Convergences in state positions on human control

MAY 2023
REPORT
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>What we looked at</td>
<td>6</td>
</tr>
<tr>
<td>Autonomous weapons systems and human control</td>
<td>6</td>
</tr>
<tr>
<td>The ‘building blocks’ of human control</td>
<td>9</td>
</tr>
<tr>
<td>Predictability</td>
<td>10</td>
</tr>
<tr>
<td>Understandability</td>
<td>13</td>
</tr>
<tr>
<td>Temporal and spatial limitations</td>
<td>15</td>
</tr>
<tr>
<td>Conclusion</td>
<td>18</td>
</tr>
<tr>
<td>References</td>
<td>19</td>
</tr>
</tbody>
</table>
Convergences in state positions on human control

INTRODUCTION

This paper presents an examination of convergences in state positions on human control in the context of autonomy in weapons systems. Since 2017, states have been meeting to discuss autonomous weapons systems at the Convention on Certain Conventional Weapons (CCW) Group of Governmental Experts on Lethal Autonomous Weapons Systems (GGE on LAWS). The creation of this group was preceded by four years of CCW discussions and expert meetings on the issue. In the course of deliberations on autonomous weapons at the CCW, one of the central concerns to have emerged from discussions is the importance of 'the human element' in relation to autonomous weapons systems, and over the use of force itself. Although the preferred terminology in this area remains contested, there is general agreement that the form and nature of human involvement is a critical area for consideration.

The importance of human involvement was further underscored on 21 October 2022, when a cross-regional group of 70 states delivered a joint statement on autonomous weapons systems at the 2022 United Nations General Assembly (UNGA) First Committee meeting. This was the first joint statement on autonomous weapons made at the UN General Assembly, and the largest group statement on autonomous weapons made in any forum to date. The statement, delivered on behalf of the group by Austria, recognised the ‘urgent need’ for the international community to adopt rules and measures on autonomous weapons, and emphasised ‘the necessity for human beings to exert appropriate control, judgement and involvement in relation to the use of weapons systems’. While acknowledging that states may have ‘different understandings of terms like human judgment, control or involvement’, the statement also noted that many states share a recognition that ‘the human element is and must remain central in the use of force’.

That recognition highlights the importance of this ‘human element’ in the design, development, and use of autonomous weapons systems.

Meanwhile, at the recent Latin American and the Caribbean Conference on the social and humanitarian impact of autonomous weapons, held in Costa Rica in February 2023, more than 30 states adopted the Belén Communiqué - the first communiqué to be adopted by a regional meeting on the topic of autonomous weapons systems. The Belén Communiqué, which calls for the urgent negotiation of an international legally binding instrument on autonomous weapons systems.
systems, also states that ‘it is paramount to maintain meaningful human control to prevent further dehumanization of warfare, as well as to ensure individual accountability and state responsibility.’

The Belén Communiqué and the UNGA First Committee joint statement demonstrate the now widespread acceptance that autonomous weapons raise serious humanitarian, ethical, legal, security and technological concerns, and that international rules and measures are necessary to address these concerns. This has already been accepted by the majority of states involved in the CCW GGE on LAWS discussions.

For many states involved in the international discussion on the autonomous weapons issue, the necessity of a human element is represented by this standard of human control over autonomous weapons systems, with states variously referring to such control as ‘meaningful’, ‘appropriate’, ‘sufficient’, or ‘effective’. Other states may use different phrasing when describing the human element, using phrases such as ‘meaningful and context-appropriate human involvement’, or preferring to simply refer to ‘human-machine interaction’. Despite differences in phrasing when discussing the human element as it relates to autonomous weapons systems, there is much common ground among states in their understandings and requirements as to what ‘meaningful’, ‘sufficient’, or ‘appropriate’ levels of human control or involvement requires. This paper aims to shed light on where states share such understandings, in particular in three key areas which provide technical underpinnings for rules on human control in autonomous weapons systems. These include:

- Predictability;
- Understandability;
- Temporal and geographic limitations on use.
WHAT WE LOOKED AT

For the purposes of this paper, we looked at states’ contributions to the UN CCW GGE on LAWS. While, as is noted above, some states have commented on the need to maintain human control over autonomous weapons systems in other fora (such as at the UNGA), it is at the CCW where states have further elaborated on their understandings of the human element.

