

U.S. Commentaries on the Guiding Principles

This paper provides U.S. commentaries on the eleven Guiding Principles adopted by the Group of Governmental Experts (GGE) on emerging technologies in the area of lethal autonomous weapons systems (LAWS) and endorsed by High Contracting Parties to the Convention on Certain Conventional Weapons (CCW).¹

The guiding principles serve as a foundation for the GGE's future work and can also guide States in the responsible development and use of emerging technologies in the area of LAWS. The guiding principles are a cohesive framework with each principle reinforcing others.²

(a) International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems.

This guiding principle is a foundational one for the GGE's work. Understanding how IHL applies to the potential development and use of lethal autonomous weapons systems is critical for effectively implementing the other guiding principles, including guiding principles (c), (d), (e), and (h).

The GGE should build on its successful 2019 work on IHL by further clarifying IHL requirements applicable to the use of emerging technologies in the area of LAWS. What IHL requires often depends on how weapons or tools are being used. Thus, clarifying IHL requirements can be done by considering how militaries have used autonomous functions in weapon systems.

In its 2019 working paper, the United States described three general scenarios for the use of autonomous functions in weapon systems: 1) using autonomous functions to effectuate more accurately and reliably a commander or operator's intent to strike a specific target or target group; 2) using autonomous functions to inform a commander or operator's decision-making about what targets he or she intends to strike; 3) using autonomous functions to select and engage specific targets that the commander or operator did not know of when he or she activated the weapon system.

¹ The United States reaffirms its support for the GGE's relevant conclusions in previous years' reports and the views previously expressed in U.S. working papers to the GGE, which may elaborate on the points in this submission. See *Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems*, March 28, 2019, [CCW/GGE.1/2019/WP.5](#); *Human-Machine Interaction in the Development, Deployment and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*, August 28, 2019, [CCW/GGE.2/2018/WP.4](#); *Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapon Systems*, March 28, 2018, [CCW/GGE.1/2018/WP.4](#); *Characteristics of Lethal Autonomous Weapons Systems*, November 10, 2017, [CCW/GGE.1/2017/WP.7](#); *Autonomy in Weapon Systems*, November 10, 2017, [CCW/GGE.1/2017/WP.6](#).

² For example, further work on elaborating how international humanitarian law (IHL) applies to the potential development and use of lethal autonomous weapons systems (principle (a)), can assist States in conducting legal reviews of new weapons (principle (e)), and such legal reviews also provide an opportunity to consider good practices in human-machine interaction to ensure compliance with IHL (principle (c)), as well as risk assessments and mitigation measures (principle (g)).

The United States proposes the following draft conclusions for the GGE’s consideration.

1. Consistent with IHL, autonomous functions may be used to effectuate more accurately and reliably a commander or operator’s intent to strike a specific target or target group.
 - a. The addition of autonomous functions, such as the automation of target selection and engagement, to weapon systems can make weapons more precise and accurate in striking military objectives by allowing weapons or munitions to “home in” on targets selected by a human operator.
 - b. If the addition of autonomous functions to a weapon system makes it inherently indiscriminate, *i.e.*, incapable of being used consistent with the principles of distinction and proportionality, then any use of that weapon system would be unlawful.
 - c. The addition of autonomous functions to a weapon system can strengthen the implementation of IHL when these functions can be used to reduce the likelihood of harm to civilians and civilian objects.
2. Consistent with IHL, emerging technologies in the area of LAWS may be used to inform decision-making.
 - a. When making a decision governed by IHL, commanders and other decision-makers must make a good faith assessment of the information that is available to them at the time.
 - b. IHL generally does not prohibit commanders and other decision-makers from using tools to aid decision-making in armed conflict. Whether the use of a tool to aid decision-making in armed conflict is consistent with IHL may depend on the nature of the tool, the circumstances of its use, as well the applicable rules and duties under IHL.
 - c. Reliance on a machine assessment to consider a target to be a military objective must be compatible with the decision-maker’s duty under IHL to exercise due regard to reduce the risk of harm to civilians and civilian objects. Such compatibility depends on the relevant circumstances ruling at the time, including:
 - i. how accurately and consistently the machine performs in not mischaracterizing civilian objects as military objectives (*i.e.*, false positives);
 - ii. the decision-maker giving the machine assessment appropriate weight relative to other information relevant to whether the target was a military objective (*e.g.*, operational context, intelligence reporting of the threat identified by the system); and

