



Focus on Ethical Implications for a Regulation of LAWS

“Focus on” Report No. 4

August 2018

International Panel on the Regulation of Autonomous Weapons (iPRAW)

coordinated by:

Stiftung Wissenschaft und Politik (SWP) – German Institute for International and Security Affairs
Ludwigkirchplatz 3-4
10719 Berlin, Germany

August 2018

www.ipraw.org
mail@ipraw.org

This project is financially supported by the German Federal Foreign Office.

ABOUT IPRAW

Setting and Objectives: The International Panel on the Regulation of Autonomous Weapons (iPRAW) was founded in March 2017. iPRAW is an independent group of experts from different nation states and scientific backgrounds. The panel will complete its work by the end of 2018.

The mission of iPRAW is to provide an independent source of information and consultation to the Group of Governmental Experts (GGE) within the framework of the United Nations CCW (Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects) during the ongoing process toward a possible future regulation of LAWS (Lethal Autonomous Weapon Systems). This work includes, but is not limited to, the provision of expertise on the military, technical, legal, and ethical basis for practical and achievable policy initiatives regarding LAWS. The mandate of the CCW's open-ended GGE on LAWS will guide the work of iPRAW.

iPRAW seeks to prepare, support, and foster a frank and productive exchange among participants, culminating in perspectives on working definitions and recommendations on a potential regulation of LAWS for the CCW GGE. iPRAW is independent from the GGE and does not function in any official capacity regarding the CCW.

Funding, Organization, and Participants: iPRAW is financially supported by the German Federal Foreign Office. The views and findings of iPRAW do not reflect the official positions of the German government or any other government. Stiftung Wissenschaft und Politik – The German Institute for International and Security Affairs (SWP) and the Johns Hopkins University Applied Physics Laboratory (JHU APL) are jointly organizing the panel. The participants have been selected on the basis of their expertise and the perspectives they bring from a wide range of professional and regional contexts. iPRAW represents the diversity of views on the topic of autonomy in weapon systems. Its members have backgrounds in natural science, engineering, law, ethics, political science, and military operational analysis.

Scope: The panel acknowledges that LAWS may pose a number of considerable legal, ethical and operational challenges and that they might change the security environment in a fundamental way. The full potential of these weapon systems is yet unknown and a mutually agreed definition on LAWS does not exist.

In order to support the CCW GGE process, iPRAW will work on how LAWS should be defined as well as on suggesting possible approaches to regulation. The panel's working sessions will cover the following topics

- state of technology and operations as well as existing definitions of LAWS
- computational systems within the scope of LAWS
- autonomy and human control
- ethics, norms and public perception
- risks and opportunities
- IHL and other fields of law.

iPRAW will publish working documents on each of these topics and will, in addition, publish the panel's final recommendations aimed at informing the CCW process.

Procedure: The participants commit themselves to actively participate in and contribute to all meetings and the scientific dialogue in-between meetings. The panel will meet seven times over the course of two years, starting in March 2017. Each meeting will take two and a half days and will be hosted by SWP in Berlin. Papers with agreed upon recommendations on relevant issues will be drafted and published via the project's website (www.ipraw.org) in-between meetings.

Communication and Publication: The participants discuss under the Chatham House Rule: participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed. As a matter of confidentiality, photographs, video or audio recordings as well as all kinds of activities on social media are not allowed during iPRAW meetings.

The results of the panel discussions will be published. iPRAW members will strive to reach consensus on their recommendations and to reflect that in the panel's publications. Media inquiries with regard to official iPRAW positions should be directed to the steering group. Apart from that, the panel members are free to talk about their personal views on participation and the topics of the panel.

Learn more about iPRAW and its research topics on www.ipraw.org. Please direct your questions and remarks about the project to mail@ipraw.org.

CONTENT

- Executive Summary.....5
- 1 Introduction7
- 2 Human Dignity in International Law9
- 3 Moral Agency and Human Dignity: Ethical Reasons for Human Control11
 - 3.1 Requirements to Safeguard Human Dignity in the Use of LAWS.....11
 - 3.2 Implementation of Human Dignity Through Human Control.....13
- 4 Ethical Positions in the Debate about LAWS.....15
 - 4.1 Skepticism towards the Relevance of Human Dignity16
 - 4.2 Deontological Ethics.....16
 - 4.3 Consequentialist Ethics16
 - 4.4 Consequences for a Regulation of LAWS.....17
- 5 Conclusion19
- 6 Annex21
 - 6.1 Literature.....21
 - 6.2 Members of iPRAW23

EXECUTIVE SUMMARY

This report focuses on the impact of ethical considerations on the principle of human control over weapon systems. As articulated in previous reports, and as supported by parties' statements at the CCW, the notion of human control over weapon systems has emerged (or persisted from prior practice) as a key principle in the debate. This principle is poised to serve as a powerful catalyst for state parties to discuss how human control can be implemented and form the basis for more specific norms.

