Data and Privacy in Autonomous Weapons Systems

March 2026

Data and Privacy in Autonomous Weapons Systems

CONTENTS

INTRODUCTION

As states, together with other stakeholders, deepen their engagement with the development and deployment of autonomous weapons systems (AWS), it is imperative that efforts to regulate these systems recognise privacy and data protection not as peripheral technical considerations, but as core human rights obligations integral to any meaningful regulatory framework. With this briefing, Privacy International seeks to contribute to the current negotiations—including the on-going work of the Group of Governmental Experts of the High Contracting Parties related to emerging technologies in the area of lethal autonomous weapons systems (LAWS)[1]—by highlighting the urgent need to embed privacy and data protection safeguards in all stages of AWS design, deployment, and oversight.

The conversations around the prohibition and regulation of AWS[2] have dominantly focused on legal and ethical considerations of digital dehumanisation, loss of dignity, compliance with international humanitarian law, accountability and meaningful human control over such systems.[3] These well-grounded concerns over AWS often focus too narrowly on their possible deployment in a context of armed conflict. As debates and practices evolve, PI aims to ensure that future governance of AWS reflects both existing international human rights obligations and the practical risks posed by data-driven military technologies in conflict and non-conflict situations.

AWS are not a distinct category of technology or weapon. It is the integration of the technology, through the upgrade or addition of software, that renders a weapon autonomous, adding the functions of identification, selection, and engagement through the processing of vast amounts of

[1] Convention on Certain Conventional Weapons (CCW) - Group of Governmental Experts of the High Contracting Parties related to emerging technologies in the area of lethal autonomous weapons systems (LAWS), 2026, https://meetings.unoda.org/ccw/convention-on-certain-conventional-weapons-group-of-governmental-experts-on-lethal-autonomous-weapons-systems-2025

[2] Concerning the terminology of *lethal* autonomous weapons, numerous states, civil society, including the Stop Killer Robots campaign, and other stakeholders, including the International Committee of the Red Cross (ICRC) have consistently raised the view that the use of the word 'lethal' in discussions on autonomous weapons is inappropriate and limits the scrutiny on these weapons.

[3] Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons System, Geneva, 3-7 March and 1-5 September 2025, Chair's Summary – First 2025 session of the GGE on LAWS, CCW/GGE.1/2025/WP, GE.25-05542 (E), https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2025)/CCW-GGE.1-2025-WP.1_-_Chair%27s_summary.pdf; UN Human Rights Council, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions', A/HRC/23/47, 9 April 2013, https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf, p 20; Statement of the UN Secretary General, SG/SM/19512-DC/3797, 25 March 2019, https://press.un.org/en/2019/sgsm19512.doc.htm; ICRC, Position on Autonomous Weapon Systems, May 2021, https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems; Stop Killer Robots Campaign, https://www.stopkillerrobots.org/stop-killer-robots/facts-about-autonomous-weapons/

personal data.[4] In other words, data-intensive technologies are at the core of autonomous weapons and sensor-based targeting systems.[5]

Moreover, the insistence on framing the problems exclusively in an armed conflict setting leaves out fundamental concerns, such as 'their potential use in border control, policing, and surveillance contexts, and the profound ethical and moral implications of using machines to sense, process, and target people as patterns of data and objects, whether with 'lethal' force or otherwise.'[6] History suggests that technologies initially developed and deployed by the military in armed conflict tend to migrate rapidly into civilian and law enforcement contexts.[7] This is a dynamic that is already observable in relation to surveillance and enforcement systems, such as drones used for policing and border controls.[8] It is therefore imperative to broaden the debate and address the issues posed by the use of AWS in contexts other than armed conflict.

These systems come with consequential yet neglected concerns. The technology involved in AWS possesses inherent shortcomings related to privacy and personal data, machine learning, algorithmic bias and gaps, human supervision and meaningful control, accountability and responsibility. Consequences of these shortcomings can be lethal.

