can be characterised as weapons capable of selecting and applying force against targets without meaningful control. Autonomy in weapons systems should be understood as a continuum; these systems are not to be confused with unmanned aerial vehicles (UAVs), commonly referred to as drones, which are remotely piloted by a human operator. By contrast, AWS would incorporate software and algorithms which, on their own, would be able to make critical determinations about life and death. Such systems raise important legal, ethical, technological, accountability and security challenges if developed to operate without meaningful control by humans.

The concept of "meaningful human control" was coined by the NGO Article 36,[3] with the aim of setting a normative limit on autonomy in weapons systems by determining the human element required over the use of force.[4] It denotes a level of control which is not purely superficial, for example, a human pressing a button to deploy force against a target that a machine has independently identified.

In situations of armed conflict, the rules of international humanitarian law (IHL) apply alongside human rights law. These require parties to a conflict to distinguish between civilians, who are afforded protection, and combatants, who may be directly attacked. Civilians who are not directly participating in hostilities must never be deliberately targeted. Parties also must distinguish between military

1 Amnesty International. (2015). *Autonomous Weapons Systems: Five Key Human Rights Issues for Consideration*. https://www.amnesty.org/en/documents/act30/1401/2015/en

2 https://www.stopkillerrobots.org

3 www.article36.org

4 Amnesty International. (2018, 27 August). UN: Decisive action needed to ban killer robots – before it's too late. https://www.amnesty.org/en/latest/news/2018/08/un-decisive-action-needed-to-ban-killer-robots-before-its-too-late

objectives and civilian objects (such as residential buildings, schools and hospitals), and direct attacks only at military objectives. All parties to the conflict must take measures to minimise harm to civilians and civilian objects and must not carry out attacks that fail to distinguish between civilians and combatants, or which cause disproportionate harm to civilians and civilian objects.

Due to the complexity and context-dependent nature of making such assessments in dynamic and cluttered environments, AWS would not be able to comply with IHL, including the requirement to distinguish adequately between combatants and civilians and to evaluate the proportionality of an attack. As former UN Special Rapporteur on extrajudicial, summary or arbitrary executions Christof Heyns argued in his 2013 report to the Human Rights Council, such assessments require intrinsically human qualities and human judgment. They also require:

> [...] common sense, appreciation of the larger picture, understanding of the intentions behind people's actions, and understanding of values and anticipation of the direction in which events are unfolding. Decisions over life and death in armed conflict may require compassion and intuition. Humans – while they are fallible – at least might possess these qualities, whereas robots definitely do not.[5]

Similarly, in law enforcement operations, which are governed by international human rights law (IHRL) alone and elaborated through international policing standards such as the UN Basic Principles on the Use of Force and Firearms (UNBPUFF),[6] the use of lethal and less-lethal AWS without meaningful human control would result in unlawful killings and injuries.

AWS threaten various fundamental human rights, most notably, the right to life which is enshrined in Article 6(1)[7] of the International Covenant on Civil and Political Rights (ICCPR).[8] Under IHRL the use of potentially lethal force is only lawful if it meets the following cumulative requirements: it must have sufficient legal basis in line with international standards; be necessary to protect human life; constitute a last resort; be applied in a manner proportionate to the threat; and law enforcement officers must be held accountable for their use of force.

Under Principle 9 of the UNBPUFF, law enforcement officers may only use lethal force if there is an imminent threat to life or serious injury. This involves a complex assessment of potential or imminent threats, for example, who is posing the threat, identifying and using means other than force, considering whether force is needed to neutralise the threat, deploying different modes of communication to neutralise the threat, deciding on the use of weapons/equipment, etc., and of how best to protect the right to life. These are inherently human skills which cannot be automated, especially given the ever-evolving, dynamic and unpredictable nature of law enforcement operations.

When applying less lethal force, law enforcement officers must apply non-violent means before resorting to use of force, for example, by using techniques including persuasion, negotiation and de-escalation. These techniques require human empathy, negotiation skills, understanding crowd behaviour, and a high level of training and ability to respond to dynamic and unpredictable situations – skills unlikely to be replicated by algorithms.