AUTONOMOUS WEAPONS SYSTEMS AND HUMAN CONTROL

The International Committee of the Red Cross (ICRC) describes autonomous weapons systems as ‘systems which select and apply force to targets without human intervention’, wherein ‘after initial activation or launch by a person, an autonomous weapon system self-initiates or triggers a strike in response to information from the environment received through sensors and on the basis of a generalized “target profile”’. This means that the specific object to be attacked, and the exact time and place of the attack, are determined by sensor processing, not by humans. As a result, and as the ICRC has pointed out, ‘this means that the user does not choose, or even know, the specific target(s) and the precise timing and/or location of the resulting application(s) of force.’ Given the serious humanitarian, legal, and ethical concerns raised by such weapons systems, it is therefore imperative that a form of human control over the critical functions of such systems is maintained, in particular in relation to issues of target specification, and over the duration and area of system functioning.

At the first session of the 2023 GGE on LAWS, Ambassador Flavio Damico of Brazil, Chair of the group, noted that ‘the debate on human control/judgment is linked to specific challenges and risks regarding compliance with IHL.’ Ambassador Damico further noted that ‘the role of human-machine interaction, human control or human judgment, as presented in past iterations of the GGE, is to ensure first that the use of LAWS is in compliance with international law, in particular international humanitarian law; secondly that human responsibility is retained; and thirdly that the operation of such systems remains within a responsible chain of command’, while also pointing out that views may differ on ‘how exactly to characterise the quality and extent of human-machine interaction’, and that the terminology used ‘is manifold’. However, and as was noted by Switzerland at the same GGE, it is also true that the GGE is ‘reaching a point where differences in mere terminology carry less sway, whether we call it human-machine interaction, human judgment, human involvement, and whatever adjective we put in front of human control’, with discussions now:

‘...shifting from terminology to concrete and operationalizable concepts of control or involvement. What stands in the centre is really the idea that humans must take measures at different stages in the life-cycle of a weapon, including in the engagement of autonomous weapon systems, and understand their functioning in appropriate ways.’

The importance that states place on human control in the use of autonomous weapons is evident from the breadth of statements delivered by states both individually, and as part of a group, at the CCW and at the UNGA. For example (please note that this is not an exhaustive list):

...
• Algeria has stated that 'human control of weapons systems must be permanent during the entire life cycle'.

• Argentina has stated that autonomous weapons systems 'that target, engage and apply force, inter alia, in deciding on the life of human beings without meaningful human control are unacceptable and must be prohibited under international law'.

• Austria has stated that ‘there is still time to prevent humanity from crossing a very dangerous threshold by adopting legally binding norms to safeguard human control over autonomous weapons systems'.

• Belgium has stated that “if you do not have human control that is meaningful then there are questions about the capacity of those systems to be used in accordance with applicable international law, especially IHL”.

• Benin, delivering a statement on behalf of the African Group at the 2018 CCW GGE on LAWS, stated that human control over weapon systems should not be seen as a matter of good-will by States but a legal standard that they ought to fully abide by.

• Brazil has stressed ‘the need of a regulation that recognizes the centrality of human control in the development and use of autonomous systems, in accordance with the norms and principles of International Humanitarian Law'.

• Bulgaria has stated that human control should be ‘exerted and retained over the whole life cycle of an autonomous weapons system in order to ensure that deployment of such weapons systems fully complies with the international humanitarian law'.

• Cambodia has stated that autonomous weapons systems ‘should be utilised at the helm of meaningful human control in assurance to accountable and ethical judgement for preliminary deployment'.

• Chile has stated that ‘there is an implicit requirement for meaningful human control embedded in IHL, notably and vis-a-vis the principles of distinction, proportionality, and military necessity among others'.

• China has stated that ‘we must ensure that weapons systems are always under human control, we should avoid their abuse, misuse and malicious use to ensure that relevant systems and combat methods are in line with International humanitarian law'.

• Costa Rica has stated that ‘meaningful human control is essential across the life cycle of weapon systems'.

• Cuba has stated a legally binding instrument on autonomous weapons should include ‘the general obligation to ensure meaningful human control in the use of force'.

• Denmark, in a statement on behalf of the Nordic countries delivered at the UN General Assembly in 2022, stated that it is important to advance work on applicable principles and norms for autonomous weapons systems ‘in order to ensure meaningful human control and accountability for decisions over the use of force'.

• Ecuador has stated that negotiating a legally binding instrument on autonomous weapons is necessary for ‘at least ensuring that there is meaningful human control over these weapons systems'.

• El Salvador has stated that it is ‘concerned with use of weapons without human control'.

• France has stated that the ‘centrality of the human element’ is a matter of consensus in the GGE, noting specifically ‘maintaining appropriate human control during the whole life-cycle'.

• Germany has stated that ‘weapon systems operating completely outside a human chain of command and control must not be developed and used and that a framework of human control must be retained for all other weapon systems featuring autonomy'.