A critical element in the ongoing debate is the notion of human dignity: what it is, why it is relevant, and when and how it applies in the context of autonomous weapons. Indeed, the debate is at times divided by this issue. One prominent view is that human dignity requires a reasoned and justified basis for taking life, which can only be carried out by a moral agent, i.e., a human. The counter-argument to this line of thinking is that human dignity can be relegated to a lower priority in certain circumstances, such as for the protection of other rights or to preserve the lives of many. There is yet a third group that believes killing must be done lawfully (in compliance with IHL) but can occur without any tie to the notion of dignity.

Two key ethical questions arise from the notion of human dignity and the principle of human control:

1. Does autonomy in weapon systems break the link to moral agency?
2. Can consequences (the expected advantage of a specific application of force in combination with autonomous functions in weapon systems) override issues of human dignity?

With regard to those questions, iPRAW observes that the different ethical perspectives on human dignity in the use of LAWS lead to different requirements for a potential regulation of those weapons. The panel stresses that these conclusions are focused on ethical implications, ignoring other aspects that can and should play into a regulation of LAWS.

Minimum requirements for human dignity in the use of force:

1. Recognize a human being as a human, not just distinguish it from other types of objects and things but as a being with rights that deserve respect;
2. Understand the value of life and the significance of its loss; and
3. Reflect upon the reasons for taking life and reach a rational conclusion that killing is justified in a particular situation.

Depending on the moral position, one would assume or deny that autonomous functions in weapon systems break the link to moral agency. In the first case, it would be necessary to safeguard moral agency through human control, in the latter case one would want to safeguard the ability to use a weapon system lawfully at the current state of technology. In consequence, both positions would require human control in both the design of the system and its use. Inherent in both views is an acknowledgment – tacit or explicit – of the **principle of human control**.

Autonomous functions in the targeting cycle have the potential to diminish moral agency and human control. The principle of precaution therefore calls for techniques to ensure **human control** over the use of these weapon systems.

Therefore, iPRAW recommends that the principle of human control should be **internationally recognized** within the CCW and possibly other documents of international law and be the basis from which requirements can be developed as part of a norm-shaping process. Weapon systems that are developed should be done so with this principle in mind.

iPRAW notes that the critical question emerging from these findings is where in the process the link to human control gets interrupted or broken. An analysis of this issue is not the topic of this report, but is at the heart of the debate and should be the focus of ongoing CCW discussions.

1 INTRODUCTION

The International Panel on the Regulation of Autonomous Weapons (iPRAW) is an independent, interdisciplinary group of scientists working on the issue of lethal autonomous weapon systems (LAWS).¹ It aims to support the current debate within the UN Convention on Certain Conventional Weapons (CCW) with scientifically grounded information and recommendations for a potential regulation of LAWS.

iPRAW publishes interim reports that each focus on different aspects or perspectives on LAWS.² This report focuses on the ethical dimension of the human-machine relation in LAWS. Building on the observations stated in this report, iPRAW makes the following conclusion for aspects of a potential regulation of LAWS:

Depending on the moral position, one would assume or deny that autonomous functions in weapon systems break the link to moral agency. In the first case, it would be necessary to safeguard moral agency through human control, in the latter case one would want to safeguard the ability to use a weapon system lawfully at the current state of technology. In consequence, both positions would require human control in both the design of the system and its use. Inherent in both views is an acknowledgment – tacit or explicit – of the **principle of human control**.

This report explores this link between human control and **human dignity**. This is relevant to the debate on LAWS as it investigates the requirements to lawfully kill according to international humanitarian law (IHL), but also encompasses human rights law (HRL) considerations. Concretely, the examination takes into account the design and use of weapon systems.

To determine how ethical considerations are pertinent to the analysis of existing IHL, HRL and other legal frameworks, we will look at the debate through two different

¹ While focusing on *lethal* autonomous weapon systems, iPRAW does not exclude the potential need to regulate autonomous weapon systems that apply non-lethal force.

² This particular report is based on the fifth and sixth meeting of iPRAW (“LAWS – Risks and Opportunities” in April 2018 and “Ethics, Norms, and Public Perception” in June 2018). The panel thanks **Jürgen Altmann** and **Elvira Rosert** for their valuable contributions to the meetings.

lenses: on the one hand, those that consider that LAWS violate human dignity; on the other hand, those that do not perceive it this way. These two positions can be further examined depending on the value that is ascribed to human dignity: **deontological** approaches that human dignity supersedes other objectives and must not be violated and **consequentialists**, who assume that if human dignity is violated, it can be justified under certain circumstances. The analysis illuminates how different legal tools are applicable to the global debate on LAWS, for example a comprehensive international ban, a political declaration and case-by-case decisions or no necessity for any (international) regulation.