AWS could also be used to facilitate violations of the right to freedom of peaceful assembly.[9] Indeed, as Heyns has stated:

> [O]n the domestic front, LARs [Lethal Autonomous Robotics] could be used by States to suppress domestic enemies and to terrorize the population at large, suppress demonstrations and fight "wars" against drugs. It has been said that robots do not question their commanders or stage coups d'état.[10]

5 Heyns, C. (2013). *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns*. www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf; see also General comment No. 36 (2018) on article 6 of the International Covenant on Civil and Political Rights, 30 October 2018, para 65. https://tbinternet.ohchr.org/Treaties/CCPR/Shared%20Documents/1_Global/CCPR_C_GC_36_8785_E.pdf: "For example, the development of autonomous weapon systems lacking in human compassion and judgement raises difficult legal and ethical questions concerning the right to life, including questions relating to legal responsibility for their use. The Committee is therefore of the view that such weapon systems should not be developed and put into operation, either in times of war or in times of peace, unless it has been established that their use conforms with article 6 and other relevant norms of international law."

6 https://www.ohchr.org/en/professionalinterest/pages/useofforceandfirearms.aspx

7 Article 6(1), ICCPR: "Every human being has the inherent right to life. This right shall be protected by law. No one shall be arbitrarily deprived of his life."

8 https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx

9 Article 21, ICCPR: "The right of peaceful assembly shall be recognized. No restrictions may be placed on the exercise of this right other than those imposed in conformity with the law and which are necessary in a democratic society in the interests of national security or public safety, public order, the protection of public health or morals or the protection of the rights and freedoms of others."

10 Heyns, C. (2013). Op. cit. The implications of this can be seen in protests in Gaza in 2018, during which Israel deployed semi-autonomous drones to fire tear gas indiscriminately at protesters – it is likely that more autonomous systems will be deployed by law enforcement agencies in the future.

AWS would also undermine the right to privacy (ICCPR article 17) and the right to equality and non-discrimination (ICCPR article 26). Masses of data will need to be collected to train targeting algorithms to profile personal data and create patterns on the basis of which AWS would make decisions on when to use force and against whom. AWS could therefore fuel the bulk collection of data and result in indiscriminate mass surveillance, which is never a proportionate interference with the right to privacy.

The mass collection and profiling of personal data could also have an impact on the right to equality and non-discrimination. Systems employing machine-learning technologies can vastly and rapidly reinforce or change power structures, as the data sets used to teach algorithms contain historical biases which are then reproduced and amplified.[11] For example, in a study by the American Civil Liberties Union, the facial recognition tool called "Rekognition" incorrectly matched 28 members of the US Congress, identifying them as other people who have been arrested for a crime.[12] The false matches were disproportionately of people of colour, including six members of the Congressional Black Caucus. AWS would therefore have the potential to entrench systemic discrimination, with potentially lethal consequences.

Delegating life-and-death decisions to machines

Quite apart from serious concerns as to whether autonomous technologies would be technically capable of conforming to international law, AWS raise numerous important ethical and social concerns – especially since AWS would not be able to refuse orders – about the delegation of human decision-making responsibilities to an autonomous system designed to injure and kill. As Heyns asserts, "[T]here is widespread concern that allowing [autonomous weapons] to kill people may denigrate the value of life itself."[13] Thus the right not just to life, but to a life with dignity, is undermined.[14] Lowering the threshold for the use of force would further depersonalise the use of force, which has already begun through the use of armed drones.

There is also a wider question about the future of our humanity. Is it acceptable to delegate human decision-making responsibilities to use force to a machine? Proponents of AWS argue that removing humans from the equation would increase speed, efficiency, accuracy, stealth and would also cut out emotions – panic, fear, revenge – which can lead to mistakes and unlawful actions. But this is a false dichotomy, as human biases are reflected in algorithms, and therefore neither humans nor machines are infallible. Indeed, human emotions such as empathy can lead to acts of mercy.

Risks to international security

AWS are also vulnerable, as without human oversight they are prone to design failures, errors, hacking, spoofing and manipulation, making them unpredictable. As the complexity of these systems increases, it becomes even more difficult to predict their responses to all possible scenarios, as the number of potential interactions within the system and with its complex external world is simply too large.[15] This would be compounded by autonomous machines interacting with other autonomous machines, posing a risk not only to civilians, but also soldiers and police officers.