• Guatemala has stated that it must be ensured ‘that no weapon or arm can operate without significant meaningful human oversight and control at the time of deciding whether an arm will be deployed, because that is making a decision between life and death'.

• The Holy See has called for ‘a moratorium on the development and use of LAWS pending the negotiation of a legal instrument that prohibits such systems from targeting humans and ensures that all weapons systems...
remain under meaningful, human control.\textsuperscript{28}

- Ireland has stated that ‘we should be clear that a degree of human control must be retained over weapons systems in order to allow for compliance with IHL’ and that it is ‘essential to ensure that meaningful human control is exerted and retained over the critical functions of the weapons system’.\textsuperscript{30}

- Japan has stated that ‘Regarding the form of human control, it is indispensable that a lethal weapon system be accompanied with meaningful human control by securing proper operation and be operated by persons with sufficient information on such weapons systems’.\textsuperscript{31}

- Liechtenstein has stated that a legal instrument should ‘place prohibitions on lethal autonomous weapon systems that cannot be used with meaningful human control across its entire life cycle’.\textsuperscript{32}

- Luxembourg, in a statement with 7 other countries at the 2021 CCW GGE on LAWS, said that ‘human control must be retained over weapons systems in order to allow for compliance of IHL and that control over critical functions of the weapon system should be applied throughout the life-cycle of the weapon’.\textsuperscript{33}

- Mexico has stated that there should be a focus on ‘preserving meaningful human control in weapons which incorporate autonomous functionalities’.\textsuperscript{34}

- The Netherlands has stated that the main objective of regulation for autonomous weapons should be ‘to ensure that sufficient human judgement and control is maintained when such systems are developed and used’.\textsuperscript{35}

- New Zealand has stated that autonomous weapons which fall below the threshold of prohibitions should be ‘regulated through positive obligations to ensure that we safeguard appropriate standards of human involvement or control and preserve accountability under IHL’.\textsuperscript{36}

- Nigeria, while not a state party to the CCW, has said in a working paper submitted along with 22 other states to the CCW GGE on LAWS in 2022 that ‘human beings must make the decisions with regard to the use of force, exert control over weapons systems that they use, and remain accountable for decisions over the use of force in order to ensure compliance with International Law, in particular International Humanitarian Law’.\textsuperscript{37}

- Norway has stated that ‘we must ensure that weapons systems featuring autonomy remain under meaningful human control’.\textsuperscript{38}

- The State of Palestine has stated that ‘meaningful human control is a fundamental and necessary condition that must govern any and all autonomous weapon systems’.\textsuperscript{39}

- The Philippines has stated that ‘we contend that any weapon systems whose autonomous functions are designed to be used to conduct attacks outside meaningful human control must be prohibited’.\textsuperscript{40}

- Qatar, delivering a statement on behalf of the Arab Group at the first session of the 2023 CCW GGE on LAWS, reiterated the ‘importance of maintaining meaningful human control’ over autonomous weapons systems.\textsuperscript{41}

- Russia, in its working paper ‘Concept of Activities of the Armed Forces of the Russian Federation in the Development and Use of Weapons Systems with Artificial Intelligence Technologies’, circulated at the CCW GGE on LAWS in March 2023, has stated that there is a ‘need to maintain human control over a weapon system with AI technologies’.\textsuperscript{42}

- Sri Lanka has stated that it continues to stress ‘the importance of centrality of meaningful human control in the full lifecycle of the weapon system’.\textsuperscript{43}

- Sweden has stated that it ‘is of the strong conviction that international law applies fully to all weapon systems and that meaningful human control over the use of force always must be upheld; the area of LAWS is no exception’.\textsuperscript{44}

- Switzerland has noted that there is ‘strong support for human control and judgement and specific risk mitigation measures all through the life cycles of these kinds of weapon systems’.\textsuperscript{45}
Convergences in state positions on human control

- Turkiye has stated that ‘the development and use of LAWS should have meaningful human control.”
- The United Kingdom has stated that it opposes ‘the creation of and use of systems that would operate without meaningful and context appropriate human involvement throughout their life cycle.’
- Venezuela, delivering a statement on behalf of the Non-Aligned Movement (NAM) at the 2023 CCW GGE on LAWS, stated that ‘weapons that function without meaningful human control should be prohibited.”

Statements delivered by the African Group, the Arab Group, and the NAM have also stressed the importance of human control. The African Group has stated that ‘there is a general consensus among States on the need to maintain human control over use of weapon systems’, noting that:

‘For that reason, in adopting General Comment Number 3 on Article 4 of the African Charter on Human and People’s Rights concerning the right to life, African states agreed in paragraph 35 that “Any machine autonomy in the selection of human targets or the use of force should be subject to meaningful human control”’.

