Introduction

Ireland welcomes the opportunity to contribute a national submission to the UN Secretary-General, as per UNGA Resolution 78/241 on Lethal Autonomous Weapons Systems (LAWS), which Ireland co-sponsored.

We associate this submission with the Chair’s Summary from the 2024 Vienna Conference on “Humanity at the Crossroads: Autonomous Weapons Systems and the Challenge of Regulation”.

The rapid development of autonomous weapons systems (AWS) poses a series of risks which require urgent and comprehensive multilateral action, including:

- **Humanitarian**
  - E.g. Risks of failing to distinguish between civilians and combatants; damage and destruction resulting from a loss of human control.

- **Legal**
  - E.g. Challenges in ensuring compliance with, accountability and responsibility under, IHL; determining state vs individual responsibility; human rights violations (including the rights to life and remedy).

- **Security**
  - E.g. Risks of fuelling an arms race / proliferation (including amongst non-state actors); lowering the threshold for armed conflict.

- **Technological**
  - E.g. Risks of malfunction and errors, cybersecurity vulnerabilities, algorithmic bias, and concerns around explainability and traceability (aka the “Black Box dilemma of AWS and Artificial Intelligence”).

- **Ethical**
  - E.g. Risks of loss of dignity; dehumanisation; discrimination; broader psychological and societal impacts.

Ireland is an active participant within the primary fora for these discussions, the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems (GGE on LAWS) – contributing to the Guiding Principle on Human-Machine Interaction and a working paper on ethical considerations.

Ireland has also participated in regional and cross-regional initiatives on AWS, including:

- Regional conferences in Costa Rica (2023), Manila (2023) and Sierra Leone (2024)¹;

---

¹ [Conferencia Latinoamericana y del Caribe – Impacto Social y Humanitario de las Armas Autónomas (conferenciaawscostarica2023.com)]; [Philippines Calls for Indo-Pacific Voices to Address Lethal Autonomous Weapons Systems Risks (dfa.gov.ph)]; [Sierra Leone Hosts First Regional Conference on Autonomous Weapons - Ministry of Foreign Affairs and International Cooperation (mofaic.gov.sl)].
• The conference on AWS in Vienna hosted by the Austrian Federal Foreign Ministry (2024).²

Ireland has also championed this issue within the UN General Assembly First Committee, through the joint statement on behalf of 70 states in 2022.

This national submission outlines Ireland’s views on how to address the various challenges with regard to AWS. It is underpinned by a human-centred approach, which focuses on:

• A functional characterisation of AWS;
• International law;
• Human control and accountability;
• Ethics;
• Bias; and
• Multilateral solutions.

**Functional Characterisation**

Autonomy is a *feature* of a weapon’s functionality, rather than a specific weapon type. Governance frameworks should therefore focus on the prohibition or regulation of the integration of autonomy into weapons systems, rather than seeking to establish a fixed definition based on specific technical characteristics. Noting that the latter approach is unlikely to remain fit for purpose within the context of rapidly evolving technology.

Ireland uses the following working definition, as outlined by the ICRC, for AWS:

> “Autonomous weapon system’ means a weapon system that is designed to select and engage one or more targets without the need for human intervention after activation”.³

We underline the following non-exhaustive list of functional characteristics which pose legal, humanitarian and ethical concerns if incorporated into AWS:

• The ability to run through a targeting cycle, with the final intention to apply force, without any human intervention;
• The ability to switch to lethal mode without any human intervention;
• The ability to redefine its mission or objective without any human intervention;
• The impossibility to interrupt or deactivate the autonomous mode;
• The absence of a mechanism to interrupt or deactivate the autonomous mode.

**International Law**

International law, including the Charter of the United Nations, international humanitarian law, international human rights law and international criminal law applies fully to all weapons systems, including the development and use of AWS.

---

² 2024 Vienna Conference on Autonomous Weapons Systems – BMEIA - Außenministerium Österreich
³ icrc_submission_on_autonomous_weapons_to_unsg.pdf
IHL compliance requires context-specific, value-based judgment by human beings, which cannot be substituted by autonomous systems. Ensuring AWS' IHL compliance requires consideration of:

- Cognitive limitations of the system (lack of common sense and human judgement);
- Epistemological limitations (i.e. the system making judgments based on biased, incomplete, or inappropriate data);
- Algorithmic bias.