[4] ICRC, 'Autonomy, artificial intelligence and robotics: Technical aspects of human control', Geneva, August 2019, https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control, p 6.

[5] Sensor-based targeting systems can be defined as "systems designed to support the targeting process by detecting and proposing potential targets to human operators, where such systems operate by matching sensor inputs from the environment against encoded profiles of intended target-types, without human assessment of those sensors inputs." See Article 36, Policy Note, Sensor-Based Targeting Systems: An Option for Regulation, November 2021, https://article36.org/updates/sensor-based-targeting-systems/

[6] See Automated Decision Research, Targeting people and digital dehumanisation, 2023, https://automatedresearch.org/news/report/targeting-people-and-digital-dehumanisation/

[7] Privacy International (PI), 'Challenging Militarisation of Tech', 2024, https://privacyinternational.org/campaigns/militarisation-of-tech

[8] PI, 'What is Militarisation of Tech?', September 2025, https://privacyinternational.org/long-read/5668/what-militarisation-tech

1. AWS THREATS TO THE RIGHT TO PRIVACY

1.1. Personal data and privacy at the core of AWS

The role of personal data in AWS is an important yet overlooked consideration in the regulatory discussions. AWS depend on the continuous collection, processing, and inferential analysis of vast quantities of data, much of it drawn from sensors, surveillance infrastructures, and data-rich environments. Data, and specifically personal data, is at the core of functions that facilitate the autonomy in weapon systems.[9] These functions, namely identification, selection and engagement are heavily reliant on mass quantities of personal data to be trained, tested, improved, and deployed. These weapons are designed to rely on and process vast amounts of observed, derived, or inferred data for their functioning.

The ability to feed more personal data into the systems is often argued to allow, in return, improved accuracy in the execution of tasks and better functionality with an increased number of scenarios that can be encountered in deployment. The assumption that the more the data the better the accuracy is however flawed. Notably it does not reflect on the impact of processing vast amounts of poor-quality data, nor the discriminatory outcomes that may result from reliance on biased datasets (more on that below). In practice, large datasets often reproduce noise, bias, and contextual errors, increasing the likelihood of misidentification and unlawful targeting, while exposing individuals and communities to disproportionate surveillance. Recent examples suggest that feeding more data to a target-recommendation tool can lead to an exponential increase in potential targets, rather than precision and accuracy.[10]

1.2. Lack of appropriate regulatory frameworks

The above underscores that AWS' data dependencies are not merely technical features but significant threats to the right to privacy, demanding scrutiny and explicit safeguarding. Despite these threats, such data is often processed without an adequate legal framework in place and without consideration for the legal basis for processing this data.[11] National data protection legislation when it exists often carves out from its scope of application data processing related to

[9] PI, 'How Data Drives the Militarisation of Tech', September 2025, https://privacyinternational.org/long-read/5667/how-data-drives-militarisation-tech

[10] Yuval Abraham, ''Lavender': The AI machine directing Israel's bombing spree in Gaza', *+972 Magazine*, 3 April 2024, https://www.972mag.com/lavender-ai-israeli-army-gaza/

[11] UN Human Rights Council, Report of the UN High Commissioner for Human Rights on The Right to Privacy in the Digital Age, A/HRC/48/31, 13 September 2021, https://documents.un.org/symbol-explorer?s=A/HRC/48/31&i=A/HRC/48/31_6383994, paras 23-24. See also, Human Rights Watch (HRW), 'Report Hazard to Human Rights', 28 April 2025, https://www.hrw.org/report/2025/04/28/a-hazard-to-human-rights/autonomous-weapons-systems-and-digital-decision-making, p 49.

national security, defence and/or military.[12] A similar blanket exclusion is contained in recent regional AI legislative initiatives.[13] As a result, the processing of personal data for AWS takes place in a regulatory vacuum, despite the serious interference with the right to privacy and other human rights that it entails.