The Arab Group, at the first session of the 2023 GGE on LAWS, reiterated ‘the importance of maintaining meaningful human control’ to limit the negative impact of autonomous weapons systems. The Non-Aligned Movement has stated that it is ‘of the view that all weapons, including those with autonomous functions, must remain under the direct control and supervision of humans at all times.” The European Union has stated that ‘human beings must make the decisions with regard to the use of lethal force’ and ‘exert control over lethal weapons systems that they use’.

THE ‘BUILDING BLOCKS’ OF HUMAN CONTROL

It is clear from the above, as Switzerland observed in the first session of the 2023 CCW GGE on LAWS, that ‘the role of humans in the use of autonomous weapon systems has emerged as the central element in ensuring compliance with international humanitarian law’, and that states:

‘...now have a shared understanding that a certain degree of control over the use of force is not only necessary to comply with IHL, but is also important from an operational perspective and in terms of ethical acceptability...’

Further, and as referenced earlier in this paper, Switzerland also noted that discussions are shifting to ‘concrete and operationalizable concepts of control or involvement’, with contributions from states making it clear that ‘in order to be controllable and to meet legal requirements, autonomy needs clear limits’.

Of course, human control over autonomous weapons is context-dependent; as the Stockholm International Peace Research Institute (SIPRI) and the ICRC have pointed out, the particular combination of measures needed to ensure human control ‘will vary according to the specific context. Where one type of control measure is inadequate, insufficient or challenging to implement, the other types may rise in prominence.” This is widely recognised by states, and is also understood by civil society.
This paper provides an overview of state positions and convergences in three of the more concrete components of control, or ‘building blocks’ of human control, including:

- Predictability
- Understandability
- Temporal and geographic limitations on use

It is by creating rules in terms of such operationalizable technical concepts that an overarching requirement for human control that is meaningful or sufficient can be formalised and codified.

These building blocks are interconnected. For a system to be ‘predictable’ implies that the user of the system is able to sufficiently predict the behaviour of that system in a given context of use. In order to adequately predict the behaviour of a system, the user must sufficiently understand how the system functions and they must sufficiently understand the contextual features that it will encounter in a specific instance of use. Temporal and geographic limitations are therefore required in order to render those contextual evaluations possible.

**PREDICTABILITY**

As identified above, a foundational building block of human control emerging from discussions on autonomy in weapons systems is the necessity for systems with autonomous functions to be sufficiently predictable. The ICRC argues that ‘a degree of unpredictability is inherent in the effects of using all AWS due to the fact that the user does not choose, or know, the specific target(s), and the precise timing and/or location of the resulting application(s) of force’, and that ‘unpredictability in AWS poses a fundamental challenge to IHL.’ As such, the ICRC recommends that ‘unpredictable autonomous weapon systems should be expressly ruled out,’ with:

‘a prohibition on AWS that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained. This prohibition would build on the recognition by States of the need for sufficient predictability in the use of AWS for compliance with IHL and for practical military operational reasons.’

Linked to the concept of predictability is the requirement that systems should be reliable: that the system should reliably function in accordance with the user’s understanding and expectation.

States’ written contributions to the CCW GGE on LAWS also demonstrate the importance of predictability. For example:

- In its 2020 contribution on the Guiding Principles, Austria stated that ‘Predictability and reliability are crucial for IHL compliance as both contribute to the ability to estimate the expected effects and results of a particular weapon use.’
In 2021, a group of 10 states wrote that ‘Sufficient levels of predictability, foreseeability, reliability, oversight, and explainability of weapon systems, as well as spatial and temporal constraints, are needed to enable operators to exercise meaningful human control, ensure legal compliance and avoid technical vulnerabilities.’ In their 2022 joint working paper, Finland, France, Germany, the Netherlands, Norway, Spain and Sweden stated that during the development phase of an autonomous weapons system, states should ‘evaluate the reliability and predictability of the system, by applying appropriate testing and certification procedures, and assess compliance with IHL through legal reviews’, and that humans should ‘at all times have sufficient assurance that weapons systems, once activated, act in a foreseeable manner.’