2 HUMAN DIGNITY IN INTERNATIONAL LAW

The Martens Clause has been part of IHL since the first Hague Convention held in 1899. It was later codified in several subsequent international treaties even though there is no consensus on its interpretation, meaning, and impact on the practice of IHL.³ The CCW refers to the Martens Clause in its preamble:

“Confirming their determination that in **cases not covered** by this Convention and its annexed Protocols or by other international agreements, the civilian population and the combatants shall at all times remain under the protection and authority of the principles of international law derived from **established custom, from the principles of humanity and from the dictates of public conscience.**”

The Martens Clause continues to represent the influence of natural law that privileges custom as an underpinning foundation. At the turn of the 20th century, most of international law, including IHL became codified under written treaties, thus positivism became the prevalent perspective in legal thinking. Even though the Martens Clause remains a polarizing idea, it is useful to be considered as it brings other elements beyond codified treaties into global governance-making processes. The Martens Clause possesses three elements that are subject to great controversy:

1. Customary law always applies.
2. Principles of international law and humanity shall always prescribe behavior: In IHL, these have been interpreted as those that prohibit the means and methods of warfare that do not serve the fulfilment of a clear military

³ Preamble, 1907 Hague Convention (IV) respecting the laws and customs of war on land, reprinted in Adam Roberts; Richard Guelf (1989), *Documents on the Laws of War*, 2nd ed., Clarendon Press, Oxford, p. 45; the four 1949 Geneva Conventions for the protection of war victims (GC I: Art. 63; GC II: Art. 62; GC III: Art. 142; GC IV: Art. 158), *op. cit.*, pp. 169-337; 1977 Additional Protocol I, Art. 1(2), *op. cit.*, p. 390, and 1977 Additional Protocol II, Preamble, *op. cit.*, p. 449; 1980 Weapons Convention, Preamble, *op. cit.*, p. 473. See Rupert Ticehurst (1997), “The Martens Clause and the Laws of Armed Conflict”, in: *International Review of the Red Cross*, 317.

advantage.⁴ In other words, the principles represent universal ideas about humane treatment.

3. Public conscience shall be considered and serve as guidance. Today, this is sometimes understood as public opinion as canvassed in polls. While these polls may be useful as sign posts, they, like all polling, sometimes lack representativeness, and they are susceptible to methodological pitfalls.

The principles of humanity, meaning the universal recognition of the need to respect humane treatment of each individual in peace and war, represent the basis for the relevance of the concept of human dignity. It was the horrors of World War II that led to the subsequent codification of the idea of human dignity. The 1945 United Nations Charter advances in its preamble:

“We the peoples of the United Nations determined [...] to reaffirm faith in fundamental human rights, **in the dignity and worth of the human person**, in the equal rights of men and women and of nations large and small, and to establish conditions under which justice and respect for the obligations arising from treaties and other sources of international law can be maintained.”⁵

For the first time, an international treaty on peace and security associates its goals with the need to uphold dignity. Equal rights became the mechanism for the attainment of dignity.⁶

Other human rights treaties, known as the ‘Magna Carta’ of human rights, further enshrined the principle of human dignity in the post-War world order, making it clear that this inalienable right is the foundation for world peace: the 1948 Universal Declaration of Human Rights (UDHR), the 1966 International Covenant on Civil and Political Rights (ICCPR) and the 1966 International Covenant on Economic, Social and Cultural Rights (ICESCR). The UDHR has sole hortatory force, however, it is widely recognized as customary law and therefore binds all states. Both ICCPR and ICESCR are legally binding upon states and are fairly universal due to the number of High Contracting Parties that have ratified them (116 and 168 countries respectively). They contain an identical formulation in their preambles:

“Considering that, in accordance with the principles proclaimed in the Charter of the United Nations, recognition of the **inherent dignity and of the equal and inalienable rights** of all members of the human family is the foundation of freedom, justice and peace in the world **recognizing that these rights derive from the inherent dignity of the human person.**”

With this ‘Magna Carta’, human dignity finds a prominent place in the architecture of the current global order. Given the elements of customary law force in the UDHR and the binding universal force of the ICCPR and the ICESCR, it is unquestionably hard to omit human dignity from the current discussions on LAWS.

⁴ Edward Kwakwa (1992), *The International Law of Armed Conflict: Personal and Material Fields of Application*, Kluwer Academic, Dordrecht, p. 36.

⁵ Preamble, United Nations Charter, emphasis added.

⁶ Antônio Augusto Cançado Trindade (2013), *International Law for Humankind Towards a New Jus Gentium*, Martinus Nijhoff, 2nd ed.