- iii. the urgency to make a decision (*e.g.*, whether the decision occurred in combat operations or in the face of an imminent threat of an attack, or whether more time could be taken before making a decision).
3. Consistent with IHL, weapons systems that autonomously select and engage targets may be used where the human operator has not expressly intended to strike a specific target or group of targets when activating the weapon system.
- a. The commander or operator could act consistently with the principle of distinction by:
 - i. Using weapon systems that autonomously select and engage targets in areas that constitute military objectives; or
 - ii. Using weapon systems that autonomously select and engage targets with the intent of making potential targets constituting military objectives (*e.g.*, potential incoming projectiles in an active protection system) the object of attack, provided that the weapon systems perform with sufficient reliability (*e.g.*, an active protection system consistently selecting and engaging incoming projectiles) to ensure that force is directed against such targets.
 - b. The expected loss of civilian life, injury to civilians, and damage to civilian objects incidental to the employment of weapons systems that autonomously select and engage targets must not be excessive in relation to the concrete and direct military advantage expected to be gained.
 - i. The expected loss of civilian life, injury to civilians and damage to civilian objects is to be informed by all available and relevant information, including information about: (i) the presence of civilians or civilian objects within the area and during the time when the weapon system is expected to be operating; (ii) the performance of the weapon's autonomous functions in selecting and engaging military objectives; (iii) the risks posed to civilians and civilian objects when the weapon engages military objectives; (iv) the incidence of military objectives that could be engaged by the weapon system in the operational area; and (v) the effectiveness of any precautions taken to reduce the risk of harm to civilians and civilian objects.
 - ii. The concrete and direct military advantage expected to be gained is to be informed by all available and relevant information, which may include information about how the employment of the weapon system: (i) threatens military objectives belonging to the adversary; (ii) contributes to the security of the operating forces; (iii) diverts enemy resources and

attention; (iv) shapes or diverts the movement of enemy forces; and (v) supports military strategies and operational plans.

- c. Feasible precautions must be taken in use of weapon systems that autonomously select and engage targets to reduce the expected harm to civilians and civilian objects. Such precautions may include:
 - i. Warnings (*e.g.*, to potential civilian air traffic or notices to mariners);
 - ii. Monitoring the operation of the weapon system; and
 - iii. Activation or employment of self-destruct, self-deactivation, or self-neutralization mechanisms (*e.g.*, use of rounds that self-destruct in flight or torpedoes that sink to the bottom if they miss their targets).

(b) Human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines. This should be considered across the entire life cycle of the weapons system.

This guiding principle reflects the fundamental importance of human responsibility in using machines. The GGE should elaborate on guiding principle (b) by addressing how well-established international legal principles of State and individual responsibility apply to States and persons who use weapon systems with autonomous functions. Such work could inform practical measures to promote accountability for such decisions, addressed under guiding principle (d).

The United States proposes the following conclusions for the GGE's consideration.

1. Under principles of State responsibility, every internationally wrongful act of a State, including such acts involving the use of emerging technologies in the area of LAWS, entails the international responsibility of that State.³
2. A State remains responsible for all acts committed by persons forming part of its armed forces, including any such use of emerging technologies in the area of LAWS, in accordance with applicable international law.
3. An individual, including a designer, developer, an official authorizing acquisition or deployment, a commander, or a system operator, is responsible for his or her decisions governed by IHL with regard to emerging technologies in the area of LAWS.
4. Under applicable international and domestic law, an individual remains responsible for his or her conduct in violation of IHL, including any such violations involving emerging technologies in the area of LAWS. The use of machines, including emerging

³ Adapted from Article 1 of the International Law Commission's Draft articles on Responsibility of States for Internationally Wrongful Acts.

technologies in the area of LAWS, does not provide a basis for excluding legal responsibility.

