

A Path towards the Regulation of LAWS

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:

(2020). A Path towards the Regulation of LAWS. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-77411-1>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

A Path towards the Regulation of LAWS

iPRAW Briefing – May 2020

This paper summarizes the recommendations by the *International Panel on the Regulation of Autonomous Weapons* (iPRAW) on steps towards a regulation of lethal autonomous weapon systems (LAWS) within the Convention on Certain Conventional Weapons (CCW). It breaks down the path to regulation into six steps:



Why do we even start walking?

The legal objective is to abide by IHL and to protect humans from arbitrary decisions. Ethical, operational, and security considerations add to that.



Where do we start?

A regulation of LAWS should focus on the human involvement in the use of force, not on a technical definition.



What will we find along the way?

A regulation needs to explore the human element, e.g. human control: The subject would be the process of use enabled by design specifications.



How big are our steps?

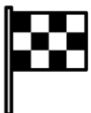
It might be a rather abstract regulation with regard to the type and level of human control, but it could encompass all steps in the weapon's life cycle.

The operational context is crucial and could be explored further in more dynamic documents.



Bonus – Should we get ice cream (here: verification) along the way?

There is probably no one-size-fits-all solution to verification. It would be challenging but not necessarily impossible.



Where does that leave us?

A – politically or legally binding – regulation should operationalize and apply the principle of human control.

iPRAW recommends to operationalize and apply the principle of human control in a more concrete norm. This could be done by a legally or politically binding agreement. The next steps for the GGE 2020/21 could be to (1) **abandon** the distracting debates about a technical definition of LAWS, (2) **focus** on the human element (e.g. human control), and (3) **analyze** the impact of the operational context on the necessary level of human involvement.

1. THE OBJECTIVE: WHY DO WE EVEN START WALKING?

The core of any regulation of LAWS within the framework of the CCW would be the adherence to **international humanitarian law** (IHL) during attack. That includes the protection of civilians and wounded from attacks (i.e. distinction), an adequate ratio between expected effect and benefit (i.e. proportionality), a military reason behind every attack (i.e. necessity), and the application of those principles not only during preparation but also during execution of an attack (i.e. precaution). As the International Committee of the Red Cross (ICRC) argues, such a decision making process calls for human involvement:¹ Those legal decisions require an understanding of the situation at hand, they are highly subjective and context-dependent as pointed out by Guiding Principles c and d.²

In addition, humans – civilians, hors de combat, and active soldiers alike – are to be protected from arbitrary decisions. This is rooted in international human rights as well as ethical standards calling for human dignity. A central argument on that regard is the necessity for moral agency to make a deliberate decision and to avoid degrading the human target to an object. Moral agency, however, cannot be offered by a machine hence calling for human control.³

2. FOCUS ON THE HUMAN ELEMENT: WHERE DO WE START?

The considerations on the objective of a regulation of LAWS reveal the necessary starting point for a politically or legally binding document: the human element. The control exercised by the operator must be sufficient to reflect their intention for the purpose of establishing the legal accountability and ethical responsibility for all ensuing acts.⁴

This stands in opposition to requests for a technical definition of LAWS as a first step towards any regulation. A technical definition would most likely not be future-proof as technology progresses or would have to remain quite vague. More importantly, though, **in the CCW context a technical definition would be beside the point**: even though technologies like data-driven computational methods (i.e. artificial intelligence, machine learning) enable many autonomous functions, **the challenges described above do not stem from a particular technology but from the lack of human involvement**. If CCW states parties wanted to create a definition of LAWS, a technology agnostic one with a focus on the ‘autonomous’ functions (instead of platforms) would suffice. The ICRC presented such an admittedly broad definition of an autonomous weapons system already: “Any weapon system with autonomy in its critical functions—that is, a weapon system that can select (search for, detect, identify, track or select) and attack (use force against, neutralize, damage or destroy) targets without human intervention.”⁵ It might need some additions to exclude rather unproblematic existing systems like purely defensive, anti-materiel ones.