3 MORAL AGENCY AND HUMAN DIGNITY: ETHICAL REASONS FOR HUMAN CONTROL

Previous iPRAW reports discussed numerous reasons why human control is needed in the use of force: for operational, legal, and ethical considerations. This report focuses on the ethical dimension, namely the link between human control in the use of force and the human dignity of the targeted person.

3.1 REQUIREMENTS TO SAFEGUARD HUMAN DIGNITY IN THE USE OF LAWS

Human dignity is a concept that often appears in discussions of human rights, but is rarely considered in detail. It is sometimes described as a sort of property that attaches to a person, and which can be stripped from them. But this metaphor only captures a part of what constitutes human dignity, and does little to help us understand its importance in armed conflict. Indeed, it is often said that there can be little dignity in war, or in dying in war, and certainly the manner in which many people are killed in wars by flames, explosions, shrapnel, bullets, etc., can be undignified. But it is not the physical means of death which determines whether a death is dignified, any more than the manner of death justifies whether the death is lawful or moral. The human right to dignity, much like the human right to life, inheres not in a property and its loss, or the physical-causal means of its loss, but in the reasons for that right being violated or overridden.

The right to life has been described as the fundamental human right, and the right which all others depend upon.⁷ But the right to human dignity may be more fundamental. Indeed, dignity persists even after death – the right to a proper funeral, and the prohibition of desecrating of dead bodies are clear examples of the human right to dignity extending beyond the loss of life. At its core, dignity is the right to be respected as a human being, with the full rights of a human being. It is the right to be

⁷ See Christoph Heyns (2017), *Autonomous Weapons in armed conflict and the right to a dignified life*, in: *An African perspective. South African Journal on Human Rights*, Vol. 33 (1), pp. 46-71.

treated as an equal being and an end in oneself not a means to some other end – accordingly the classic violations of human dignity are slavery and torture. While the practices of slavery and torture are clearly very bad for those who are enslaved and tortured, from a consequentialist moral perspective, there may be arguments that some people benefit greatly, or many people might benefit, from such practices in certain instances. But from the perspective of a right to human dignity, the very existence of slavery or torture diminishes the value of human dignity for all people, including the slavers and torturers. By adopting an action, and any reasons to justify these practices, one makes it easier for all people to become slaves or be tortured, and diminishes the collective dignity of all people.

When it comes to the question of autonomous weapons, it may seem easy to argue that “it does not matter how one is killed in war.” **But clearly it does, from both a moral and legal perspective.** What, then, is required to ensure that killing in war does not violate human dignity? As Heyns has argued, it must not be arbitrary, which is also to say that it must be justified and must therefore have reasons to justify it.⁸ Moreover, the decision to use lethal force must be taken with justification and reason. The question for autonomous weapons is whether a calculated machine decision, i.e., computations based on sensor data, can meet these criteria.

Some have argued that computational systems are neither moral nor legal agents and thus cannot legitimately determine when it is appropriate to take human life.⁹ While computational systems may be able to accurately and reliably apply a computational rule to a set of data, in so doing they are not respecting human dignity. In order to make a moral judgment to take a life, while respecting human dignity, it is **minimally required** that a **moral agent** can

1. **Recognize** a human being as a human, not just distinguish it from other types of objects and things but as a being with rights that deserve respect;
2. Understand the **value of life** and the significance of its loss; and
3. **Reflect** upon the reasons for taking life and reach a rational conclusion that killing is justified in a particular situation.¹⁰

Currently, only humans are capable of meeting these criteria, which is why it is morally and legally required that humans take responsibility for decisions to use lethal force, and should continue to be in the future. Some argue that in conflict, humans have taken human lives without respect for human dignity, therefore it is a standard we ought not to apply to autonomous weapons systems. However, this argument is inadequate unless the goal is to regress to a lower standard for justification in killing. Others argue that if an autonomous weapon system can be used with greater compliance for standards of IHL than a human, the fact that it cannot reflect on the killing as a moral agent (and therefore cannot recognize human dignity) should not be

⁸ See Heyns (2017), fn. 7.

⁹ See Peter Asaro (2012), “On Banning Autonomous Weapon Systems: human rights, automation, and the dehumanization of lethal decision-making”, in: *International Review of the Red Cross*, Vol. 94 (886), pp. 687-709.

¹⁰ See Peter Asaro (2016), “Jus Nascendi, Robotic Weapons and the Martens Clause”, in: Ryan Calo, Michael Froomkin and Ian Kerr (eds.), *Robot Law*, Edward Elgar Publishing, pp. 367-386.

a reason to forbid its use. These differing ethical perspectives lead to different evaluations of the value of human dignity, which impacts potential regulatory requirements (see below in chapter 4).

