Defense Innovation Board Listening Event on Artificial Intelligence and Ethics

Public Remarks by Chuck Allen, Deputy General Counsel for International Affairs, Department of Defense, Stanford University, April 25, 2019

I'm happy to have this opportunity to give some background about the Department of Defense's (DoD) legal practice, especially our strong commitment to law and ethics. We are in fact committed to legality, to compliance with the law in all DoD programs, activities, and operations, including compliance with the law of war in all military operations. The law of war is also known as the law of armed conflict and international humanitarian law or IHL – which is the term I'll mostly use today.

Our leaders have especially focused on ethics. Acting Secretary Shanahan's recent memo to the Department stated: "A key component of leadership is reinforcing ethical behavior across the full spectrum of our work and recognizing ethics principles as the foundation upon which we make sound, informed decisions."

These words reflect the approach that DoD also applies to Artificial Intelligence (AI). Earlier this year, DoD released an unclassified summary of its strategy on AI. One of the strategy's five pillars is "Leading in military ethics and AI safety." The strategy describes DoD's intention to lead in AI ethics and safety by articulating and realizing its vision and guiding principles for using AI in a lawful and ethical manner, including to improve civilian protections under the law of war, investing in research and development for resilient, robust, reliable, secure, and explainable AI, and promoting transparency.

Our work in the responsible use of AI builds on existing DoD policies. One is DoD's directive on Autonomy in Weapon Systems, initially issued in 2012 and updated in 2017. Then-Deputy Secretary of Defense Ash Carter issued the Directive – which is binding on the entire Department – after careful study of past DoD practice in using autonomy in weapon systems, such as the Patriot missile system and the Aegis Combat System. The Directive captured lessons learned from past experience in using weapon systems with autonomous functions.

It establishes strong guidelines to minimize the probability of failure and the consequences of any failure in autonomous and semi-autonomous weapon systems that could lead to "unintended engagements" – such as unexpectedly harming friendly forces or civilians.

The Directive provides that "autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force." It requires realistic and rigorous testing, clear human-machine interface, and training for commanders and operators that is commensurate with their responsibilities.

In addition, compliance with the law, including international humanitarian law, underlies all DoD programs and operations. It is foundational as well to the Department's responsible use of Artificial Intelligence.

International humanitarian law itself is a well-established body of law that governs conduct during armed conflict. The U.S. military has a long history of law of war compliance. This is

touched on in the DoD Law of War Manual, which I'll say more about. I'm glad that the chief author of that Manual is here today – Karl Chang.

For the United States, the law of war includes treaties the United States has accepted, such as the 1949 Geneva Conventions as well as customary international law. The law of war seeks to reduce unnecessary suffering in war; it protects civilians and other persons who are *hors de combat,* or "out of the fight." The law of war permits targeting enemy combatants and military objectives, like enemy tanks or aircraft, but prohibits targeting civilians and protected objects like hospitals or museums.

As I mentioned, DoD's long-standing policy is to comply with IHL in all military operations. DoD has robust processes to implement IHL, including for training, required reporting of incidents involving alleged violations by anyone, including our forces or foreign forces, investigations and reviews of incidents, and corrective actions, which include discipline and judicial processes as appropriate. We have our 1,200-plus-page DoD manual that collects and explains law of war requirements. It is authoritative guidance for DoD personnel and is available publicly online at the DoD General Counsel's website along with other official documents on the law of war. DoD also has thousands of military and civilian lawyers who advise commanders and decision-makers on legality, including IHL compliance, wherever our people are around the world. They review the intended acquisition of weapons by the myriad systems commands of the Military Departments around the country and up to the sub-Cabinet and in some cases the Cabinet level in Washington, DC. They also review all military plans, operations orders, and rules of engagement.

Let me mention three ways in which international humanitarian law can inform the DoD's responsible use of Artificial Intelligence. I don't mean to suggest that IHL provides all the legal answers on AI. As noted in the DoD AI strategy, the Department is seeking to "harness the potential of AI to transform <u>all functions</u> of the Department positively" (emphasis added) – and not just as AI relates to weapons, or to the Nation's wars. Depending on the type of application, AI could raise other legal issues, for example, privacy, civil liberties, or medical ethics issues. As for IHL, however, DoD has deeply internalized compliance with it. As the DoD Law of War Manual puts it – "the law of war is part of who we are." You will find this is true among all the Military Departments including U.S. Soldiers, Sailors, Airmen, and Marines around the world and Pentagon senior leaders alike. Because the law of war reflects the Department's ethical values, like accomplishing the mission and protecting civilians, IHL absolutely is an important part of the Department's approach to using AI in armed conflict.