All five working papers submitted to the March 2023 meeting of the GGE on LAWS referenced the concepts of reliability and/or predictability. For example:

- The State of Palestine’s 2023 proposal for the normative and operational framework on LAWS states ‘an AWS must be sufficiently predictable. This means that the human user(s) must be able to reasonably anticipate how the system will function (so that its effects are in accordance with the user’s intent);’ and that ‘an AWS must be sufficiently reliable. This means that it must be capable of performing as expected to, consistently, safely, and securely.’
- Pakistan’s 2023 proposal for a legal instrument states that an autonomous weapons system which has effects that ‘cannot be adequately predicted, understood and explained’ should be prohibited’, and that ‘weapons with “human control over the decision to use force” must “be designed to be transparent, indubitable, reliable and predictable in the identification and selection of potential targets”.
- The ‘Draft articles’ submitted by Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States includes the assertion that ‘to ensure effective implementation of the principle of distinction in attacks involving the use of autonomous weapon systems’, the autonomous weapon system ‘needs to perform with adequate reliability to enable, in circumstances of its use, force to be directed against such targets or to remain within such locations.’

Oral statements delivered at the March 2023 meeting of the GGE on LAWS further underscore the importance of predictability and reliability:

- Switzerland noted that ‘most proposals revolve around comparable, even complimentary concepts such as predictability.’
- Canada suggested that one of the ways to mitigate risk or respond to contingencies includes ‘improve safety, reliability, and predictability of the system’.
- The Philippines stated that the Austrian and Palestinian working papers provided a ‘crisp analysis’ of the elements that the GGE could further develop, namely ‘predictability, reliability, understandability, explainability and traceability’.
- Tunisia noted that, in the case of autonomous weapons systems, ‘after activation, the context and environment can be changed at any time, and the predictability can fail. Also, the reliability of the system can be degraded. Therefore, for lethal autonomous weapons systems, human supervision and control need to be exercised and retained even after activation.’
- Ireland stated that technical considerations to be considered in the future normative and operational framework on autonomous weapons systems should include ‘whether adequate limits on tasks and types of targets are in place to allow the weapon system to be operated with sufficient degrees of reliability and predictability.’
Automated Decision Research

- The United Kingdom noted that there is a need for those involved in the use of LAWS to understand the 'performance and predictability of autonomous functions' within its context of use.71

- The Netherlands stated that 'at all times, humans must have sufficient assurance that weapons systems, once activated, act in a foreseeable manner', and noted 'the importance of predictability and controllability, as stated in many of the proposals' and 'the importance of understandability, traceability and reliability'.72

- The State of Palestine said that, to be able to use an autonomous weapon system with meaningful human control, the system must be 'sufficiently predictable, reliable, understandable, explainable and traceable', noting further that 'if the user of an autonomous weapon system cannot reasonably anticipate what would trigger an application of force, its use would inter alia undermine the prohibition of weapons that are indiscriminate by nature. That is why a fundamental component of meaningful human control is for these systems to be sufficiently predictable.' On reliability, the State of Palestine added that 'meaningful human control requires that an autonomous weapon system must be sufficiently reliable. This means that it must be capable of performing as expected as affected consistently, safely and securely.'73

- Austria stated that 'ensuring meaningful control sets out a multidimensional approach which relates to the level of predictability and reliability of autonomous weapons...'.74

- Norway noted that, at the first session of the 2023 GGE on LAWS, there was 'broad convergence' around the relevance of 'predictability and foreseeability' and 'explainability, reliability and related testing'.75

- Denmark noted that 'appropriate testing and certification procedures, including legal reviews' can 'enhance the reliability and predictability of the system...'.76

- The Holy See stated that the principles of adequate, meaningful and consistent human supervision 'imply the requirements of predictability and reliability of autonomous systems'.77

- Italy stated that it is 'crucial to maintain the reliability and verification of autonomous weapons systems throughout their entire lifecycle'.78

- Pakistan stated that there should be 'transparency, reliability and predictability in the identification and selection of potential targets'.79

- Sri Lanka noted that weapons systems which are 'inherently unpredictable and cannot be used in accordance with IHL are de facto prohibited'.80

Predictability arises from the ability of the user to understand the system and to understand the context of use. As SIPRI and the ICRC have noted:

‘...understanding and predicting the behaviour of an AWS is essential for achieving safety, military efficiency and most importantly legal compliance. In order for an AWS to be approved for use by both a system safety and legal review, the functioning of the system must be understandable, predictable and explainable.’81
UNDERSTANDABILITY

In order to ensure the adequacy of the human element in the use of autonomous weapons systems and to ensure predictability, it is broadly agreed that it is necessary that those responsible for designing, deploying and operating such systems sufficiently understand how the system functions and the likely effects of attack resulting from the use of that system in an operational context. This means that the human operator and those in command must adequately understand how a system works, how it is likely to function in the intended context of operation, and be able to sufficiently understand and explain, after the use of such a system for an attack, why a system operated in a particular way.