3.2 IMPLEMENTATION OF HUMAN DIGNITY THROUGH HUMAN CONTROL

Distinguishing a “target” in a field of data is not recognizing a human person as someone with rights. Nor is discriminating between combatants and non-combatants sufficient for recognizing someone as a human being with rights to dignity and life. When it comes to making proportionality decisions, the value of human life is not quantifiable in any deep sense. As human beings, who have experienced loss and are ourselves mortal, we have access to the value of human life. And finally, machines are not capable of deciding questions of military necessity – whether a state of war exists, whether a person in a given situational context can be justifiably killed, or what reasons justify the necessity of destroying a military objective.

Distinguishing a “target” in a field of data is not recognizing a human person as someone with rights.

In terms of the discussions of the CCW, the focus on meaningful human control of weapons, or appropriate levels of human judgment, or other formulations of this general concept of human control, all speak directly to the central concept of human dignity. **Requiring humans to make targeting decisions**, or the ultimate decision to use lethal force, **is to recognize humans as the only appropriate and legitimate agents for making such decisions**: “control” and “judgment” speak directly to the need to ensure reasoned deliberation over the necessity and justification for the use of force in every instance and with the specific context and situation in mind. The focus on humans is based on the recognition that humans are the only appropriate sorts of moral and legal agents for making such judgments. In addition, “meaningful” and “appropriate” underline that preprogrammed decision rules, however well designed and intentioned, cannot determine their own appropriateness in a given situation, nor can they be held accountable or take moral responsibility (though humans may fail to do so).

In the previous report, iPROW operationalized the term human control:¹¹ The design of weapon systems with autonomous functionalities must enable the operator to understand the operational situation to allow for informed decisions over the use of force. The necessary monitoring of the environment and the system includes system diagnostics, internal and external sensors for system and environmental monitoring as well as methods for communicating that information. In addition, the ability for humans to actively intervene prior to the ultimate decision over the use of force must be given and should be a default feature. The need for situational understanding and intervention is not limited to one single weapon system, but should also refer to systems of multiple robots executing a mission, which is quite likely how these capabilities will be developed and fielded.

As technology evolves it is still disputed whether an obligation exists to not only use the most advanced weapon systems available (e.g. to improve discrimination

¹¹ See iPROW (March 2018), *Focus on the Human-Machine Relation in LAWS*, <https://www.ipraw.org/wp-content/uploads/2018/03/2018-03-29_iPROW_Focus-On-Report-3.pdf>.

between combatants and civilians), but to implement all possible options to ensure human control over the use of these weapon systems.

In its sixth meeting in June 2018, iPRAW discussed whether a mortar shell with an electronic target selection function based on an advanced target recognition library calls for more human control than a conventional GPS-guided round when applied beyond the line of sight. We focused on the question of whether the operator loses human control and moral agency in the former case (automated target selection functionality) as compared to the latter due to the fact that a hypothetical possibility of a such weapon system to communicate back detected target options (situational understanding) exists. This would leave the opportunity for the operator to make legal and moral judgments based on a more detailed situational understanding and to override the selection process (intervention) if necessary.

iPRAW recognizes that autonomous functions in the targeting cycle have the potential to diminish moral agency and human control and that the principal of precaution therefore calls for techniques to ensure human control over the use of these weapon systems. The level of such human control remains to be determined and may vary with the operational context.¹²

¹² For more details see iPRAW (August 2018 – forthcoming), *Add-on to “Focus on the Human-Machine Relation in LAWS”*.

4 ETHICAL POSITIONS IN THE DEBATE ABOUT LAWS

In the international and CCW debate various topics have been raised to address the ethical challenges of autonomous functions in weapon systems. A first strand of this debate focuses on challenges related to international security and stability, potential arms dynamics and races, the threshold for the use of force in and to an armed conflict, proliferation of technologies relevant to LAWS and future weapon system with autonomous functions, and the question if LAWS are inherently unethical because they cannot be applied in adherence to IHL. The second strand circles around the question whether a possible transfer of control from the operator/commander to a weapon system may be unethical in general or in specific cases. Through the lens of ‘human control’ this means: **How much human control is necessary to safeguard moral agency in the use of force?**

This iPRAW report focuses on the second strand of the debate. From our observation of the discussion at the CCW, we found, amongst others, two important ethical questions raised:

1. Does autonomy in weapon systems break the link to moral agency?
2. Can consequences (the expected advantage of a specific application of force in combination with autonomous functions in weapon systems) override issues of human dignity?