Ireland recognises that AWS that cannot be used in accordance with IHL, must not be developed, deployed or used and are de facto already prohibited under international law. However, new rules, and guidelines, are needed to clarify how IHL applies to AWS.

**Human Control & Accountability**

Ireland adopts a human-centred approach to addressing the challenges posed by AWS, i.e.:

- Prioritising the wellbeing, safety, and dignity of individuals and communities;
- Considering the human impact of arms and armed conflict; and
- Being guided by the principles of international humanitarian law, human rights, ethics, and the protection of civilians.

Ireland considers that IHL compliance requires that human beings must make and remain accountable for decisions on AWS, and must exert full control over AWS – retaining the ability to intervene, interrupt and deactivate AWS throughout their lifecycles.

In determining the extent and quality of human-machine interaction, consideration must be given to:

*Contextual considerations:*

- Whether an AWS is capable of reading the operational context correctly and whether it demonstrates a sufficient level of situational awareness (i.e. able to perceive and react to changing circumstances).

*Technical considerations:*

- Whether adequate limits on tasks and types of targets are in place to allow AWS to be operated with sufficient degrees of reliability and predictability in the identification, selection and engagement of targets;
- Whether adequate environmental limits, including spatial and temporal limits, are in place to ensure that the decisions, made at the planning stage, including legal assessments, are respected throughout the execution stage;

*Forms of human control:*

- Whether human control is exerted and retained over the critical functions of AWS, e.g. identification, selection and engagement of targets;

---

4 2023 March GGE LAWS national statement
5 Microsoft Word - 23 10 23 1C Final UNGA78 EU statement on Conventional Weapons Cluster IV FINAL (unmeetings.org)
• Whether the degree of human control allows for human supervision and intervention to prevent redefinition of AWS' mission without human validation and to interrupt or deactivate the carrying out of autonomous functions if needed.

Individuals remain responsible and accountable for applying the law and this responsibility cannot, under any circumstances, be transferred to machines. This responsibility extends to developers and manufacturers, who bear responsibility in the design and programming stages of the weapon. This is particularly relevant for issues related to data bias, which can impact targeting, and malicious or careless programming.

AWS must not be designed, deployed or used without a clear and unambiguous line of responsibility – any ambiguity would increase the risk of impunity and undermine confidence in the efficacy of any governing framework.

A human chain of command and control must always be ensured during the deployment and use stages of the lifecycle of AWS. It must also be ensured that commanders and operators are informed about any new characteristics, functions and parameters of AWS and are trained accordingly to be able to recognise, understand and interpret these new elements, before deployment or use of AWS in the field.

**Ethics**

Any AWS governance framework must consider not only if utilising an AWS is legally acceptable but also, whether its use would be ethically acceptable.

An ethically aligned approach to AWS should recognise that:6

• Ethical considerations, including the principles of humanity and dictates of public conscience, are central to determining the acceptability of and limits on AWS;
• Ceding human control, agency, responsibility and/or intent on decisions regarding the application of force, to algorithms and computer-controlled processes, poses fundamental ethical concerns;
• Ethical considerations underpin the necessity of establishing limits on autonomy in weapons systems and should be taken into account to determine the necessary type and degree of human control that must be retained over AWS;
• Ethical responsibilities on the decision to apply force require context-specific, value-based judgment by human beings which are neither legally feasible, nor ethically desirable, to delegate to AWS;
• There must be a direct connection, with a clear line of human responsibility and accountability, between the human rationale for AWS deployment and the direct consequences of its use, in order to prevent the erosion of existing ethical standards.

**Bias**

Ireland notes the risks and recommendations pertaining to algorithmic bias in AWS, as outlined in a [joint submission to the GGE on LAWS](https://www.icrc.org/en/ICRC_submits_recommendations_to_UN_Secretary-General) in March 2024:

---

6 These points have been made elsewhere by the ICRC, among others: Autonomous weapons: ICRC submits recommendations to UN Secretary-General | ICRC
“Algorithms and related machine learning risk repeating, amplifying, or contributing to unjust biases that programmers may not be aware of, or that are the result of narrow data selection. Industry and practitioners already acknowledge, as an uncontroversial fact, that human beings consciously or unconsciously encode their own biases into the programmes they write.

