

CCW comments – Richard Moyes

I am going to comment briefly on approaching the issues raised by autonomous weapons systems with a focus on "human control." We have printed up some more detailed background to the comments I am making here – and that is available at the back of the room.

I will comment on:

- the general approach based on human control,
- on how we see this in relation to international humanitarian law (IHL),
- on the key elements of human control,
- and a final point on approaches to working definitions.

× **Consideration of the form and nature of human control considered necessary is the most useful starting point for discussions on this issue.**

So, the first point I want to make is that we think that focusing discussions on the human control that is necessary over technologies is the best starting point. Debating hypothetical future technologies is inevitably very difficult. But our anxieties about 'autonomy' stem from the concern that such systems would remove a human element that might be considered essential. So we should start by focusing on that.

And although there might be some debate about which terms are preferred, it feels like there is some broad agreement that human control is necessary – in some form.

As an extreme example, I think we would all agree that a weapon system operating with no human control whatsoever would be unacceptable. And at the same time, I think we would recognize that the mere engagement of a person in a loop is not enough to be sufficient. For example, if a person in a dark room simply pressed a fire button when told to by a computer, they would hardly be asserting a meaningful level of human control.

So we probably all recognize that some human control is necessary, and that it needs to be substantive. We have used the term 'meaningful' to express that – but delegations shouldn't fixate on that word, because it really functions to say: you need to further explain what the necessary form of human control is.

There was a comment yesterday to the effect that meaningful human control might be considered rather subjective. But I think the same could be said of other formulations. Appropriate human judgment for example. All such approaches require further description or definition. In fact it should be the work of this group to move from our separate subjective ideas about what form of human control is necessary, to a shared understanding. That is the work that needs to be done and should be the focus of state's work next year.

× **International humanitarian law provides a framework that should be understood as requiring human judgment and control over individual "attacks" as a unit of legal management and tactical action.**

The second point I want to make is related to "control over what"? We call for recognition that there needs to be meaningful human control over "individual attacks." This links our concept of control to the legal framework of IHL.

IHL is addressed to humans, collectively or individually. Humans apply its legal rules. We are talking about machines, not legal agents – and humans have obligations to apply the legal rules in their use

of these machines.

And in the law, there are obligations that bear on commanders when they plan and decide upon "an attack." It is clear in the law that a human commander needs to assess and decide upon a specific military objective against which an attack will be directed. There is more detail on this in our paper – but essentially, at the tactical level of military operation, the term "attacks" provides both a unit of legal management and of military action.

An attack may contain multiple applications of kinetic force to specific objects, but it must have some boundaries to it if the structure of the law is to function. And it follows in our thinking, that because there is an obligation for human legal judgment in relation to individual attacks, a machine cannot set a military objective, nor can it undertake multiple "attacks" in this sense, without human legal judgment and control being applied.

This is not an argument about whether future technologies have this capacity or that capacity, but this is about the structure of the law as it is written now, addressed to humans.

- × **That without recognizing a requirement for human control to be in some way substantial or meaningful, the existing legal framework does not ensure that human legal judgment will not be diluted to the point of being meaningless, as a result of the concept of "an attack" being construed more and more broadly.**

And this brings us on to a key concern regarding developing autonomy and the existing legal framework. A concern that means that simply demanding 'compliance' with the law, or asserting the adequacy of legal reviews, is inadequate as a response to the issue.

Because our analysis is that without some affirmation of the need for human control, developing autonomy will result in the term "an attack" being interpreted more and more broadly – to allow for more specific applications of kinetic force, over a wider area and against a more diverse set of target objects - which would effectively be an erosion of the case-by-case application of the law.

- × **Vehicle target slides**

And here I have just tried to illustrate this very simply. If we imagine a weapon system that targets fighting vehicles.

In this first scenario, there is a group of vehicles that a commander will target. And because they are all geographically quite close together we can probably imagine quite easily how a commander would assess the legality of this as an individual attack.

