



Statement of the UN Institute for Disarmament Research

at the CCW Meeting of Experts on Lethal Autonomous Weapon Systems

13 May 2014

Delivered by Kerstin Vignard, Chief of Operations, UNIDIR

check against delivery

Thank you, Mr Chairman, for giving us the opportunity to offer a short brief on research underway at the UN Institute for Disarmament Research.

In 2013 the Secretary-General's Advisory Board on Disarmament Matters recommended that the Secretary-General consider commissioning a comprehensive analysis on the development, proliferation and use of increasingly autonomous weapons technology—and suggested UNIDIR as being able to carry out such a study.

As Member States have yet to decide whether to commission such a study, in the interim — with the support of the governments of the Netherlands and Switzerland—UNIDIR has launched an 18 month project on the weaponization of increasingly autonomous technologies. This project is part of UNIDIR's Research Programme on "Security and Society". I should also note that this project is considering the issue in a broad way, and isn't specifically focused on the CCW discussions.

The project is focused on considering **what information is relevant** for good decision making in this domain and **whether that information is presently available to policy-makers**. Rather than being primarily concerned with offering specific policy recommendations, the project's aim is to help policy-makers think about autonomy, to challenge assumptions and to ask better questions of themselves and others..

At the heart of the project is a small group of experts who have agreed to explore—in a collaborative way —various facets of this topic. These experts represent a range of disciplines—artificial intelligence and computer science, ethics, science and technology, policy and programme design, humanitarian concerns, IHL, human rights, regional security, defense, diplomacy, and policy research. This approach helps to identify areas where different stakeholders might work together to refine the areas of concern, as well as identify useful research and linkages in other domains that may be of relevance.

We held the first meeting of the expert group in March, where we focused on mapping the issue. The conversation was very rich and served as the basis for the observation paper which I will turn to in a moment. The group will meet twice more this year. Our next meeting will be dedicated to two themes—first, consideration of human control over weapon systems, and particularly the decision to use force as this cannot be divorced from discussions on increasing autonomy. The second topic will be autonomy in the marine and underwater environment. We hear very little in international fora about undersea autonomy yet we note that many experts agree that the underwater environment is where we will see the most rapid advances in autonomous technologies.

In the later stages of this project, UNIDIR intends to consider questions related to how the proliferation of increasingly autonomous systems might alter regional security dynamics; whether autonomous technologies might drive development of other weapons, counter-measures or methods of concern such as cyber-conflict; and what might be the impact of these technologies on asymmetric warfare and terrorism, to name a few.

We are actively seeking funds for both the continuation of this work and its expansion. I am of course happy to discuss with any delegation who might be interested in joining with the Netherlands and Switzerland to support the Institute's work in this area.

At the end of April, UNIDIR released a short paper entitled "Framing discussions on the weaponization of increasingly autonomous technologies." If you haven't already seen it, there are copies available in the back of the room and the paper can be downloaded from UNIDIR's website.

The paper makes **four observations** for policy makers' consideration as they approach discussions both in multilateral fora such as the CCW and the Human Rights Council, as well in their national policy dialogues. I will just briefly mention the observations—these are explored in greater detail in the Framing Paper itself.

1. Consider the variables that make up autonomy

In the initial stage of discussions, **States might consider a functional approach focused on identifying the critical functions of concern and the interactions of different variables that limit or augment the characteristic of autonomy**—we suggest some of these variables in the paper. This approach would anchor the discussion and set its boundaries. It would also allow discussions to bypass—for the time being—a technology-centric definitional exercise. If definitions are later determined to be necessary, taking a functional approach now will help inform the definition so that the truly essential elements are captured.

2. Consider the drivers

The drivers behind **military interest** in increasingly autonomous technologies are well known, such as

- greater force projection;
- risk reduction for "dirty and dangerous" missions;
- increased speed of decision making;
- dwindling defence budgets coupled with the high costs of military personnel; and
- a belief by some that autonomous weapon systems may eventually be able to respect international humanitarian law or human rights law better than humans do.

However, a second critical yet underappreciated driver is the **civilian technology sector**—where the robotics revolution is well underway. Industry, the scientific community and even consumers will drive expectations and investment in further advances in autonomous technologies. This is not just a question of dual-use aspects technology. It is also about

socio-cultural perceptions of technology--and these are not universal. As civilians become more familiar with autonomous machines in their daily lives, perceptions of—and trust in—the capabilities of machine decision-making is likely to change.

3. Shift from a technological and innovation-centric frame to one that addresses acceptability, impacts and longer-term consequences of use

Peter Singer, author of *Wired for War*, has noted the danger of technology-centric frames: "... too often in discussions of technology we focus on the widget. We focus on how it works and its direct and obvious uses. ... Indeed, with robotics, the issues on the technical side may ultimately be much easier to resolve than dilemmas that emerge from our human use of them."

We must acknowledge that humans have a poor record of predicting the full range of benefits and risks associated with new technologies. Often technologies are developed for one set of tasks but then adopted in other fields for missions not envisaged by their designers. Concerns have been raised, for example, that deployment of increasingly autonomous weapon technologies will begin in uncluttered environments and steadily migrate into more complex ones, perhaps without a State undertaking a new Article 36 or policy review that takes into consideration the different operating environment. In addition, humans, including soldiers, have a tendency to modify technologies to overcome safety features and controls.

4. Beware of how you define the area of concern

The widely used verbal shortcuts we use in these discussion (such as the man "in", "on" or "out" of the loop description, or the categories of remotely controlled, automatic, automated and autonomous), are perhaps not the only, or even the most useful, distinctions to make.

For example, much focus has been placed on questioning the acceptability and legality of robots taking lethal decisions. However might it also be an issue that machines take decisions to use so-called non-lethal force, for example firing rubber bullets, riot control agents or beanbag rounds?

Mr Chairman, distinguished colleagues,

There is considerable disagreement among experts on the state of the technological development—and the timelines for their integration in weapon systems. That said, robotic autonomy writ large is an area of extremely high investment by the private sector as well as the military and thus **the components of increasingly autonomous technologies (for both civilian and military applications) will continue to improve, even if the pace of improvement is open to question.**

Therefore, the technological capability to design autonomous weapon systems that can detect and analyse complex environments, select targets and carry out an attack is likely to be reality one day—even if that day is far in the future. But **the decision to weaponize these capabilities is not inevitable.**

It is crucial to give significant consideration today to the question “If a weapon system were ABLE to do X, would we WANT it to do so?” This question offers opportunity for reflection that go beyond legal assessments to other fundamental considerations such as the right to life and protection of human dignity. This discussion will benefit from increasing interactions between the security community with that of human rights, defence, and other disciplines more widely, including ethics and philosophy, technologists, psychologists, sociologists, scientists and engineers, and others.

Broad and well informed discussions such as those starting today will help us identify the areas that evoke clearer responses that can be the basis of policy development, and those that entail greater ambiguity or uncertainty—and therefore require deeper consideration by States, researchers, the private sector and civil society. Therefore, we wish you a productive week of exchange on this very important topic. UNIDIR stands ready to assist Member States in this endeavour.

Thank you, Mr Chairman.