First, existing IHL rules apply when new technologies, such as cyber operations or new types of AI, are used in armed conflict. The law of war actually anticipates technological innovation and contemplates that its existing rules will apply to new means and methods of warfare. IHL rules can apply to new technologies because, although some rules are specific to certain types of weapons like mines or incendiaries, IHL rules generally are not framed in terms of specific technological means. For example, the rules on conducting attacks do not depend on what type of weapon is used to conduct the attack. So, if AI is used to help the DoD perform an activity that is subject to IHL, like conducting an attack, that activity continues to be subject to the applicable IHL rules – including that attacks must be directed against military objectives, attacks

must not be conducted when they are expected to cause excessive collateral damage, and feasible precautions must be taken to reduce the risk of harm to civilians.

Another way IHL can inform DoD's use of AI is through its principles. The law of war is based on certain fundamental principles – including military necessity, humanity, distinction, proportionality, and honor.

These principles provide a general guide for conduct during war, where no more specific rule applies. Said another way, even if there is not a specific international law rule that applies, by no means is it the case that "anything goes." Rather, it is still – always -- important to consider whether, under IHL, a contemplated action is justified by military necessity, and to ensure an action is not unreasonable or excessive under the principle of proportionality.

In this way, the fundamental principles of IHL provide a framework to consider novel legal and ethical issues posed by emerging technologies, such as AI. In international discussions on emerging technologies in lethal autonomous weapon systems, the United States has noted that the principles of the law of war provide a basis for assessing new technologies. We have said in Geneva at the Group of Governmental Experts (GGE) on Emerging Technologies in Lethal Autonomous Weapon Systems (LAWS) of States Parties to the Convention on Conventional Weapons:

> [I]f the use of a new technology advances the universal values inherent in the law of war, such as the protection of civilians, then the development or use of this technology is likely to be more ethical than refraining from such use.

This leads to my third point about how IHL informs the DoD's use of AI – using AI expressly to improve our implementation of civilian protections.

The law recognizes that civilian casualties are a tragic, but, at times, unavoidable part of armed conflict. IHL requires that militaries take feasible precautions to reduce the risk of civilian casualties.

Just as new advancements in AI have the potential to save lives in peacetime by reducing the risk of car accidents or improving medical diagnoses, advancements in AI could also help save the lives of civilians during war by improving precautions we can take to protect civilians.

One way AI could do this is by making weapons more accurate. "Smart" weapons that use computers and autonomous functions to deploy force more precisely and efficiently have been shown to reduce risks of harm to civilians and civilian objects. Weapons systems with automated target recognition allow the weapon to lock on to targets and strike military objectives more accurately and with less risk of harm to civilians and civilian objects. And, when the weapon is more accurate, fewer weapons need to be fired to achieve the same military advantage.

Another way that AI could create humanitarian benefits is by improved military awareness of civilians during military operations. Civilian casualties can result from a lack of awareness of where civilians are on or near the battlefield due to the "fog of war" – information available during war is limited and unreliable often due to the of chaos of combat and the opposing sides'

efforts to deceive one another. Due to the "fog of war," commanders might be unaware that civilians are in or near a military objective. AI very likely could help reduce these kinds of mistakes.

As a case in point, many of you may have heard about Project Maven – a DoD effort to use AI to improve analysis of video from intelligence, surveillance, and reconnaissance platforms. By using AI to identify objects of interest from imagery autonomously, analysts can search through larger quantities of data and focus on the more sophisticated and important tasks requiring human judgment. We believe this kind of work could help improve the commander's battlespace awareness and help cut through the "fog of war." This could mean better identification of civilians and civilian objects on the battlefield, which allows our commanders to take steps to reduce the risk of harm to them.

A common theme underlying the potential humanitarian benefits of AI on the battlefield is a convergence between military and humanitarian interests. When use of force is necessary to protect and defend our country and our people, the military wants to use force more accurately, precisely, efficiently. We want to make faster and more accurate decisions. These military advantages also bring humanitarian benefits in terms of reducing the risk to civilians from military operations. Even more directly, perhaps, DoD is already working on ways to use AI capabilities for humanitarian purposes. The Joint Artificial Intelligence Center has a national mission initiative on humanitarian assistance and disaster relief that is applying lessons learned and reusable tools from Project Maven to field AI capabilities to help first responders and local agencies save lives when responding to wildfires and hurricanes.

I'll close now -- eager to listen and learn from this session today. We recognize the need to be open and to embrace a diversity of viewpoints. DoD respects the right of individuals and U.S. companies and their workforces to express their views. Of course, this is one of the very freedoms guaranteed by our Constitution that we in DoD are sworn to defend. On an issue like artificial intelligence, we have much to learn from perspectives outside the Department, and we look forward to learning from many of you today.