There is a growing body of literature documenting and analysing examples of gender and racial biases in AI, which have shown that algorithmic models do this in at least two ways. First, data-based systems reproduce existing inequalities. A 2016 study\(^7\) of a computer program designed to evaluate the potential for recidivism for the criminal justice system found that the system unevenly predicted recidivism among race and genders.

Second, existing data sets and algorithms skew toward white males, meaning that women of colour, for example, are significantly less likely to be intelligible to machine learning programs trained to recognize images\(^8\) and voices\(^9\). When translating between languages, machine learning models are more likely\(^10\) to use the pronoun “he” as opposed to “she” for gender-neutral terms; women, moreover, are less likely\(^11\) to be shown ads for higher-paying employment positions.

Such occurrences of bias are not one-off events. A review\(^12\) of publicly available information on 133 biased AI systems, deployed across different economic sectors from 1988 to 2021, found that 44.2 percent (59 systems) exhibited gender bias and that 25.7 percent (34 systems) exhibited both gender and racial biases.

The consequences of bias in machine learning are amplified in a military context, as a recent UNIDIR\(^13\) report has shown. In the context of autonomous weapons, the criteria that will inform who is and who is not a combatant or target will likely involve factors including gender, age, race, and ability. Biased data sets and poorly programmed algorithms could mean that, for instance, women of colour may be misrecognized at a higher rate, leaving them exposed to differential risks, or that an autonomous system may miscategorise civilian men as combatants, due to their traditional roles in warfare. […]

Recommendations:

Due to the high risk that the integration of algorithmic biases into autonomous weapons poses to fundamental rights and safety, it is essential that any future regulatory framework clearly identifies and recognizes those risks, in a factual manner, and develops appropriate measures to tackle the issue.

The future normative framework on autonomous weapons should include a number of positive obligations and commitments with regard to bias in the algorithms used in autonomous weapons, in order to mitigate the significant risks they pose to fundamental rights and to compliance with international law. Such obligations and commitments may include, but are not limited to:
a) Comprehensive testing and reviews, in order to identify and correct potential biases;
b) Rigorous documentation of the data sets used in autonomous weapons, in order to enhance traceability and transparency, and provide information regarding motivation, the collection process and recommended use;
c) The testing of algorithmic models against benchmarks, that evaluate their operation against gender, age, and race, and in scenarios that are distinct from the data sets used to train the machine learning model; and
d) Comprehensive training and awareness of this issue by those testing and using the system.”

**Multilateral Solutions**

Ireland supports calls for legally-binding rules, and guidelines to address AWS, including the [October 2023 joint call by the UN Secretary-General and the ICRC](https://www.icrc.org). Ireland welcomes and supports the GGE LAWS’ ongoing and critical work, and supports parallel initiatives which are complimentary to the GGE and facilitate an inclusive, global approach.

Ireland supports the ‘two-tier approach’ developed in the GGE on LAWS, whereby certain AWS must be prohibited outright, with all other AWS subject to multilaterally agreed regulation.

AWS must be prohibited:
- If they cannot comply with IHL, including with regard to distinction, proportionality, or if they cause superfluous injury or unnecessary suffering;
- If their effects are not sufficiently predictable, controllable or understandable;
- If they are capable of selecting and engaging targets without the necessary degree of human control to ensure compliance with IHL.

Multilateral limits and regulation should be applied to all other AWS, including:
- Limits on the type of target;
- Limits on the duration, geographical scope and scale of use;
- Ensuring the maintenance of human control to, *inter alia*, ensure effective oversight of AWS, including timely intervention and de-activation.

**Conclusions**

Ireland:
- Joins calls for multilateral rules, and guidelines on AWS – and supports a legally binding solution;
- Underlines the urgency of addressing this issue;
- Underscores the legal, ethical and bias risks associated with the development and use of AWS, and calls for any future framework to take these into account;
- Reiterates the core principles of multilateralism and inclusivity, and the essential roles of the ICRC, civil society, academia, industry and the UN General Assembly.