1.3. Normalising mass surveillance

The indiscriminate processing of personal data for AWS is putting entire populations within and outside conflict under constant surveillance.[14] The context of many current armed conflicts, including military occupation, is characterised by methods of mass surveillance, including automated checkpoints, constant facial recognition and video surveillance, and biometric ID systems in exchange of humanitarian aid.[15] The development of data-intensive systems, such as AWS, requires the uncontrolled fusion of civilian and military data,[16] as none of these systems can function without the massive surveillance infrastructures that support their functionality and their constant demand for personal data at population scale, be it satellite data, telecommunications data or other.[17] All this data is now allowed to feed into AWS. The reliance of AWS on personal data fits within a trend among state security and military agencies since 9/11 towards the use of predictive surveillance technologies to identify potential targets/suspects.

The data fed into those systems are no longer confined to traditional understandings of what constitutes military intelligence. All personal data is now considered relevant for military purposes. Users' data from social media platforms are now feeding into war machines. Characteristically, Big-Tech companies as well as start-ups that initially developed commercial and civilian infrastructure technologies are entering the defence sector. Many reneged on previous promises not to build AI for surveillance or war-related purposes.[18] The role of civilian data in conflict zones has been evident

[12] Article 2(2)(b), Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (GDPR), *OJ L* 119, 4 April 2016.

[13] Article 2(3), Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (EU AI Act), OJ L, 2024/1689, 12.7.2024; Article 3(2), Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series - No. 225, Vilnius, 5.IX.2024.

[14] HRW, 'Report Hazard to Human Rights', note above, p 49.

[15] PI, 'Biometrics and counter-terrorism: Case study of Israel/Palestine', 28 Mary 2021, https://privacyinternational.org/report/4527/biometrics-and-counter-terrorism-case-study-israelpalestine

[16] Henning Lahmann, 'Gaza and the Collective Political Costs of Algorithmic Warfare', 25 March 2025, https://www.ejiltalk.org/gaza-and-the-collective-political-costs-of-algorithmic-warfare/?utm_source=mailpoet&utm_medium=email&utm_campaign=ejil-talk-newsletter-post-title_2

[17] PI, 'Mass surveillance', https://privacyinternational.org/learn/mass-surveillance

[18] For example, in 2025, Google updated its ethical guidelines by removing previous promises to not pursue 'weapons, surveillance, and technologies that can cause or are likely to cause overall harm'. Nitasha Tiku & Gerrit De Vynck, 'Google drops pledge not to use AI for weapons or surveillance', 4 February 2025, https://www.washingtonpost.com/technology/2025/02/04/google-ai-policies-weapons-harm/ Meta also announced in November 2024 that it would permit use of their AI for national security.

in both Ukraine[19] and Gaza.[20] In simple terms, anyone's personal data may end up building a lethal algorithm and we would all be oblivious to this.

1.4. Privacy violations enable the violation of other rights

To be clear: it is not only the right to privacy that is at stake here. All these surveillance methods and tools come already with extensive human rights concerns.[21] The failure to strictly regulate these AWS risks encouraging more mass surveillance than ever, further exacerbating the unchecked state power and control over individuals. Unchecked surveillance and the widespread, often unlawful exploitation of personal data do not just pose privacy risks—they can directly enable life-and-death decisions by weapons systems, with devastating and irreversible consequences that strike at the heart of the right to life and human dignity. Yet, it is not only the often-lethal consequences of the reliance on personal data by AWS to determine their actions (e.g. targeting individuals) that come into play. By relying on generalised harvesting of data the AWS support forms of mass surveillance that lead to serious violations of the right to privacy and other human rights, including freedom of peaceful assembly, freedom of expression, freedom of movement, principle of non-discrimination, as well as political participation.[22]

1.5. Private sector dependencies

The role of private companies in collecting and processing the personal data that feed these AWS is staggeringly complex and opaque. Household names such as Microsoft, Amazon and Google are increasingly offering data processing services to militaries involved in armed conflicts.[23] Others are developing dual-use technologies, benefitting both from the military's increased appetite for data and the expansion of some military technologies into civilian spaces.[24] All of them thrive in a poorly

---

Nick Clegg, 'Open Source AI Can Help America Lead in AI and Strengthen Global Security', 4 November 2024, https://about.fb.com/news/2024/11/open-source-ai-america-global-security/ See further PI, 'What is Militarisation of Tech?', note above.

