

**Third CCW meeting of experts on lethal autonomous weapons  
systems (LAWS) Geneva, 11-15 April 2016**

**Autonomous Weapons Systems:  
Risk Management and State Responsibility**

**Robin Geiss\***

I am pleased to contribute my views to this third informal CCW Meeting of experts on “Lethal Autonomous Weapons Systems”. The CCW is grounded in the laws of armed conflict. My remarks are focused on the issue of **risk management and state responsibility** arising from violations of this legal order in the context of military operations involving autonomous weapons systems.

1. Traditional accountability models are typically premised on some form of control and/or foreseeability. Higher levels of autonomy in weapons systems, however, mean lower levels of control and foreseeability. Accordingly, the more autonomous a (weapons) system is, the more difficult will it be to establish accountability on the basis of traditional accountability models. This challenge exists with regard to civil uses of autonomous technology (e.g. self-driving cars) in the same way that it exists for military uses of autonomous systems.
2. This, however, does not mean that there is an inevitable or insurmountable “accountability gap”. Especially in the area of state responsibility – the conceptual challenges are greater when focusing on individual criminal responsibility – accountability challenges can be overcome by way of regulation and clarification of existing laws. There is no conceptual barrier for holding a state (or individual human being) accountable for wrongful acts committed by a robot or for failures regarding risk minimization and harm prevention. There is therefore no need to devise a new legal category of “e-persons”, “non-human actors”, or “virtual legal entities” and any idea that a robot could be held accountable should be rejected.
3. It is clear that autonomous weapons systems that cannot comply with the laws of armed conflict must not be fielded. Article 36 API is therefore a logical starting point for any discussions about autonomous weapons systems. But even weapons that have been thoroughly reviewed may fail or malfunction in combat. This is true for any type of weapon. Accidents can happen. There is no such thing as absolute certainty. With regard to autonomous weapons systems there is concern that because of their autonomy – even after thorough testing – there remains **a higher than usual risk and degree of unpredictability** as to how exactly they will operate under actual battlefield conditions and a resultant higher risk of accidents and wrongful conduct.

---

\* Professor of International Law and Security, University of Glasgow, UK.

**The question is how to manage these risks and how to allocate (state) responsibility if something goes wrong?**

5. The answer to this question is obviously highly controversial. Some have argued that because of these risks autonomous weapons systems should be banned. Others have argued that residual risks, i.e. risks that remain after the required Article 36 API weapons review, are generally acceptable. In my view, both approaches go too far. Article 36 API should be understood as an important first step. But further fine-tuning and additional risk mitigation is required.

6. The deployment of autonomous weapons systems is not (per se) unlawful but it is (at least in certain deployment scenarios) a high-risk activity. This novel technology – especially if it is used in a complex battlefield environment – is not (yet) fully understood. **There is predictable unpredictability.** It follows that a State that benefits from the various (strategic) gains associated with this new technology should therefore be held responsible whenever the (unpredictable) risks inherent in this technology are realized.

**On the basis of this rationale a State could be held responsible for failures regarding risk prevention and harm reduction (at the pre-deployment stage) as well as for specific (wrongful) actions of the autonomous weapon system.**

7. **Prevention is better than cure.** The identification and specification of detailed (due diligence) obligations aiming at risk prevention and harm reduction is central. Violations of such obligations also lead to state responsibility. With respect to the Laws of Armed Conflict such obligations could e.g. be derived from common Article 1 GC I-IV (and corresponding customary international law), which requires States to *ensure respect* for the laws of armed conflict in all circumstances. The problem is not the lack of a legal basis but the lack of clarity as to what exactly it is that the due diligence obligation to *ensure respect* requires with regard to autonomous weapons systems. As is well known due diligence obligations typically require what a reasonable actor would do under the given circumstances. But it is hard to know what is considered reasonable when dealing with a new technology for which clear standards, practical experiences and benchmarks do not (yet) exist. Without such clarification due diligence obligations aimed at risk mitigation remain empty shells.

8. **It is therefore recommended that in addition to the clarification of Article 36 API more emphasis is put on the specification and clarification of due diligence obligations aimed at risk prevention and harm reduction.** There are various ways in which risks resulting from unpredictable robot behavior could be mitigated (e.g. by implementing automatic deactivation devices, real-time monitoring, conservative programming (“shoot second”, “double-check”). As a general rule the higher the risk, the stricter the obligation to mitigate risks. There is thus **a graduated set of risk mitigation obligations** depending on deployment scenarios, the range of tasks to be fulfilled, and the specific features of the weapons system at issue. In other words, risk mitigation obligations will be rather low when a robot is deployed in a predetermined area where no human beings are present. Conversely, if a robot were to be deployed in a complex, highly dynamic (urban) area risk mitigation obligations would be very high.

9. In accordance with general rules on state responsibility, a State is responsible for internationally wrongful acts that are attributable to it. **No particular legal challenges arise with regard to the attribution of acts committed by autonomous weapons systems.** For as long as human beings decide on the deployment of these systems, accountability can be determined on the basis of established rules on attribution. Thus, if a member of the armed forces (i.e. a state organ) of State A decides to deploy a robot on a combat mission, all activities carried out by the robot are attributable to that State. The mere fact that a weapons system has (some) autonomous capabilities does not alter this assessment.

10. The determination whether an internationally wrongful act has been committed, i.e. whether a (primary) norm has been violated by an autonomous weapons system **can be more problematic.** Some rules of international (humanitarian) law are violated whenever their objective requirements are met. In this case no particular challenges arise. Other primary rules of international (humanitarian) law, however, in order to be violated, require an element of “fault” (negligence, recklessness, intent). Which rules belong to the first or the second category may not always be clear and is in my view not fully settled. It depends on the specific requirements and interpretation of the (primary) rule in question. If the rule in question belongs to the second category, i.e. if it is a rule that requires an element of “fault” in order to be violated, it may be difficult or impossible to establish state responsibility for robotic activity.

11. The following scenario may help to illustrate the problem: A state/military commander may field a thoroughly tested and duly authorized autonomous weapons system, which – because it operates autonomously in a complex and dynamic battlefield environment – nevertheless unexpectedly violates the laws of armed conflict. There is no indication that the military commander acted with intent or negligence. And intent and negligence denote human mental states that are by definition absent in a robot. Given the complexity of these systems it may in any case be difficult to prove what exactly went wrong. As a consequence, it may be impossible to establish or prove state responsibility.

12. Conceptually, there are two principal ways how this particular accountability challenge associated with autonomous systems could be overcome. These two approaches are not mutually exclusive. Ideally they should complement each other.

First, a (future) liability regime for autonomous weapons systems could be designed so as to not require any proof of fault (“strict liability”) or reverse the burden of proof (“presumed liability”). Strict liability regimes for unpredictable, high-risk activities are not without precedent in international (see e.g. Outer Space Treaty 1967, Space Liability Convention 1972). With respect to civil uses of autonomous systems the Swedish automaker Volvo recently pledged to be “fully liable” for accidents caused by its self-driving technology. Further reflection, however, is required as to how such a liability regime could be applied in the context of an armed conflict and with regard to systems that are by definition designed to cause certain damage.

Second, instead of focusing (only) on the specific act in question the focus could be shifted to risk mitigation and harm reduction obligations (at the pre-deployment stage) and state responsibility arising from failure to abide by these obligations.