¹ See International Committee of the Red Cross (April 2018), *Statement at the CCW GGE on LAWS: Further consideration of the human element in the use of lethal force; aspects of human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems*, <[https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/5216D20D2E98E7AAC12582720057E6FC/\\$file/2018_LAWS6b_ICRC1.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/5216D20D2E98E7AAC12582720057E6FC/$file/2018_LAWS6b_ICRC1.pdf)> (May 06, 2020)

² See High Contracting Parties to the CCW (2019): *Revised Draft Final Report. CCW/MSP/2019/CRP.2/Rev.1*, <[https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/815F8EE33B64DADDC12584B7004CF3A4/\\$file/CCW+MSP+2019+CRP.2+Rev+1.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/815F8EE33B64DADDC12584B7004CF3A4/$file/CCW+MSP+2019+CRP.2+Rev+1.pdf)> (May 06, 2020), p. 10.

³ See iPRAW (August 2018), *Ethical Implications for a Regulation of LAWS*.

⁴ From iPRAW (March 2018), *Focus on the Human-Machine Relation in LAWS* p. 18.

⁵ ICRC (2016), *Views of the International Committee of the Red Cross on Autonomous Weapon Systems*, <<https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system>> (May 06, 2020), p. 1.

3. HUMAN CONTROL: WHAT WILL WE FIND ALONG THE WAY?

The human element can be coined as human control, meaning e.g. situational understanding and the option to intervene by design and during use. In iPRAW's understanding, human control does not necessarily equal direct manipulation. The directness of the means whereby the agent seeks to control some object is only contingently related to the degree of control. Under some circumstances more direct manipulation enables greater control, while in other circumstances the presence of an intervening mechanism might be the better option to reach the desired outcome. The increasing number of assisting systems does not necessarily increase precision, though, as they make the weapon system also more complex and possibly less predictable. A prudent balance of operational needs and situational understanding is crucial.⁶

Control is also not to be understood as a singular event during or at the end of the targeting process, but as a process that requires at least a frequent understanding of the situation. The adequate type and level of human control depends on the individual operational context.

A regulation with a focus on human control would address the process of use which is quite unusual for an arms control regulation, be it humanitarian or otherwise. The design and every step of the weapon system's life cycle would be enablers to allow for human control during attack. To unpack the concept of human control, a common understanding of what constitutes the start of an attack can be useful. In this context, the most relevant point in the mission thread is not defined by the launch or activation, but by the final necessary decision on target selection and engagement by a human. **Weapon systems with autonomous functions potentially move the final human decision to a very early stage of the operation.** With regard to the legal judgments to abide by IHL principles this effect could be challenging for two reasons: First, it can increase the level of abstraction in the target selection process (i.e. class of targets instead of specific target). Second, the environment might change during this extended timespan between targeting decision and engagement, e.g. outdated the initial proportionality assessments. The underlying notion of attack will therefore influence the understanding of the principle of human control in a regulation of autonomous weapon systems. This is because IHL principles like distinction and proportionality are legally required during the planning phase, but, to a certain extent, become a question of feasibility in attack. This would alter the need or necessary level of human control in attack.⁷

4. LEVEL OF GRANULARITY: HOW BIG CAN OUR STEPS BE?

To account for a multitude of battlefield applications a regulation of LAWS might have to remain rather abstract with regard to the type and level of human control. It could, however, encompass all steps in the weapon's life cycle. Those could be, for example, design requirements, training, rules of engagement as well as the explicit call for human control in the targeting process and during the actual attack.

The operational context is crucial for defining the necessary type and level⁸ of human control and multiple factors contribute to the determination of what level of human control is adequate in a given situation. A 'one-size-of-control-fits-all' solution that addresses all concerns raised by the use of autonomous weapon systems will thus most likely not be achievable. Looking at this multitude of relevant factors from the perspective of IHL, the crucial lens would be the **risk for violations of IHL** (due to a lack of situational understanding or timely intervention). One option to address this would be a classification of factors that define the operational context in order to derive consequences for the

⁶ From iPRAW (March 2018), *Focus on the Human-Machine Relation in LAWS*, p. 13.