[19] In Ukraine, the use of Clearview AI demonstrates how social-media-scraped facial images have become a core component of wartime intelligence, enabling authorities to identify Russian soldiers and other individuals by linking battlefield photos to their social-media profiles. PI, 'The Clearview/Ukraine partnership - How surveillance companies exploit war', 8 March 2022, https://privacyinternational.org/news-analysis/4806/clearviewukraine-partnership-how-surveillance-companies-exploit-war

[20] Similarly, in Gaza, the Israeli military's digital targeting systems reportedly rely on extensive and systematic surveillance of Palestinians' personal data—including information drawn from their social-media activity—to inform threat predictions and attack-related decision-making. HRW, 'Questions and Answers: Israeli Military's Use of Digital Tools in Gaza', 10 September 2024, https://www.hrw.org/news/2024/09/10/questions-and-answers-israeli-militarys-use-digital-tools-gaza

[21] *ibid*. See also HRW, 'Report Hazard to Human Rights', note above.

[22] PI, 'Mass surveillance', note above.

[23] PI, 'Big Tech's bind with military and intelligence agencies', 1 October 2025, https://privacyinternational.org/long-read/5683/big-techs-bind-military-and-intelligence-agencies

[24] Corporate Watch & PI, 'Investigating dual-use technology and the darker side of innovation', 2025-26, https://privacyinternational.org/long-read/5705/investigating-dual-use-technology-and-darker-side-innovation

regulated space, lacking transparency and accountability. All of them have or will likely develop a stake in the development of AWS.

Personal data plays a fundamental role in the design, development and deployment of AWS. It is not a byproduct that can be easily bypassed and current privacy-preserving technologies have not yet proved effective to address the risks to privacy. As it currently stands, the personal data processed by AWS fails the test of legality, necessity, and proportionality required under the human right to privacy. Any debate on the banning/regulation of AWS must take this aspect into consideration. These concerns have further implications when it comes to machine learning, algorithmic bias and discrimination, human supervision and accountability gaps.

2. MACHINE LEARNING AND PRIVACY CONCERNS

Machine learning can be understood as 'the technology that allows systems to learn directly from examples, data, and experience.'[25] Conventional programming methods depend on explicitly written rules that outline step-by-step instructions to solve a problem. On the other hand, machine learning systems are assigned a task and provided with a large dataset, which they use to learn how to perform the task or to identify patterns within the data.[26]

Machine learning takes place through three different modalities, namely supervised machine learning, unsupervised machine learning, and reinforcement learning. Whereas labelled data is fed to the system in supervised learning, unsupervised learning aims to identify and categorise unlabelled data based on their similiarities. Lastly, reinforcement learning aims to learn from the simulated experience through mistakes and wins, leading the system to learn the consequences of its decisions.[27]

There are canonical problems starting at the very early stages of machine learning functions. Objects and situations are identified and described through computer vision algorithms embedded in autonomous weapons systems faster than humans can, allowing a quicker reaction time to a perceived threat. Beyond object recognition, certain AI systems are designed to infer emotional or intentional states from visual cues — including assessments of whether an individual poses a threat. Such systems are deeply unreliable even in controlled settings; their deployment in the chaotic conditions of an armed conflict is ever more concerning. However, an algorithm cannot understand the meaning and the context of an object,[28] and less so in armed conflict. Even for the most

[25] The Royal Society, 'Machine learning: the power and promise of computers that learn by example', April 2017, https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf, p 16.
[26] *ibid*, p 19.
[27] *ibid*, p 20.
[28] ICRC, 'Autonomy, artificial intelligence and robotics: Technical aspects of human control', Geneva, August 2019, https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control, p 20.