The first question is linked to human dignity as discussed in chapter 3. Weapon systems can lack human control for various reasons, for example the loss of communication, technical failure, negligence or misconduct. In such cases the loss of human control in the use of force does not automatically violate the dignity of the targeted person. Those who argue that it does, refer to circumstances in which weapon systems are deliberately designed and used to apply (lethal) force without sufficient human control. The second question addresses the importance of human dignity in the application of (lethal) force. While some would say that certain ends can justify the violation of dignity, others cannot accept such a violation in any case. **Answering these two questions may foster a better understanding of individual positions along the lines of two major ethical theories to be described in the following and offers guidance for the type of regulation appropriate for LAWS.**

4.1 SKEPTICISM TOWARDS THE RELEVANCE OF HUMAN DIGNITY

Some actors argue that LAWS will not violate human dignity or that human dignity is not relevant to the debate or to an analysis under IHL. This argument does not see human dignity as a central part of the debate either because it does not really exist, or because it is reducible to the formal respect of other rights (e.g. life). For some, it is an ephemeral quality that can never be adequately understood and therefore need not be respected. In the case of some states, actions in conflict are interpreted strictly through the lens of IHL and whether it is violated, absent a tie between IHL and human dignity or IHL and HRL. While it is notoriously difficult to argue against skepticism, the fact that human dignity has been articulated in various world philosophies and religions has been integral to legal theory and practice, provides strong support that human dignity has both a defined structure and is broadly recognized as integral to law and morality.

4.2 DEONTOLOGICAL ETHICS

Chapter 3 shows how the process of decision making influences the dignity of the target. It does not say anything about the value attached to this quality as that depends on the ethical perspective – deontological and consequentialist ethics come to very different conclusions on that regard.

Deontological ethics address the morality of human choices. Those cannot be justified by beneficial outcomes, but only assessed by their compliance to a moral norm. Following this logic, human dignity is an inviolable right. While documents of international HRL do not state the inviolability of human dignity explicitly, a deontological interpretation would imply just that. One national example would be the German constitution which explicitly grants human dignity as an absolute and untouchable right and a supreme value that supersedes any other claim.¹³

4.3 CONSEQUENTIALIST ETHICS

One of the most influential approaches standing in opposition to deontological ethics is known as ‘consequentialism’. As the name suggests, consequentialists believe that the best way to establish if some action or state of affairs is ethically sound is to weigh up the consequences of that action: if those consequences are, on balance, good, then the action is ethical. The dominant consequentialist theory, utilitarianism, makes that assessment on the basis of the overall balance of ‘utility’ (or in some variants ‘happiness’) for those affected by the action. Accordingly, utilitarians are generally willing to accept that an action can be ethically sound even if it results in harm or other negative consequences for some of those affected by that action, so long as more utility than disutility results therefrom.¹⁴

Consequentialist approaches, e.g. utilitarianism, do not necessarily deny the violation of human dignity through a loss of human control – though principle of human dignity

¹³ See also: Frank Sauer et al. (2018), *Autonomy in Weapons Systems. The Military Application of Artificial Intelligence as a Litmus Test for Germany’s New Foreign and Security Policy*, Heinrich Böll Stiftung, Berlin, pp. 32-33.

¹⁴ See Deane-Peter Baker (2015), “Consequentialist Ethics”, in: Deane-Peter Baker (ed.), *Key Concepts of Military Ethics*, 2015.

and human control still applies. They allow, however, other claims to supersede this violation of rights. Therefore, benefits enabled by autonomous functions or the lives of many can be reasons to choose less controlled settings. For example, machines with autonomous functions might be better at discrimination, proportionality or even military necessity calculations. However, this would depend on a consequentialist analysis, whereby “better performance” consists of making automated choices more accurately or reliably than a human. But this overlooks the reasons and justifications for the decisions, as opposed to the consequences, or choices. When it comes to human dignity, what is crucial is both the manner in which the decision is made and the legitimacy of who is making the decision, not the final outcome of the decision.

1. **Does autonomy in weapon systems break the link to moral agency (and therefore constitutes a violation of human dignity)?**
 - a. Skepticism: no.
 - b. Deontological ethics: yes.
 - c. Consequentialism: possibly yes.
2. **Can consequences override issues of human dignity?**
 - a. Skepticism: irrelevant.
 - b. Deontological ethics: no.
 - c. Consequentialism: yes.

4.4 CONSEQUENCES FOR A REGULATION OF LAWS

The previous discussion stated the ethical but also legal relevance of human dignity, moral agency, and human control for evaluating the compliance of weapon systems with international law. For this report’s purpose, the concept of human control serves as an indicator to illustrate in which situations moral agency might be lost in the human-machine interaction and the use of force on the one hand. On the other hand, the principle of human control might serve as a scheme to investigate the minimum legal requirements to lawfully applying force according to IHL but also encompassing HRL considerations.

If one does not agree that autonomous functions in weapon systems can lead to a violation of human dignity, there might be no (ethical) reason for any regulation of this type of weapon.

