Federal Republic of Germany
Foreign Office

CCW Expert Meeting

Lethal Autonomous Weapons Systems (LAWS)

Geneva 13 – 16 May

Technical Issues – Summary by Michael Biontino

During the whole debate on technical issues of lethal autonomous weapons systems, the notion of autonomy and its definition was at the center of interest. It became quite obvious that there is no ready-made, generally accepted definition of what is an "autonomous system" and as to where to draw the line between "autonomous" and "automatic" or "automated".

In the first place, I will line out the key messages, as I understood them, of the different presentations we have heard. Then, I will highlight some of the main topics of the quite lively debate.

Prof. Raja Chatila supported the thesis that autonomy is a relative concept. It is related to the complexity of the environment with machines having greater autonomy the simpler the environment is. Furthermore, he underlined that a clear-cut border between automaticity and autonomy is difficult to establish and that there is rather a continuum between the two. Prof. Chatila furthermore distinguished two types of autonomy: operational and decisional autonomy. While operational autonomy is already present in existing systems, decisional autonomy would require some sort of reasoning by machines which would be very difficult to achieve.

Mr. Paul Scharre elaborated on how autonomy is used in weapons today. He elaborated on the three different but interrelated indicators which are machine freedom, complexity and the tasks performed. Depending on these indicators, the inherent risk of autonomous machine action might be evaluated differently. Mr. Scharre also briefly presented the US Directive 3000.09 that could give orientation in defining at which points human involvement was necessary in order to minimize risks of machine action depending on the environment. He encouraged more states to develop rules for the appropriate use of autonomy in weapons.

Prof. Ron Arkin and Prof. Noel Sharkey discussed their very different views on whether lethal autonomous weapons systems would ever be able to comply with international humanitarian law. Prof. Arkin pointed to the possibility to program, in the future, a weapons system to act more humanely than a human soldier and therefore advocated that it would be an imperative to use such a system in order to prevent unnecessary casualties. Some of the advantages of LAWS would be that they could act more conservatively than humans in case of low certainty

of target identification, that they would not have emotions such as revenge, that they would be able to integrate more information from more sources far faster than men and that they would be capable of objectively monitoring ethical behavior in the battlefield by all parties and reporting infractions. However, Prof. Arkin advocated for regulating the use of LAWS and limiting it to specialized missions with a minimal likelihood of civilian encounter.

Prof. Sharkey, on the other hand, called for a ban of LAWS as he did not believe that machines could ever comply with international humanitarian law. In his views, computers are just not able to make qualitative assessments as required, for example, by the principle of proportionality. Prof. Sharkey stressed the principle of meaningful human control when deploying LAWS. The role of the human in the kill chain should be clarified and command and control structures be made more transparent. This would also lead to clearer accountability and responsibility.

Prof. Jean-Paul Laumond elaborated on the state of the art in humanoid robotics. He described the enormous technological challenges presented by the development of such robots as even quite simple movements require the coordination of a multitude of different factors. However, in my opinion, the key message of his presentation was that robots, even the ones we might call autonomous, in the end are and will always be machines that execute automatically a program defined by humans. Robots cannot think or decide. It is always the human behind who thinks or decides.

Mr. Hajime Wakuda stressed as others that there is a continuum between automation and autonomy. A fully autonomous machine would have to be able to create purpose. He furthermore explained that the research and technology necessary to develop robots would be the same for civilian and military uses. Robot technology itself does not have the function of killing. The difference between civilian and military would lie in the use of the technology, not in the technology itself.

Mr . Yong Woon Park underlined that autonomy should be measurable, implementable, comparative to human capacity and handled by an operator or commander. Autonomous systems should always be linked to a human in order to use all resources of the human and of the machine to achieve the best performance. Humans should always be in the responsible loops (mission, operation, control even task handling) when a robot is working.

With that, I want to come now to the main points of our debate. There was a really lively debate and I will just point out some of the key notions that have been raised as I recollect them. This does not imply any prioritization or assessment of the points made.

It was underlined that LAWS are still an emerging technology. Thus, object of this conference are weapon systems that do not exist yet. Already existing highly automated weapons systems were explicitly exempted from this debate by many delegations.

The complexity of lethal autonomous weapons systems was underlined. They consisted of a combination of very different technologies, such as sensors and recognition, mobility, payload integration etc. All of these technologies were still in the phase of development. Before

LAWS became a reality, substantive advances in all of these technologies would be required. Thus, talking about LAWS was talking about a distant future.

Despite the acknowledgement that there is no generally-accepted definition of lethal autonomous weapons systems, there has been no debate about what this definition should be. Some delegations even questioned the necessity to start the discussion on a definition at this stage or argued that a definition at this stage was impossible because there was no clarity at all about what kind of autonomous weapons systems would be possible to achieve in the future. Some possible elements of a definition were however mentioned such as the lethality, the self-learning capacity of LAWS and thus the unpredictability of the concrete steps to achieve a predetermined outcome.

When thinking about the possible future emergence of fully autonomous weapons systems, there was a widespread understanding that some sort of human control should always be maintained. However, there was no clear definition of what this control should look like. In the eyes of many delegates, this meaningful human control required, however, a sufficient contextual awareness, sufficient time to deliberate and the ability to intervene and possibly abort an attack. The required human control would also depend on the environment in which autonomous weapons systems were used. In simple, uncluttered environments there was more room for machine autonomy than in complex situations.

Interventions also underlined the inherent dual-use character of robot technology with many civilian uses of robots being very positive and welcome. Thus, an outright research ban or ban of the technology did not seem feasible. The only option would be to regulate the use of robot technology.

The notion of “responsible innovation” was mentioned calling on scientists to consider all possible effects of their research from the beginning.

The mentioned military necessity for weapons that are able to react much faster und thus in the end for autonomous weapons was questioned. In developing and fielding always faster weapons man could create this military necessity himself. It would be like a self-fulfilling prophecy. Thus, further progress in always faster reacting, more sophisticated and in the end autonomous weapons would possibly lead to a new arms race.

Additionally, the risk of proliferation of LAWS was referred to by various delegations. As soon as one state would have LAWS, others would follow. There could also be possible proliferation of LAWS to other sectors than the defence sector, such as prosecution and policing and, as well, to non-state actors and terrorists.

Finally, the question of the error margin of LAWS was raised and how safeguards against failures could be built in the systems.

To sum up: In my opinion, we had a very fruitful discussion that pointed out a number of difficult questions we will have to elaborate on further. A positive outcome was the emerging consensus on the need for meaningful human control over lethal weapon systems. But I want to immediately add a word of caution: This apparent consensus on the need for meaningful human control, however, does not mean that all delegations have the same understanding of

what meaningful human control actually means. So I would like to conclude, that we had a very good debate which gave us a lot of input for further discussions in the future.