CHALLENGES TO INTERNATIONAL HUMANITARIAN LAW

Human-machine interaction in terms of various degrees of autonomy as well as political and legal responsibility for actions of autonomous systems

Neha Jain, Associate Professor of Law, University of Minnesota Law School
njain@umn.edu

Thank you, Mr. Chairperson and your Excellencies for the opportunity to speak to you on the very important issue of political and legal responsibility for the actions of autonomous weapon systems. Lawyers tend to have a hard time grasping technology, so it is not surprising that the prospect of machine autonomy in diverse areas of our lives, ranging from competitors who can beat us at go, and maybe even poker, to chaufferless cars, to LAWS, threatens to disrupt our conventional understandings of civil and criminal responsibility. Of these, LAWS have attracted heightened legal scrutiny quite simply because the potential harm and loss of human life arising from their use are much greater. If this harm does indeed come to pass, will it be possible for the law to allocate responsibility for this conduct to one or multiple human beings?

In the few minutes I have, I want to emphasize how the law may think about responsibility for the conduct being effectuated through LAWS.

(Next slide) Here, the term "autonomy" is a bit of a distraction for a lawyer – or rather, more like a red rag to a bull - because it plays on the sense of LAWS as full-blown moral agents who are tireless, merciless, and capable of making their own decisions and executing them with precision. However, as was apparent in the discussions during the past few days, this strong sense of autonomy does not currently exist, and it seems unclear whether it will ever exist. Instead, the sense of autonomy we are working with here is more easily characterized as "emergent". Emergence refers to the ability of the system to engage in unpredictable useful behaviour, which arises due to the interaction of thousands of parallel processes, where small program fragments examine and extract data from a range

of inputs depending on the context. As the system exhibits 'learning', the context used to extract and relate this data may itself undergo changes under the influence of new inputs. Thus, the system will follow pathways that are unpredictable due to the constant slippage between the data and the context.

This weaker sense of "autonomy" helps in narrowing the focus of enquiry to what exactly ails the lawyer obsessed with responsibility for the use of LAWS: it is not because LAWS possess agency that is similar to human agency, but because epistemic uncertainty is built into the very design of the system and is in fact crucial to its successful functioning.

How does the law try to allocate responsibility for this unanticipated action if it results in tremendous harm? Here, two concepts will be crucial: the first is control, and the second is foreseeability.

(Next slide)

The concept of "meaningful human control" has already been raised on a number of occasions, but there is still no consensus on what it would require. In the past few days, several delegates have raised concerns different aspects of the phrase: how do we understand the term "meaningful"/at what point is control relevant, who is exercising control etc.. Very little has been said, however, of how we conceptualise what it means to control something, anything. And this term – "control" – has a pre-existing deeply theorized content in the criminal law. (Next slide) For instance, in civil law systems such as Germany, in order to be responsible as the perpetrator of a crime, an individual should have "control" over the act in question. This element of control signifies perpetrator's ability to execute or obstruct the commission of the offence according to his will and it can take different forms. (Next slide) Control is also central to another form of responsibility in domestic and international law: command responsibility, where a military or civilian superior must have 'effective control' over the subordinate at the time of commission of the act – that is, he should have the material ability to prevent and punish the offences. In each case, the criminal law emphasizes the element of temporality – the control must exist at the *time of commission* of the act.

And this dimension of temporality creates challenges for allocating legal responsibility based on control for acts undertaken through LAWS. (Next slide) For example, the programmer is too remote from the crime to control the conduct at the time of the commission of the offence. This is especially true given that the system will be able to adapt the means and methods it uses to achieve programmer-designed ends. Similarly, the individual who deploys or operationalizes the AWS may not be well equipped to direct or predict the decision-making operation carried out by the AWS in real time.

One possible solution to this problem is to reconceptualise the temporality requirement in evaluating the existence of "control". Thus, the law should concentrate on control at the point when the process to develop the LAWS and the decision to deploy them in certain field operations is made. The control will then be exercised by the commander who reviews the ability of the AWS to function within the limits of the law and gives the authorization to deploy it for a certain operation. This shift in focus will only prove workable if the commander has adequate information about the nature, design and functioning of the weapon to be able to take an informed decision that it is capable of being deployed in a lawful manner. Also, rigorous testing and inspection of the weapon and training of operators will be essential to ensure that they have the requisite information and capability to exercise control and when necessary, abort the use of the weapon.

However, an additional element that may still limit the ability to exercise control is the level of unpredictability associated with the operation of LAWS. Here again, we may look to command responsibility in international criminal law, which imposes liability on the superior for reckless, and in some cases, even negligent conduct of the subordinate (Next slide)

The reckless perpetrator and the negligent perpetrator are both considered to be responsible because they disregarded a substantial and unjustifiable risk that the conducted that the prohibited conduct or result will occur. However, while a reckless defendant is aware of the risk, the negligent defendant is not, but should have been aware of it. If we adopt a

broad concept of recklessness in the law, then even if the defendant was unaware of the exact nature of the risk posed by AWS conduct, he could still be deemed reckless and, hence, liable for the harm. However, we will still need to demonstrate some level of subjective awareness of some kind and degree of risk, and depending on the level of uncertainty we are dealing with for LAWS, this might be difficult to prove in individual cases. (Next slide) Negligence liability for AWS conduct would alleviate this problem to some extent. The adjudicator will have to determine whether the commander/field officer/deploying soldier should have been aware of a substantial and unjustifiable risk of harm resulting from AWS conduct and if, given their circumstances and knowledge, their failure to advert to this risk constituted a gross deviation from the standard of care.

While accepting these broad concepts of recklessness and/or negligence will certainly broaden the potential scope of legal liability, there are important policy questions that will have to be addressed as to the costs of this proposal: most crucially, any steps to embrace the negligence standard more widely must be attentive to the scepticism towards negligence liability in most domestic criminal legal systems.

(Next slide)

The above analysis mostly relates to the criminal responsibility of civilian and military superiors, field officers and the individuals who are in charge of deploying the AWS. Due to their remoteness from the crime, the responsibility of other actors, such as software developers and programmers should be considered under civil law principles of product liability based on a standard of negligence and/or strict liability for harm that results from malfunctions or from poor or error-prone software or hardware design features of the AWS. (Next slide) The foreseeability issue will, however, continue to be relevant even for civil/tortious liability. Even "strict liability" cases of tort liability typically require proof of the foreseeability of one or more of three elements: the kind or type of risk of harm, the person likely to be harmed and the manner of harm.

(Next slide) To conclude, given the current state of technology, the responsibility challenge in LAWS arises from the ability of the system to rely on non-deterministic reasoning in

order to operate in unstructured and dynamic environments. This necessarily entails a level of epistemic uncertainty and unpredictability. For civil as well as criminal responsibility, the concepts of "control" and "foreseeability" point to issues that will be central to developing a paradigm for legal responsibility for the conduct of LAWS and how we should allocate the attendant risks amongst the different actors developing, supervising, and operationalizing their use.