In the background paper to its position on autonomous weapons systems, the ICRC has noted that certain ‘machine learning techniques make it extremely difficult for humans to understand and, therefore, to predict and explain the process by which an AWS functions (the “black-box” challenge), irrespective of its environment of use.’ Therefore:

‘...if an AWS’ functioning is opaque, then humans responsible for the application of IHL rules - both persons entrusted with the legal review of an AWS and persons responsible for compliance with IHL during its use - could not reasonably determine its lawfulness under IHL. The functioning could be opaque notably due to reliance on artificial intelligence and machine learning techniques, or because it changes during use in a way that affects the use of force (e.g. machine learning enables changes to targeting parameters over time).’

For a system to be amenable to lawful use, it must be sufficiently understandable and explainable, so that users are able to sufficiently predict the likely effect and outcome of a system’s use.

In the context of human-machine interaction/control/involvement, understandability has emerged as a key component of state’s understanding of the ‘human element’ in the use of autonomous weapons systems. This is readily apparent in written submissions to the CCW GGE on LAWS.

For example:

- The United States, in its 2020 ‘U.S. commentaries on the Guiding Principles’ document, proposed in relation to human-machine interaction that ‘Relevant personnel should properly understand weapons systems based on emerging technologies in the area of LAWS. Training, doctrine, and tactics, techniques, and procedures should be established for the weapon system. Operators should be certified by relevant authorities that they have been trained to operate the weapon system in accordance with applicable rules.’

- In 2021, in a joint submission by a group of 10 states, the group suggested that ‘Inability to understand or explain, which both reduce the operator’s control over the system and prevent investigation after the fact’ would render an autonomous weapon system incapable of being used in compliance with IHL.

- In 2022, a group of 7 states, in a joint working paper, stated that, in order to retain appropriate human control, ‘developers, commanders and operators - depending on their role and level of responsibilities - must have a sufficient understanding of the weapons systems’ way of operating, effect and likely interaction with its environment. This would enable the commanders and operators to predict (prospective focus) and explain (retrospective) the behavior of the weapons systems.’
At the 2023 meeting of the CCW GGE on LAWS, four of the five written submissions referenced understandability. For example:

- Austria’s working paper states that those authorising the use of a ‘weapons system that integrates autonomy in its critical functions of selecting and applying force to a target’ must have ‘an adequate functional understanding of the system under consideration. They must understand what circumstances or conditions will trigger an application of force by the system, including conditions that would trigger an unintended engagement’.

- The State of Palestine wrote that in order to be able to use a system with meaningful human control, ‘the autonomous weapons system must be sufficiently predictable, reliable, understandable and explainable, and traceable’, and further stated that an autonomous weapon system ‘must be sufficiently understandable and explainable. This means that the human user(s) must have a sufficient understanding of how the system functions and what circumstances will trigger an application of force, to comply with legal and other requirements.’

Further, the first session of the CCW GGE on LAWS in March 2023 saw numerous oral statements made in a discussion on human-machine interaction/meaningful human control, including:

- Switzerland stated that ‘what stands in the centre [of human control] is really the idea that humans must take measures at different stages in the life cycle of a weapon including an engagement of autonomous weapons systems and understand their functioning in appropriate ways.

- Canada stated that the ‘education and training of the human operators and relevant decision-makers, as well as their understanding of and experience with the weapon system in question’ should be considered when formulating policies and doctrine for autonomous weapons systems.

- The Philippines noted that elements that should be further developed in terms of meaningful human control include ‘predictability, reliability, understandability and explainability, and traceability’.

- Germany noted that in its 2022 working paper, one of the suggested technical prerequisites for human-machine interaction is ‘proper training for commanders and operators so that they properly understand the system and its effects’.

- Turkie reiterated that states should ensure that ‘operators and commanders are sufficiently trained in order to understand the capabilities and limitations of the autonomous weapons systems they use’.

- Tunisia underscored that after deployment, an operator should activate a system ‘based on full understanding of the system’.

- Argentina noted that there is a ‘high level of convergence between delegations on the need for an operator to be part of a responsible chain of command and that he should have sufficient knowledge of how the system functions’.

- The United Kingdom said that ‘there appears to be broad agreement that there is a need for those involved in the use of LAWS to understand the context, situation and environment within which the system is being used’ and ‘have an understanding of the performance and predictability of autonomous functions within this context of use’.

- The Netherlands stated that ‘developers, commanders and operators must have a sufficient understanding of the weapon system’s way of operating, effect, and likely interaction with its environment’.