advanced militaries with these tools, it is a challenge to identify attempts to surrender or distinguish a walking cane from a gun,[29] or make a nuanced determination as to whether an assistive device of a person with disabilities is not a threat.[30] Understanding the human intent requires a sophisticated understanding of humanity, a task for which machines fall short.[31]

The problem is not limited to the development phase. One specific modality – reinforcement learning – continuously uses feedback mechanisms to help the system learn positive or negative consequences about its actions, constantly informing the model to process data in pursuit of its results.[32] This learning continues until the problem is solved (e.g. *machine wins the game)* or too many mistakes are made. Mistakes in this process may lead to disproportionate/unlawful death and injury to civilian populations and damage to civilian property in the context of armed conflict, or excessive use of force in law enforcement.

All these concerns are worsened by the fact that opaque machine learning systems produce their results without tangible explanation. Often referred to as "black boxes", these systems prevent human intervention in the learning process—intervention that could otherwise mitigate the harm by bringing common sense and ethical considerations to bear, values known to be exclusively human. Militaries and other entities that employ such opaque machine-learning systems are at risk of "over-reliance on AI-generated outputs" without the agency to effectively question the learning process.[33]

Throughout the process of reinforcement learning, the machine may develop unforeseen solutions to the task, including unlawful conduct under IHL to reach its goal. Beyond that, translating learning from simulation to the real-world, known as "sim-to-real", remains highly complex in armed conflict environments. Experts have drawn attention to even more problems that may occur. The algorithm may learn to exploit errors in its restraints and goals ("reward hacking"), carry out actions unrelated to its main goal, and even override human control and supervision to reach its goal.[34] Currently, it is proposed that the AWS should be regulated so that its mission parameters regarding the target selection and engagement functions cannot be modified by the system without context-appropriate human control and judgement.[35] While this may mitigate some of the concerns related to

[29] Mariana Díaz Figueroa et al, 'The Risks of Autonomous Weapons: An Analysis Centred on the Rights of Persons with Disabilities', 105(922) *International Review of the Red Cross*, April 2023, pp 278–305.
[30] Report of the Special Rapporteur on the rights of persons with disabilities, Artificial intelligence and the rights of persons with disabilities, 2021, A/HRC/49/52, https://www.ohchr.org/en/documents/thematic-reports/ahrc4952-artificial-intelligence-and-rights-persons-disabilities-report,
[31] The Royal Society, 'Machine learning…', note above, p 30.
[32] PI, 'Algorithms, Intelligence, and Learning Oh My', December 2016, https://privacyinternational.org/node/863; ICRC & Stockholm International Peace Research Institute (SIPRI), 'Limits on Autonomy in Weapon Systems, Identifying Practical Elements of Human Control', June 2020, https://www.sipri.org/sites/default/files/2020-06/2006_limits_of_autonomy_0.pdf, p 12.
[33] ICRC, 'Artificial intelligence and machine learning in armed conflict: A human-centred approach', 102 (913) *International Review of the Red Cross*, 2020, pp 463–479.
[34] ICRC, 'Autonomy, artificial intelligence and robotics…', note above, p 17.
[35] See the Chair's summary – First 2025 session of the GGE on LAWS, note above.

deployment of AWS, it does not address the overall concern related to the development phase and the current limits of machine learning embedded in AWS to reliably identify and distinguish targets.

3. ALGORITHMIC BIAS, DISCRIMINATION, AND GAPS

The performance of AWS as a data-intensive technology is heavily reliant on the quality and integrity of the large sets of training data. Consequently, the systems are very likely to exhibit training data bias.[36] This means that data that is over-representative of certain demographics, groups or areas, poor in quality or misinterpreted by humans to provide lawful compliance can lead to unlawful actions.