5. The responsibilities of any particular individual in implementing a State or a party to a conflict's obligations under IHL may depend on that person's role in the organization or military operations, including whether that individual has the authority to make the decisions and judgments necessary to the performance of that duty under IHL.
6. Under IHL, a decision, including decisions involving emerging technologies in the area of LAWS, must be judged based on the information available to the decision-maker at the time and not on the basis of information that subsequently becomes available.
7. Unintended harm to civilians and other persons protected by IHL from accidents or equipment malfunctions, including those involving emerging technologies in the area of LAWS, is not a violation of IHL as such.
8. States and parties to a conflict have affirmative obligations with respect to the protection of civilians and other classes of persons under IHL, which continue to apply when emerging technologies in the area of LAWS are used. These obligations are to be assessed in light of the general practice of States, including common standards of the military profession in conducting operations.

(c) Human-machine interaction, which may take various forms and be implemented at various stages of the life cycle of a weapon, should ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with applicable international law, in particular IHL. In determining the quality and extent of human-machine interaction, a range of factors should be considered including the operational context, and the characteristics and capabilities of the weapons system as a whole.

This principle recognizes that human-machine interaction should ensure IHL compliance and as well as the need to consider human-machine interaction comprehensively, across the life cycle of the weapon system. The GGE should elaborate on good practices in human-machine interaction that can strengthen compliance with IHL.

The United States proposes the following conclusions on human-machine interaction for the GGE's consideration.⁴

1. Weapons systems based on emerging technologies in the area of LAWS should effectuate the intent of commanders and operators to comply with IHL, in particular, by avoiding

⁴ These and other U.S. practices to ensure that the use of machines helps effectuate human intent are discussed in greater detail in the U.S. working paper, *Human-Machine Interaction in the Development, Deployment and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems*. These practices are reflected in U.S. Department of Defense Directive 3000.09, *Autonomy in Weapon Systems*, November 21, 2012 (updated May 8, 2017), available at www.esd.whs.mil.

unintended engagements and minimizing harm to civilians and civilian objects. This can be effectuated through the following measures:

- a. Weapons systems based on emerging technologies in the area of LAWS should be engineered to perform as anticipated. This should include verification and validation and testing and evaluation before fielding systems.
- b. 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.
- c. User interfaces for weapons systems based on emerging technologies in the area of LAWS should be clear in order for operators to make informed and appropriate decisions in engaging targets. In particular, interface between people and machines for autonomous and semi-autonomous weapon systems should: (i) be readily understandable to trained operators; (ii) provide traceable feedback on system status; and (ii) provide clear procedures for trained operators to activate and deactivate system functions.

(d) Accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law, including through the operation of such systems within a responsible chain of human command and control.

This guiding principle recognizes that State and individual responsibility must be ensured through the effective implementation of accountability measures, including the military chain of command. Such implementation is an essential part of the responsible use of emerging technologies in the area of LAWS. The GGE should elaborate on guiding principle (d) by articulating good practices to help ensure accountability.

The United States proposes the following conclusions on human-machine interaction for the GGE's consideration.

1. The following general practices help ensure accountability in military operations, including operations involving the use of emerging technologies in the area of LAWS:
 - a. Conducting operations under a clear operational chain of command.
 - b. Subjecting members of the armed forces to a system of military law and discipline.
 - c. Establishing and using procedures for the reporting of incidents involving potential violations.

- d. Conducting assessments, investigations, or other reviews of incidents involving potential violations.
 - e. Disciplinary and punitive measures as appropriate.
2. The following practices with respect to the use of weapons systems, including those based on emerging technologies in the area of LAWS, can promote accountability:
- a. Rigorous testing of and training on the weapon system, so commanders and operators understand the likely effects of employing the weapon system.
 - b. Establishing procedure and doctrine applicable to the use of the weapon system, which provide standards for commanders and operators on responsible use and under which they can be held accountable under the State's domestic law.
 - c. Using the weapon system in accordance with established training, doctrine, and procedures and refraining from unauthorized uses or modifications of the weapon system.

(e) In accordance with States' obligations under international law, in the study, development, acquisition, or adoption of a new weapon, means or method of warfare, determination must be made whether its employment would, in some or all circumstances, be prohibited by international law.

This guiding principle reaffirms the principle in Article 36 of the 1977 Additional Protocol I to the 1949 Geneva Conventions. The United States is not a party to the Additional Protocol I and does not regard Article 36 as reflecting customary law, but engages in robust practice of conducting reviews of the legality of weapons. Such reviews are a good practice to facilitate the implementation of international law applicable to weapons and their use in armed conflict.