But our concern regarding growing autonomy is that a similar weapon system could start to be used over a wider area and over a longer period of time. And so in this example, with the vehicles spread out it becomes much more difficult to be confident that this should really be approached as one individual attack.

If the commander does not know the specific locations where force will be applied, the legal assessment becomes more difficult.

Autonomous weapons threaten to expand this concept of an attack, in a way that will dilute the law.

- × **Against that background, delineation of the key elements of human control should be the**

**primary focus of work by the international community.**

So, given broad agreement that some form of human control is needed. And given that without some articulation of that human control we may see an erosion of the legal framework, the international community should begin to lay out the key elements of the human control that is necessary. And towards that process we have some suggestions:

- × **Key elements can be proposed:**
  - × **Predictable, reliable and transparent technology.**
  - × **Accurate information for the user on the outcome sought, the technology, and the context of use.**
  - × **Timely human judgement and action, and a potential for timely intervention.**
  - × **Accountability to a certain standard**

We think that human control is driven by four key elements – relating to the technology, information, human action and accountability.

**Firstly** the technology itself needs to operate predictably and reliably. Of course nothing is completely reliable, but the more predictable and reliable it is, the more easily its use can be controlled.

**Secondly** the user needs to understand the outcomes that they are working towards, they need to understand the technology, and they need to understand the context in which it will be used.

I mentioned the predictability of technology before. Predictability is partly a characteristic of the technology, but also of the interaction of that technology with the operating environment.

So information on context is very significant. We should have some understanding of the environment in which the technology will operate, the presence of civilians and civilian objects for example.

Of course we may not achieve complete predictability – already in the use of weapons we accept degrees of uncertainty about the actual effects that will occur. And we know that there may be limitations on the information available about the context.

But our ability to understand the context is directly linked to both the size of the area within which the technology will operate, and the duration over which it will operate.

For any given environment – it follows logically that greater area and longer duration of independent operation result in reduced predictability and reduced control.

But we have recognized that different environmental domains present different general characteristics – with land, air and sea being different. Such that a large area of operation in the sea, may still facilitate better contextual understanding than a smaller area on land.

And from our understanding of the technology, and our understanding of the context we should be able to assess likely outcomes – including the risk of civilian harm that is the basis for a legal assessment.

**Thirdly**, - based on this information – which may itself be the product of wider systems - we need humans to apply their judgment – as required by the law – and to choose to use a specific weapon system in a specific way. And we note in a timely manner here – because the accuracy and relevance

of this information about context, for example, gets less relevant over time.

And timely action might be to intervene and to stop the independent operation of a system.

**And finally**, all of this needs to exist within a framework of accountability. I will not go into that here, but structures of accountability are vital to the overall system of human control over the use of force.

So: Technology, information, human judgment and action, and accountability.

**To sum up on this**, consideration of these key elements does not provide immediate answers regarding the form of control that should be considered sufficient or necessary, but it does provide a framework within which certain normative understandings should start to be articulated.

And really I think that further developing those understandings should be the main focus of work for this group next year. To produce a focused discussion and one that allows us to make progress.

× **Approach working definitions based on understanding 'lethal autonomous weapons systems' as weapons systems operating with elements of autonomy and without the necessary forms of human control.**

And to conclude – I want to link this sense of the necessity of some human control to how we might approach working definitions.

And strategically, at this point, we think it is useful to adopt a working definition of LAWS that matches that term to the problem we are concerned with.

If we choose to consider lethal autonomous weapons systems, LAWS, as systems operating without meaningful human control, then we will greatly simplify the debate we are engaged in.

Some states in their interventions are already doing that – even though it is not being spelled out in those terms.

This does not preclude other, broader definitions, as indicated in the Swiss paper. And I would agree strongly with their recognition that the term lethal should not be taken to limit the scope of what we are discussing.

But taking this approach to the term LAWS would allow for a focused debate.

It would not be the end of the conversation – but it would make description of the necessary human control a central element of future work.

Thank you.