Provided that the use of LAWS violates the dignity of the targeted person, a **deontological** point of view would assume that this type of weapon must be forbidden on an international level. This would include a norm on human control to safeguard the inviolable right/principle to human dignity.

From a **consequentialist** perspective, the violation can be accepted in certain – but not all – cases in which the benefit of the outcome compensates for this. Here, an international agreement concerning the principle of human control would be necessary to establish how human dignity can be safeguarded. On this basis, a case-by-case decision on a national level would assess if a violation of human dignity can be justified: Weapon systems that lack adequate human control to meet ethical and/or legal thresholds should be subject to more rigorous evaluation. The mechanism for conducting such an evaluation should be determined through discussion among the States in the course of developing best practices.

The described consequences for the regulation of LAWS are limited to the ethical perspective. If other angles are applied *in addition to the ethical evaluation, different outcomes might be more adequate*. Apart from legal concerns regarding accountability and the principles of IHL, those other angles could be the risks to international security, the high potential for proliferation or the challenges to a verification of any regulation of LAWS.¹⁵ We will take all those perspectives into consideration in our final report in late 2018 to provide comprehensive recommendations.

¹⁵ For details see e.g. Jürgen Altmann; Frank Sauer (2017), “Autonomous Weapon Systems and Strategic Stability”, in: *Survival* 59 (5), pp. 117-142; Mark Gubrud; Jürgen Altmann (2013), *Compliance Measures for an Autonomous Weapons Convention*, ICRAC Working Paper No. 2, <http://icrac.net/wp-content/uploads/2013/05/Gubrud-Altman_Compliance-Measures-AWC_ICRAC-WP2.pdf>.

5 CONCLUSION

The fact that human dignity has been articulated in various world philosophies and religions, has been integral to legal theory and practice, and has been codified in the constitutions of countries, including the Preamble of the United Nations Universal Declaration of Human Rights and Article 1 of the German constitution, provides strong support that human dignity has both a defined structure and is broadly recognized as integral to law and morality. Human dignity is an essential piece of the global human rights architecture. Concomitantly, universal principles, such as precaution in war, and the distinction between civilians and combatants give rise to new norms of behavior that guide the conduct of states at war and peace. These have been codified under IHL. In the current debate on LAWS, the principle of human control is firmly rising as a prescriptive blueprint for action. The overwhelming majority of states seem to agree that human control is the principle that should constitute the basis of new norms of behavior regarding LAWS.

Depending on the moral position, one would assume or deny that autonomous functions in weapon systems break the link to moral agency. In the first case, it would be necessary to safeguard moral agency through human control, in the latter case one would want to safeguard the ability to use a weapon system lawfully at the current state of technology. In consequence, both positions would require human control in both the design of the system and its use. Inherent in both views is an acknowledgment – tacit or explicit – of the principle of human control.

Minimum requirements for human dignity in the use of force:

1. Recognize a human being as a human, not just distinguish it from other types of objects and things but as a being with rights that deserve respect;
2. Understand the value of life and the significance of its loss; and
3. Reflect upon the reasons for taking life and reach a rational conclusion that killing is justified in a particular situation.

Therefore, **iPRAW recommends that the principle of human control should be internationally recognized within the CCW and possibly other documents of international law**¹⁶ and be the basis from which requirements can be developed as part of a norm-shaping process. Weapon systems that are developed should be done so with this principle in mind. iPRAW notes that the critical question emerging from these findings concerns where in the process the link to human control is interrupted or broken. An analysis of this issue is not the topic of this report, but is at the heart of the debate and should be the focus of ongoing CCW discussions.

¹⁶ For suggestions on that see Elvira Rosert (2017), *How to Regulate Autonomous Weapons. Steps to Codify Meaningful Humanitarian Control as a Principle of International Humanitarian Law*, PRIF Spotlight 6/2017, Frankfurt/M.