“Emerging technologies” are novel by definition and thus may be construed as “new” under this principle. The use of autonomy in weapon systems, however, is not necessarily new. There is substantial State practice in using autonomous functions and features in weapon systems for decades.

In that light, the United States proposes the following good practices for the legal review of weapons systems based on emerging technologies in the area of LAWS for the GGE's consideration.

1. Legal advisers should be consulted regularly in the development or acquisition process as decisions that could pose legal issues are being made so that legal issues can be identified and more in-depth reviews can be conducted where necessary.
 - a. A weapon system under modification should be reviewed to determine whether the modification poses any legal issues.

- b. New concepts for the employment of existing weapons should also be reviewed, when such concepts differ significantly from the intended uses that were considered when those systems were previously reviewed.
2. The nature of the legal review and advice should be tailored to the stage of the process of developing or acquiring the weapon.
 - a. Providing legal advice early in the development or acquisition process allows IHL considerations to be taken into account early in the life cycle of the weapon.
 - b. At the end of the development or acquisition process, formal legal opinions can memorialize relevant conclusions and analysis while also being useful to consider in subsequent reviews.
3. The legal review should consider the international law obligations applicable to the State intending to develop or acquire the weapon system, including prohibitions or other restrictions applicable to specific types of weapons, and whether the intended or expected uses of the weapon system can be consistent with those obligations under IHL.
4. The legal review should consider whether the weapon is illegal per se, *i.e.*, whether the use of the weapon is prohibited in all circumstances.
 - a. The legal review should consider whether the weapon is of a nature to cause superfluous injury or unnecessary suffering, or if it is inherently indiscriminate, or is otherwise incapable of being used in accordance with the requirements and principles of IHL.
 - b. Analyzing whether a weapon is “inherently indiscriminate,” should consider whether the weapon is capable of being used in accordance with the principles of distinction and proportionality.
 - c. In considering whether a weapon with new autonomous features or functions is consistent with the prohibitions against weapons calculated to cause superfluous injury or against weapons that are inherently indiscriminate, it may be useful to compare the weapon to existing weapons not falling under these prohibitions.
5. The legal review should advise those developing or acquiring the weapon system or its concepts of employment to consider potential measures to reduce the likelihood that use of the weapon will cause harm to civilians or civilian objects.
6. Persons conducting the legal review should understand the likely effects of employing the weapon in different operational contexts. Such expectation should be produced through realistic system developmental and operational test and evaluation.

7. Bearing in mind national security considerations or commercial restrictions on proprietary information, States should share good practices on weapons reviews or legal reviews of particular weapons where appropriate.

(f) When developing or acquiring new weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, physical security, appropriate non-physical safeguards (including cyber-security against hacking or data spoofing), the risk of acquisition by terrorist groups and the risk of proliferation should be considered.

The responsible development and use of new weapons systems based on emerging technologies in the area of LAWS should consider a variety of issues, including those not addressed specifically by IHL. In U.S. military practice, DoD Directive 3000.09 requires that in order to mitigate the potential consequences of an unintended engagement or loss of control of a system to unauthorized parties, “physical software and hardware will be designed with appropriate [...] safeties, anti-tamper mechanisms, and information assurance [...]”⁵

(g) Risk assessments and mitigation measures should be part of the design, development, testing and deployment cycle of emerging technologies in any weapons systems.

Risk assessments and mitigation measures are useful tools to address the uncertainty in the anticipated pace and trajectory of the future development of emerging technologies. Risk assessments allow for a weighing of the benefits of the emerging technologies against potential risks and allow for adjustments to be made as further research and development occurs. Risk assessments can also support the training of commanders and operators by helping them understand the function, capabilities, limitations, and likely effects of using a weapon system.

The GGE should build on the work reflected in paragraphs 23(a) and 23(b) of its 2019 report by further cataloging potential risks and mitigation measures that should be considered in the design, development, testing, and deployment of weapons systems based on emerging technologies in the area of LAWS.

(h) Consideration should be given to the use of emerging technologies in the area of lethal autonomous weapons systems in upholding compliance with IHL and other applicable international legal obligations.