The development of AWS would inevitably spark a new high-tech arms race between world superpowers, with each state wanting to keep up with new technologies and seeking to secure them for their arsenals. Given the intangible nature of the software, AWS may also proliferate widely to unscrupulous actors, including non-state actors. In addition, the ease of deploying these weapons may result in an unintended escalation in conflicts.

Therefore, human control and the autonomy of systems should not be viewed as mutually exclusive. The strengths of humans (legal and moral agents, fail-safe) and strengths of machines (data processing, speed, endurance, etc.) should be combined to ensure compliance with the law and predictability, reliability and security.

11 Lum, K., & Isaac, W. (2016). To predict and serve? *Significance, 13*(5), 14-19. https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x

12 Eleven of the 28 false matches misidentified people of colour (roughly 39%), including civil rights leader Rep. John Lewis (D-GA) and five other members of the Congressional Black Caucus. Only 20% of current members of Congress are people of colour, which indicates that false-match rates affected members of colour at a significantly higher rate. Snow, J. (2018, 26 July). Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots. *American Civil Liberties Union*. https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

13 Heyns, C. (2013). Op. cit.

14 Article 10, ICCPR: "All persons deprived of their liberty shall be treated with humanity and with respect for the inherent dignity of the human person."

15 Scharre, P. (2016). *Autonomous Weapons and Operational Risk*. Center for a New American Security. https://www.cnas.org/publications/reports/autonomous-weapons-and-operational-risk

AWS would evade accountability mechanisms

States have legal obligations to prevent and redress human rights violations by their agents, as well as a duty to prevent, investigate, punish and redress the harm caused by human rights abuses by private persons or entities. A failure to investigate an alleged violation of the right to life could in and of itself constitute a breach of this right.[16] In armed conflict, states also have obligations under international humanitarian law to investigate, and where appropriate prosecute, potential war crimes.[17]

Since it is of course not possible to bring machines to justice, who would be responsible for serious violations? Would it be the programmers, commanders, superior officers, political leaders or manufacturers? It would be impossible for any of these actors to reasonably foresee how an AWS will react in any given circumstance, given the countless situations it may face. Furthermore, without meaningful human control, commanders and superior officers would not be in a position to prevent an AWS from carrying out unlawful acts.

This accountability gap would mean victims and families of victims would not be able to access effective remedy. This would mean states' obligation to ensure that victims and families of victims of violations of IHL or IHRL receive full reparation could not be met.

Conclusion

Given the unacceptably high risk that AWS pose to human rights, as well as the ethical, moral and security threats their use would entail, Amnesty International is calling for a legally binding instrument to ensure that meaningful human control is retained over the use of force by prohibiting the development, production, transfer and use of AWS.

Momentum for a ban is steadily growing. Many states, including Austria, Brazil, Mexico, states forming the African Group, and the Non-Aligned Movement, have emphasised the importance of retaining human control over weapons and the use of force. Most states expressed support for developing new international law on AWS, and so far, 29 states[18] have called for them to be banned. These states are largely from the global South, perhaps indicating a well-founded fear that AWS are likely to be used against them.

UN Secretary-General António Guterres also voiced strong support for a ban, describing weapons that can select and attack a target as "morally repugnant".[19] In his Agenda for Disarmament[20] he pledged to support states to elaborate new measures, such as a legally binding instrument. On 12 September 2018 a large majority (82%) in the European Parliament[21] called for an international ban on AWS and for meaningful human control over the critical functions of weapons.

Despite this, a small group of states including Russia, the US, the UK, Australia, Israel, France and Germany are blocking movement towards negotiations for a ban. These are all countries known to be developing AWS.[22] France and Germany have proposed a non-binding political declaration[23] as "a first step" to gather support for the principle of human control over future lethal weapons systems and to ensure they are in full compliance with international law.