⁷ From iPRAW (August 2019), *Focus on Human Control*, p. 5-6.

⁸ Not: the need for human control.

implementation of human control.⁹ The shortcoming of such a kind of typology lies in the multitude of combinations of environmental factors, operational requirements, and weapon capabilities it cannot account for. **Instead a regulation would be more useful if it included general approximations to be specified in each case along the lines of existing IHL considerations. Best practices and other dynamic soft law measures could accompany a more abstract regulation.** In addition, table top exercises¹⁰ and other scenario based workshops could facilitate a better understanding of the context before actually fielding weapon systems with autonomous functions.

5. BONUS – VERIFICATION: SHOULD WE GET ICE CREAM ALONG THE WAY?

If States Parties find a consensus on a legally binding regulation and if they want to verify compliance, this verification will be challenging but not necessarily impossible. Different challenges arise from the verification of autonomous functions: (1) they are a qualitative feature, (2) the human role in the target selection and engagement is not visible from the outside, and (3) the software might be altered after inspection. Those challenges would call for a mix of instruments and could depend on the specific type of weapon system and the application of autonomous targeting functions. As with the requirements for human control, there is probably no one-size-fits-all solution to verification. Since other CCW protocols do not provide for verification measures either, a regulation could be useful even without hard verification measures – but it could benefit from the enhanced transparency.¹¹

6. CONCLUSION: WHERE DOES THAT LEAVE US?

The considerations presented above define the *function* of a regulation of LAWS but are not tied to a specific *form*. Overall, iPRAW's recommendation is to operationalize and apply the principle of human control via a more concrete norm. This could be done by a legally or politically binding agreement.¹² The next steps for the GGE 2020/21 could be to (1) **abandon** the distracting debates about a technical definition of LAWS, (2) **focus** on the human element (e.g. human control), and (3) **analyze** the impact of the operational context on the necessary level of human involvement.

The International Panel on the Regulation of Autonomous Weapons (iPRAW) is coordinated by:
Stiftung Wissenschaft und Politik (SWP) – German Institute for International and Security Affairs
Ludwigkirchplatz 3-4, 10719 Berlin, Germany

This project is financially supported by the German Federal Foreign Office.
Find all reports and more information online at www.ipraw.org.

⁹ For an exploratory approach to define relevant criteria see: Marcel Dickow et al. (2015), *First Steps towards a Multidimensional Autonomy Risk Assessment (MARA) in Weapons Systems*, <https://www.swp-berlin.org/fileadmin/contents/products/arbeitspapiere/FG03_WP05_2015_MARA.pdf> (May 06, 2020).

¹⁰ As announced by UNIDIR: Giacomo Persi Paoli (April 2020), *Human Element in the Decisions about the Use of Force*, <https://rethinkingarmscontrol.de/wp-content/uploads/2020/04/working-session-1_%C2%ADPersi-Paoli_presentation%C2%AD.pdf> (May 06, 2020), p. 8.

¹¹ From iPRAW (August 2019), *Verifying LAWS Regulation – Opportunities and Challenges*, p.4.

¹² See iPRAW (December 2018), *Concluding Report*; Campaign to Stop Killer Robots (November 2019), *Key Elements of a Treaty on Fully Autonomous Weapons*, <<https://www.stopkillerrobots.org/wp-content/uploads/2020/03/Key-Elements-of-a-Treaty-on-Fully-Autonomous-Weapons.pdf>> (May 06, 2020).

Icons in Figure 1: Shoe <<https://www.flaticon.com/de/autoren/vitaly-gorbachev>>; Start <<https://www.flaticon.com/de/autoren/freepik>>; Path <<https://www.flaticon.com/de/autoren/smashicons>>; Steps <<https://www.flaticon.com/de/autoren/freepik>>; Ice Cream <<https://www.flaticon.com/de/autoren/good-ware>>; Flag <<https://www.flaticon.com/de/autoren/xnimrodx>> (all retrieved on May 05, 2020).