This principle recognizes that emerging technologies in the area of LAWS can be used to provide benefits, such as strengthening the implementation of IHL and reducing the incidence of civilian casualties and other tragic outcomes in armed conflict that may occur even when all parties have complied with the law.

This principle should be implemented during legal reviews of new weapons, during the formulation of military strategies and plans, and during the planning and conduct of military operations. To facilitate such consideration and to encourage innovation that furthers the objects and purposes of the CCW, the GGE should develop examples of specific practices that those

⁵ *Id.* at paragraph 4(a)(2)(a).

involved in these activities could consider. For example, the GGE could begin this workstream by cataloging examples of ways in which emerging technologies in the area of LAWS could be used to reduce risks to civilians in military operations, such as by:

1. incorporating autonomous self-destruct, self-deactivation, or self-neutralization mechanisms into munitions;
2. increasing awareness of civilians and civilian objects on the battlefield;
3. improving assessments of the likely effects of military operations;
4. automating target identification, tracking, selection, and engagement to improve speed, precision, and accuracy; and
5. reducing the need for immediate fires in self-defense.⁶

(i) In crafting potential policy measures, emerging technologies in the area of lethal autonomous weapons systems should not be anthropomorphized.

Anthropomorphizing emerging technologies in the area of LAWS can lead to legal and technical misunderstandings that could be detrimental to the efficacy of potential policy measures. From a technical perspective, anthropomorphizing emerging technologies in the area of LAWS can lead to mis-estimating machine capabilities. From a legal perspective, anthropomorphizing emerging technologies in the area of LAWS can obscure the important point that IHL imposes obligations on States, parties to a conflict, and individuals, rather than machines. “Smart” weapons cannot violate IHL any more than “dumb” weapons can. Similarly, machines are not intervening moral agents, and human beings do not escape responsibility for their decisions by using a weapon with autonomous functions. Anthropomorphizing emerging technologies in the area of LAWS could incorrectly suggest a diminished responsibility of human beings simply by the use of emerging technologies in the area of LAWS.

(j) Discussions and any potential policy measures taken within the context of the CCW should not hamper progress in or access to peaceful uses of intelligent autonomous technologies.

Technology should not be stigmatized. Autonomy-related technologies, such as artificial intelligence (AI) and machine learning, have remarkable potential to improve the quality of human life with applications such as driverless cars and artificial assistants. The use of autonomy-related technologies can even save lives, for example, by improving the accuracy of medical diagnoses and surgical procedures or by reducing the risk of car accidents. Similarly, the potential for these technologies to save lives in armed conflict warrants close consideration,

⁶ These practices are discussed in the U.S. Working Paper, *Humanitarian Benefits of Emerging Technologies in the Area of Lethal Autonomous Weapon Systems*, March 28, 2018, [CCW/GGE.1/2018/WP.4](#). For a discussion of other potential humanitarian benefits, in addition to reducing the risk of civilian casualties in military operations, see paragraph 15 of the U.S. Working Paper, *Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems*, March 28, 2019, [CCW/GGE.1/2019/WP.5](#).

including potential applications to help uphold IHL as reflected in guiding principle (h). As a result, research and development on autonomy-related technologies should not be restricted based on the rationale that such technologies could be used for weapons systems. Moreover, although the use of technologies for the purpose of violating international law, must be condemned, the use of autonomy-related technologies for defensive or other beneficial purposes should remain unhindered.

(k) The CCW offers an appropriate framework for dealing with the issue of emerging technologies in the area of lethal autonomous weapons systems within the context of the objectives and purposes of the Convention, which seeks to strike a balance between military necessity and humanitarian considerations.

The United States strongly supports the CCW GGE as the appropriate multilateral forum for States to address emerging technologies in the area of LAWS because States can use the GGE to engage in a substantive, non-politicized dialogue on IHL issues. The GGE allows States to send technical, legal, policy, and military experts as part of their delegations, submit working papers, and exchange State practice. The CCW GGE is open to all States, includes States with relevant practice, and develops its reports by consensus. Civil society participants can observe the proceedings and participate in the discussions. The High Contracting Parties to the CCW have successfully put this framework to use in their consideration of emerging technologies in the areas of LAWS as reflected in GGE's substantive reports and the guiding principles.