- The State of Palestine said that ‘in order to be used with meaningful human control, the system must be sufficiently predictable, reliable, understandable, explainable and traceable’.
Mexico stated that ‘it is vital to guarantee the necessary, context-specific value judgment which is necessary for the application of the principles of IHL, also the viability of the human being to understand and explain what the system is doing, and why it’s doing it’;

France underlined that, in order for human control to be effective, the ‘proper understanding by operators and of the command, of the impacts of the system and of its interaction with its environment’;

Austria reiterated the importance of having ‘an adequate understanding’ of an autonomous weapons systems ‘functioning, while taking into account the concrete scenario in which it would operate’;

Denmark noted that understandability is one of the elements that underpins ‘appropriate human control’, and the importance of ‘sufficient education and training of developers and operators to enable them to reasonably predict and explain the function or behaviour of the system and its likely interaction with the environment’;

Norway noted that, at the first session of the 2023 GGE on LAWS, there was ‘broad convergence’ around the relevance of ‘understandability and related training’.

Being able to adequately predict how a system will function requires an understanding of how a system will function in a specific context of use, which in turn requires ‘a specific context’: one that is bounded by temporal and spatial limitations. As SIPRI and the ICRC have argued:

‘Greater mobility and duration of operation generally increases unpredictability, because the area of operation becomes larger and the environment changes over time. Therefore, important control measures are restricting mobility to a specific area in which the AWS can apply force to a specific type of target; and limiting the duration for which the AWS is activated.’

TEMPORAL AND SPATIAL LIMITATIONS

Linked to the above concepts of predictability and understandability is the assertion that in order to ensure meaningful/effective etc. control in the use of autonomous weapons systems, temporal and spatial limits - i.e. limits on the timeframe of operation and the geographic area of operation - should be imposed on autonomous systems.

The ICRC recommends that ‘the design and use of autonomous weapon systems that would not be prohibited should be regulated’ including through ‘limits on the duration, geographical scope and scale of use, including to enable human judgement and control in relation to a specific attack.’ Such limits would, for example:

‘...aim to enable AWS users to have the necessary situational awareness to anticipate the effects of an attack and be reasonably certain upon launching the attack that it will comply with IHL. These limits also reduce the risk that circumstances may change during an attack and facilitate supervision during the operation of the AWS.’
Many states have coalesced around limits on duration and geographical scope of use, both in written submissions and oral statements to the CCW GGE on LAWS.

For example:

· In 2020, in its ‘food for thought’ paper on considerations on the appropriate level of human involvement in LAWS, Finland wrote that: “To enable planning and governance of the autonomous functioning of LAWS, clear pre-defined boundaries need to be set. These include geographical coordinates, the allowed time window, and environmental conditions for authorised operation, but also system specific limitations, preconditions, rules of engagement, etc.”

· In 2021, in their joint working paper, Costa Rica, Panama, the Philippines, Sierra Leone, and Uruguay wrote that: ‘sufficient levels of predictability, foreseeability, reliability, oversight, and explainability of weapon systems as well as spatial and temporal constraints are needed to enable operators to exercise meaningful human control, ensure legal compliance and avoid technical vulnerabilities.’

· In 2022, a group of 23 states committed to work: ‘collaboratively to identify and agree on limits and other regulations to uphold the rules of international humanitarian law for other types of autonomous weapons systems, including through… Limits on the duration, geographical scope and scale of use.’

In 2023, all five working papers submitted to the first session of the GGE on LAWS in March 2023 referenced limits on duration and geographical scope of use. For example:

· Austria’s working paper indicates that an inability to limit the duration and geographical area of a system’s functioning, as well as the number of engagements that a system can undertake, means that a system is unacceptable. As such, ‘systems that cannot meet these conditions should be prohibited.’

· In their ‘Draft articles on autonomous weapon systems’ document, Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States listed as one of the ‘Consideration of potential precautions or features to be implemented in the design and use of the system to mitigate the risk of harm to civilians and civilian objects’ includes measures to ‘control, limit, or otherwise affect the duration, geographical scope, and scale of the operation of the weapon system...’

· Russia’s 2023 ‘Concept of Activities of the Armed Forces of the Russian Federation in the Development and Use of Weapons Systems with Artificial Intelligence Technologies’ states that ‘The Concept enshrines the need to maintain human control over machine to uphold compliance with the existing norms of international law’, and that ‘such control should be exercised’ by means of a number of limits, including ‘limitations on duration of operation, geographic scope and scale of use.’

· Pakistan’s ‘Proposal for an international legal instrument’ states that weapons ‘with human control over the use of force’ must ‘be designed to limit their scope and scale of use, including temporal and spatial limits.’