There are also significant risks of incorporating bias against racialised and historically discriminated populations, as well as gender-based bias.[37] The bias originates from the training datasets or algorithmic assumptions into the functioning of AWS,[38] and patriarchal and racist categorisations of humans as data will explicitly and inadvertently embed themselves.[39]

Beyond the initial datasets, *emergent bias* occurs throughout the deployment of the AWS due to the feedback from the environment that reflects the capacities, character, and habits of its users.[40] *Transfer context bias* is another risk that leads to unpredictability for AWS especially when the algorithm was deployed in an environment where it was not initially designed to function. For instance, artificial intelligence in facial recognition technologies employed in law enforcement or even military occupation is intrinsically distinct from armed conflict settings, which are characterised by active hostilities, fluid and uncontrolled environments, limited oversight, and rapidly changing targets.

Beyond the concerns of bias, the image classification algorithms of AWS can be manipulated by adversaries, and the AWS can be misled into wrongfully labelling a person or an object or failing to identify them at all. Such adversarial attacks (tricking) can cause unpredictable targeting and further damage the reliability of AWS. Experts consider adversarial attacks (tricking) is likely to become "a

[36] ICRC, 'Autonomy, artificial intelligence and robotics…', note above, p 20.
[37] It is observed that certain groups are already disproportionately affected by automated decision-making processes. PI, 'Artificial Intelligence', https://privacyinternational.org/learn/artificial-intelligence
[38] Article 36, 'Submission to the UN Secretary-General on the considerations on the development of an international legal instrument on autonomous weapons', 8 May 2024, https://article36.org/wp-content/uploads/2024/05/78-241-Article36-EN.pdf
[39] Ray Acheson, 'Autonomous Weapons and Patriarchy, Women's International League for Peace & Freedom (WILPF) and Stop Killer Robots, 2020, https://automatedresearch.org/news/report/autonomous-weapons-and-patriarchy/, p 12.
[40] Batya Friedman and Helen Nissenbaum, 'Bias in computer systems', *ACM Transactions on Information Systems (TOIS)*, Volume 14, Issue 3, July 1996, https://dl.acm.org/doi/10.1145/230538.230561, pp 330–347.

particularly acute problem in the inherently adversarial environments of conflict", when such algorithms are used in AWS.[41]

4. LIMITS TO THE HUMAN SUPERVISION AND MEANINGFUL HUMAN CONTROL

Context-appropriate human control and judgement regarding the use and effects of AWS has been considered essential to ensure its employment follows international law, and in particular IHL, including the principles and requirements of distinction, proportionality and precautions in attack.[42] While delegations at the Group of Governmental Experts of the High Contracting Parties related to emerging technologies in the area of lethal autonomous weapons systems try to clarify what this means in practice, PI notes that human supervision cannot act as the one solution to address the many shortcomings of AWS.

All the lethal shortcomings of AWS cannot simply be safeguarded by a human operator. First, in critical situations of life and death, the human operator will have less agency to understand, oppose or confirm in due time the actions of exceptionally complex technologies. Second, the perceived reliability of and the trust put in the system by the operator may create over-trust in the automation, decreasing the diligence in oversight required to intervene.[43]

Conversations around human control tend to focus on the one human who will 'hit the button'. Yet control can never lie with a single human operator. The deployment of AWS should be evaluated and understood through the entire lifecycle of its design and deployment, as well as the entire system within which it operates–including decision support systems it relies on, military operational command, the role of the private sector involved–for control to be meaningful.

5. LACK OF TRANSPARENCY AND ACCOUNTABILITY

Transparency and accountability go hand in hand and are indispensable for compliance with international humanitarian law and international human rights law. AWS cannot ensure accountability if they fail to be transparent, explainable and understandable.[44] The responsibility

[41] ICRC, 'Autonomy, artificial intelligence and robotics…', note above, p 21.
[42] Chair's summary – First 2025 session of the GGE on LAWS, note above.
[43] ICRC & SIPRI, 'Limits on Autonomy in Weapon Systems…', note above, p 19.
[44] European Commission Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for a Trustworthy AI, 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai; Automated Decision Research, 'Artificial intelligence and automated decisions: shared challenges in the civil and military spheres', September 2022, https://automatedresearch.org/news/report/artificial-intelligence-and-automated-decisions-shared-challenges-in-the-civil-and-military-spheres/, p 10.