6 ANNEX

6.1 LITERATURE

- Altmann, Jürgen; Sauer, Frank (2017), “Autonomous Weapon Systems and Strategic Stability”, in: *Survival* 59 (5), pp. 117-142.
- Asaro, Peter (2012), “On Banning Autonomous Weapon Systems: human rights, automation, and the dehumanization of lethal decision-making”, in: *International Review of the Red Cross*, Vol. 94 (886), pp. 687-709.
- Asaro, Peter (2016), “Jus Nascendi, Robotic Weapons and the Martens Clause”, in: Ryan Calo, Michael Froomkin, Ian Kerr (eds.) *Robot Law*, Edward Elgar Publishing, pp. 367-386.
- Baker, Deane-Peter (2015), “Consequentialist Ethics”, in: Deane-Peter Baker (ed.), *Key Concepts of Military Ethics*.
- Cançado Trindade, Antônio Augusto (2013), *International Law for Humankind Towards a New Jus Gentium*, Martinus Nijhoff, 2nd ed.
- Garcia, Denise (2018), “Lethal Artificial Intelligence and Change: The Future of International Peace and Security”, in: *International Studies Review* 20, pp. 334-341.
- Gubrud, Mark; Altmann, Jürgen (2013), *Compliance Measures for an Autonomous Weapons Convention*, ICRAC Working Paper No. 2, <http://icrac.net/wp-content/uploads/2013/05/Gubrud-Altman_Compliance-Measures-AWC_ICRAC-WP2.pdf>.
- Heyns, Christoph (2017), “Autonomous Weapons in armed conflict and the right to a dignified life: An African perspective”, in: *South African Journal on Human Rights*, Vol. 33 (1), pp. 46-71.
- Human Rights Watch, IHRC (2016), *Killer Robots and the Concept of Meaningful Human Control Memorandum to Convention on Conventional Weapons (CCW) Delegates*.

- iPRAW (March 2018), *Focus on the Human-Machine Relation in LAWS*, <https://www.ipraw.org/wp-content/uploads/2018/03/2018-03-29_iPRAW_Focus-On-Report-3.pdf>.
- iPRAW (August 2018 – forthcoming), *Add-on to “Focus on the Human-Machine Relation in LAWS”*.
- Kwakwa, Edward (1992), *The International Law of Armed Conflict: Personal and Material Fields of Application*, Kluwer Academic, Dordrecht.
- Roberts, Adam; Guelf, Richard (1989), *Documents on the Laws of War*, 2nd ed., Clarendon Press, Oxford.
- Rosert, Elvira (2017), *How to Regulate Autonomous Weapons. Steps to Codify Meaningful Humanitarian Control as a Principle of International Humanitarian Law*, PRIF Spotlight 6/2017, Frankfurt/M.
- Sauer, Frank et al. (2018), *Autonomy in Weapons Systems. The Military Application of Artificial Intelligence as a Litmus Test for Germany’s New Foreign and Security Policy*, Heinrich Böll Stiftung, Berlin.
- Ticehurst, Rupert (1997), “The Martens Clause and the Laws of Armed Conflict”, in: *International Review of the Red Cross*, 317.

6.2 MEMBERS OF IPRAW

Liran Antebi*
Research Fellow
Institute for National Security Studies
Tel Aviv, Israel

Peter Asaro
Professor
The New School
New York, USA

Deane-Peter Baker
Senior Lecturer
University of New South Wales
Canberra, Australia

Vincent Boulanin* **
Researcher
Stockholm International Peace Research Institute
Stockholm, Sweden

Thompson Chengeta*
Fellow
South African Research Chair in International Law, University of Johannesburg
Johannesburg, South Africa

Anja Dahlmann
Researcher
German Institute for International and Security Affairs
Berlin, Germany

Marcel Dickow
Head of Research Division
German Institute for International and Security Affairs
Berlin, Germany

Denise Garcia
Professor
Northeastern University
Boston, USA

Robin Geiß
Professor
University of Glasgow
Berlin, Germany

Erin Hahn*
Researcher
Johns Hopkins University Applied Physics Laboratory
Washington D.C., USA

Vadim Kozyulin
Researcher
PIR Center for Policy Studies
Moscow, Russia

Ian MacLeod
Researcher
Johns Hopkins University Applied Physics Laboratory
Washington D.C., USA

AJung Moon*
Senior Advisor of Research and Industry Outreach
UN Secretary-General's High-level Panel on Digital Cooperation
Canada/ USA

Shashank Reddy**
Researcher
Carnegie India
New Delhi, India

Heigo Sato
Professor
Takushoku University
Tokyo, Japan

Frank Sauer
Researcher
Universität der Bundeswehr
Munich, Germany

David Hyunchul Shim* **
Professor
Korea Advanced Institute of Science and Technology
Daejeon, South Korea

Lena Strauß
Research Assistant
German Institute for International and Security Affairs
Berlin, Germany

The asterisk indicates those members, who did not participate in the fifth meeting of iPRAW ("LAWS – Risks and Opportunities") in April 2018 (*) or the sixth meeting ("Ethics, Norms, and Public Perception") in June 2018 (**).

Former members: **Dong Lin**, Researcher at the National University of Defense Technology in Changsha, China and **Kelvin Wong**, Researcher at IHS Janes in Singapore.

International Panel on the Regulation of Autonomous Weapons (iPRAW)

coordinated by:

Stiftung Wissenschaft und Politik (SWP) – German Institute for International and Security Affairs
Ludwigkirchplatz 3-4
10719 Berlin, Germany

August 2018

www.ipraw.org
mail@ipraw.org

This project is financially supported by the German Federal Foreign Office.