16 General comment No. 36 (2018) on article 6 of the International Covenant on Civil and Political Rights, 30 October 2018, para 27. https://tbinternet.ohchr.org/Treaties/CCPR/Shared%20Documents/1_Global/CCPR_C_GC_36_8785_E.pdf

17 International Committee of the Red Cross, Customary International Humanitarian Law, Rule 158.

18 Campaign to Stop Killer Robots. (2019, 21 August). Country Views on Killer Robots. https://www.stopkillerrobots.org/wp-content/uploads/2019/08/KRC_CountryViews21Aug2019.pdf

19 Guterres, A. (2018, 25 September). Address to the General Assembly. https://www.un.org/sg/en/content/sg/speeches/2018-09-25/address-73rd-general-assembly

20 https://www.un.org/disarmament/sg-agenda/en /

21 https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2018-0341+0+DOC+XML+V0//EN&language=EN

22 For example, a recent report revealed that the UK Ministry of Defence and defence contractors are funding dozens of AI programmes for use in conflict, and in November 2018 the UK held exercise "Autonomous Warrior" (https://www.army-technology.com/news/british-autonomous-warrior-experiment), the biggest military robot exercise in British history, testing over 70 prototype unmanned aerial and autonomous ground vehicles. The UK has repeatedly stated that it has no intention of developing or using fully autonomous weapons (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/673940/doctrine_uk_uas_jdp_0_30_2.pdf). Yet such statements are disingenuous given the UK's overly narrow definition of these technologies ("machines with the ability to understand higher-level intent, being capable of deciding a course of action without depending on human oversight and control"), making it easier for the UK to state that it will not develop such weapons. Although Russia has said it believes the issue of AWS is "extremely premature and speculative", in 2017 Russian arms manufacturer Kalashnikov announced it would be launching a range of "autonomous combat drones" which would be able to identify targets and make decisions without any human involvement; see Gilbert, D. (2017, 13 July). Russian weapons maker Kalashnikov developing killer AI robots. *VICE*. https://news.vice.com/en_us/article/vbzq8y/russian-weapons-maker-kalashnikov-developing-killer-ai-robots

23 https://www.unog.ch/80256EDD006B8954/(httpAssets)/895931D082ECE219C12582720056F12F/$file/2018_LAWSGeneralExchange_Germany-France.pdf

Encouragingly, momentum has been growing in the private sector. The workforce of tech giants like Amazon, Google and Microsoft have all challenged their employers and voiced ethical concerns about the development of artificial intelligence technologies that can be used for military and policing purposes.[24]

In addition, nearly 250 tech companies, including XPRIZE Foundation, Google DeepMind and Clearpath Robotics, and over 3,200 AI and robotics researchers, engineers and academics have signed a Lethal Autonomous Weapons Pledge[25] committing to neither participate in nor support the development, manufacture, trade or use of autonomous weapons.

This demonstrates widespread support for a legally binding treaty, despite proposals for weaker policy responses. Just as ethical principles have not been effective in holding tech companies to account, non-legally binding principles would fall far short of the robust response needed to effectively address the multiple risks posed by these weapons.

24 For example, in April 2018 around 3,100 Google staff signed an open letter protesting Google's involvement with Project Maven, a programme which uses machine learning to analyse drone surveillance footage in order to help the US military identify potential targets for drone strikes. Google responded by releasing new AI principles (https://www.blog.google/technology/ai/ai-principles), which included a commitment not to develop AI for use in weapons, and announced it would not renew the Project Maven contract when it expired in 2019. See Wakabayashi, D, & Shane, S. (2018, 1 June). Google Will Not Renew Pentagon Contract That Upset Employees. *The New York Times*. https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html

25 https://futureoflife.org/lethal-autonomous-weapons-pledge

GLOBAL INFORMATION SOCIETY WATCH 2019

Artificial intelligence:
Human rights, social justice and development

Artificial intelligence (AI) is now receiving unprecedented global attention as it finds widespread practical application in multiple spheres of activity. But what are the human rights, social justice and development implications of AI when used in areas such as health, education and social services, or in building "smart cities"? How does algorithmic decision making impact on marginalised people and the poor?

This edition of Global Information Society Watch (GISWatch) provides a perspective from the global South on the application of AI to our everyday lives. It includes 40 country reports from countries as diverse as Benin, Argentina, India, Russia and Ukraine, as well as three regional reports. These are framed by eight thematic reports dealing with topics such as data governance, food sovereignty, AI in the workplace, and so-called "killer robots".

While pointing to the positive use of AI to enable rights in ways that were not easily possible before, this edition of GISWatch highlights the real threats that we need to pay attention to if we are going to build an AI-embedded future that enables human dignity.

GLOBAL INFORMATION SOCIETY WATCH
2019 Report
www.GISWatch.org

APC
ARTICLE 19
Sida