The widespread coalescence around the need for such limits was also evident in states’ oral statements at the March 2023 GGE meetings:

· Russia stated that ‘establishing limits on tasks, missions, duration and the zones of application of [lethal autonomous weapons systems] could enhance predictability and in so doing contribute to upholding IHL’.

· Ireland stated that technical considerations in the context of human control in the future normative and operational framework on autonomous weapons systems should include ‘adequate environmental limits including spatial and temporal limits.’
The United Kingdom stated that there is a need for those involved in the use of LAWS to ‘have the ability to impact the behaviour of the system and its effects on the environment and the entities within it, including banding the system temporally and geographically, and the targets that can be engaged.’

The Netherlands stated that control must go beyond ‘nominal control’, and as such it should be ensured that humans ‘will be in a position to inter alia set spatial and temporal limits that may vary according to the situation and context.’

The State of Palestine stated that ‘autonomous weapon systems should only be capable of operating for short periods of time; they should only be able to operate within a limited geographical space…’

France stated that ‘there is a need to maintain rigorous human control when it comes to the scope of utilisation, with a spatial, temporal and regulatory level’ and that a lethal system ‘integrating autonomy should remain subordinate to human command and its decision-making functions should remain dependent on a spatial, temporal regulatory framework’.

Austria noted the need to ‘limit the geographical area, duration of use and number of engagements’ of an autonomous weapons system.

Denmark observed that many delegations highlighted ‘setting geographical and temporal limits taking into consideration the context and the environment that the system operates in’ and noted its agreement with these.

Norway noted the ‘broad convergence’ around the relevance of ‘limits on targets, temporal and geographical scope’.

Pakistan stated that there should be limits on the scope and scale of use of autonomous weapons systems, including ‘temporal and spatial limits.’
CONCLUSION

States have demonstrated through statements at the UNGA, the Belén Communiqué, and contributions to the CCW GGE on LAWS that human control in the use of autonomous weapons systems - whatever terminology is used by each individual state - is of significant concern and importance. It is also clear that states agree that there is a need for autonomous weapons systems to:

- Be predictable;
- Be understandable;
- Have limits on duration and geographic scope of use.

The general recognition of the need to ensure human control, or the 'human element', in the use of autonomous weapons systems thus finds clarity through specific rules on predictability, understandability, and temporal and geographic limitations of use. This is coupled with an understanding that control is context-dependent and that these requirements will interact in different ways depending on the system in question and the context of use. These convergences indicate that, as Ireland observed at the March 2023 meeting of the CCW GGE on LAWS, the differences found among states in the area of human control/involvement are now 'largely differences of nomenclature.'

ABOUT AUTOMATED DECISION RESEARCH:

Automated Decision Research is the monitoring and research team of Stop Killer Robots - tracking state support for a legally binding instrument on autonomous weapons systems and conducting research and analysis on responses to autonomy and automated decision-making in warfare and wider society.


109 Working Paper submitted to the 2022 Chair of the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems (LAWS) on behalf of Argentina, Austria, Belgium, Chile, Costa Rica, Ecuador, Guatemala, Ireland, Kazakhstan, Liechtenstein, Luxembourg, Malta, Mexico, New Zealand, Nigeria, Panama, Peru, the Philippines, Sierra Leone, Sri Lanka, State of Palestine, Switzerland, and Uruguay (2022), CCW GGE on LAWS. Available at: <https://reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2022/gge/documents/WP5-Rev1.pdf>

110 Austria (2023), Revised working paper, CCW GGE on LAWS. Available at: <https://docs.library.uno.org/Convention_on_Certain_Conventional_Weapons-/Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE2023_WP1_Rev1.pdf>

111 Australia, Canada, Japan, the Republic of Korea, the United Kingdom and the United States (2023), Draft articles on autonomous weapon systems – prohibitions and other regulatory measures on the basis of international humanitarian law (‘IHL’), CCW GGE on LAWS. Available at: <https://docs.library.uno.org/Convention_on_Certain_Conventional_Weapons-/Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE2023_WP4_Rev1.pdf>


114 Statement by Russia, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=4567&channel=ENGLISH>

115 Statement by Ireland, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=5063&channel=ENGLISH>

116 Statement by the United Kingdom, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=6078&channel=ENGLISH>

117 Statement by the Netherlands, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=6338&channel=ENGLISH>

118 Statement by the State of Palestine, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=7495&channel=ENGLISH>

119 Statement by France, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=9073&channel=ENGLISH>

120 Statement by Austria, CCW GGE on LAWS, 08 March 2023. Available at: <https://conf.unog.ch/digitalrecordings/index.html?guid=public:610500/8F5153B6-AAB5-4AFC-85C2-0BA55F743D04_10h16&position=9073&channel=ENGLISH>