and accountability cannot be transferred to weapon systems. The legal requirements on conduct of hostilities must be fulfilled by those who plan, decide on and carry out attacks, not the machines or programs.[45]

Yet, software-based machine learning, especially with the advances in AI, is considerably less transparent in its functioning compared to traditional models.[46] The opacity of the software in AWS poses significant problems for regulators, and even for the designers of the system themselves. If AWS are allowed to function without explainability in terms of profiling and targeting, it will make it extremely challenging to understand the outcomes (including on people's lives and civilian infrastructure). [47]

Further, legal accountability is only possible through the attribution of unlawful actions. Yet, attribution is technically complex in AWS due to the inherent difficulties in the human-machine interaction and the level of opacity of AWS already described. It has been rightfully noted that without the technical or investigative tools states may fail to fulfil their duty to investigate violations of international humanitarian and human rights law and hold violators accountable.[48]

[45] ICRC & SIPRI, 'Limits on Autonomy in Weapon Systems…', note above, p 4.
[46] Vincent Boulanin, 'Artificial intelligence: a primer', in Vincent Boulanin (Ed), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, *vol. 1, Euro-Atlantic Perspectives*, SIPRI, Stockholm, May 2019, https://www.sipri.org/sites/default/files/2019-05/sipri1905-ai-strategic-stability-nuclear-risk.pdf
[47] PI, 'Artificial Intelligence', note above.
[48] Leif Monnett, Research Brief, Sending Up a Flare: Autonomous Weapons Systems Proliferation Risks to Human Rights and International Security, Geneva Academy, July 2024, https://geneva-academy.ch/publication/sending-up-a-flare-autonomous-weapons-systems-proliferation-risks-to-human-rights-and-international-security/, p 8.

RECOMMENDATIONS

As mentioned earlier in this briefing, it is imperative that deliberations recognise privacy and data protection not as peripheral technical considerations, but as core human rights obligations integral to any meaningful regulatory framework. AWS rely on vast ecosystems of personal data—much of it sensitive, behavioural, or biometrically derived—and operate in ways that risk expanding surveillance capabilities, entrenching opacity, and undermining individual autonomy. Without explicit and enforceable safeguards, these systems threaten to normalise intrusive data practices in military and security contexts, erode accountability, and disproportionately impact already marginalised communities. Embedding strong privacy and data protection duties within the governance of AWS is therefore essential to preventing harm, preserving human dignity, and ensuring that emerging military technologies remain subject to the rule of law.

Privacy International is recommending states:

1. Initiate a global process toward a binding international instrument that bans and strictly regulates autonomous weapons systems (AWS). This instrument should be grounded in a systemic understanding of AWS and how they amplify structural harms—including threats to the right to privacy. It must address the entire lifecycle of these technologies, from research and development to deployment, export, and oversight.
2. Ensure that ongoing international efforts to ban or limit the use of autonomous weapons systems and broader Military AI applications explicitly integrate core state obligations to respect and protect privacy and personal data, recognising that surveillance and data practices underpinning AWS directly shape risks of harm and must be governed accordingly.
3. Reject the use of "national security" as a blanket justification to bypass regulation, and instead require states to demonstrate legality, necessity, proportionality, and independent oversight in all AWS-related data processing, ensuring that security claims do not undermine fundamental rights or international legal obligations.
4. Establish robust transparency requirements and strict limitations on the categories, sources, and uses of data feeding into the development, training, deployment, and operation of AWS and their enabling systems, ensuring that civilian data—particularly social-media-derived or commercially collected data—is not repurposed for targeting or other military functions without clear, lawful, and rights-respecting safeguards.

Privacy International
62 Britton Street
London EC1M 5UY
United Kingdom

+44 (0)20 3422 4321

privacyinternational.org