How practices make norms: autonomous and AI technologies in weapon systems and diminished human control

Dr Ingvild Bode
University of Southern Denmark

Autonomous and AI technologies, including in weapon systems, have become fundamental elements of warfare. In the Russia-Ukraine War, both sides have used one-way attack drones with the potential to function as autonomous weapons systems (AWS). Once activated autonomously, AWS can recognise, track, select and strike targets with no further human intervention.

One-way attack drones are one of a growing number of weapon systems that incorporate autonomous and AI technologies. Although frequently discussed in the context of AWS, these systems are rarely 'fully autonomous' because there is always some human decision-making involved, such as people authorising attacks. However, the quality of control that human users can exercise in interacting with such systems may be *compromised* due to the complexity of the tasks human users need to perform and the demands placed on them, such as speed and overseeing many networked systems.

Investigating the governance gap

Since August 2020, the ERC AutoNorms project has been investigating these developments and their potential consequences for international norms governing the use of force. As weapon systems integrating autonomous and AI technologies develop, so too do the consequences for human control, and a governance gap emerges. The use of all weapon systems is governed by general international bodies of law, such as international humanitarian law (IHL). However, there are no specific legally binding international regulations on autonomous and AI technologies in weapon systems, and such regulations are not on the horizon in the short term. The main international forum where the debate has taken place since 2017—the Group of Governmental Experts (GGE) on lethal AWS under the Convention on Certain Conventional Weapons (CCW)—has so far shown limited progress towards agreeing on international legal norms.

Shaping norms through practice

The AutoNorms project argues that in this situation, practices make norms. The notion of norms can refer to both legal and social norms. Legal norms are typically institutionalised in some form, either in 'soft' or 'hard' international law, for example, international treaties, resolutions or declarations. Social norms are understandings of appropriateness that can be implicit and are often not openly discussed. But such social norms also communicate what states and other actors consider 'appropriate' behaviour when it comes to using force.

AutoNorms' research has theorised how *practices* of designing, training personnel for, and using weapons integrating AI and autonomous technologies shape social norms. The salience of these developments has grown since the 2020s and is frequently referred to as the 'AI revolution'. However, integrating sensor-based targeting and sharing cognitive tasks between humans and technological systems in the process of targeting decision-making is not new. Since the 1960s, states have used various weapon systems integrating predecessor technologies such as air defence systems, guided missiles, active protection systems, counter-drone systems or loitering munitions. AutoNorms, therefore, analyses a longer historical trajectory of weapon systems that include autonomous technologies.

The AutoNorms project examines the consequences of such practices of use for human control in specific use-of-force situations as an emerging norm.

AutoNorms analytically distinguishes between two processes of norm emergence based on operational practices and public deliberation. The public-deliberative process began when AWS entered the international community's agenda in the 2010s, first at the United Nations Human Rights Council and then at the CCW. Since then, the two processes have run in parallel, but the practice-based process precedes public deliberation. States often develop and use weapons integrating new technologies years or decades before there is an international debate about the weapons' appropriateness. These long trajectories of practices shape norms behind the scenes.

AdobeStock © ReneBot/Peopleimages - AI Generated.

This does not happen in a normative vacuum. States use new weapon systems inside a normative structure, such as IHL, which limits how weapons are designed and used. This establishes general behavioural standards rather than weapon system-specific requirements. For example, IHL does not explicitly specify that weapons need to be used under human control. Many scholars understand human control as a constitutive norm of IHL located in its spirit rather than in its letter. However, its specific absence has created legal room for states to manoeuvre and decrease the quality of human involvement in use-of-force decisions.

Insights into human control dynamics

The AutoNorms project has investigated AWS and human control across both the operational and the public-deliberative processes of norm emergence. From the practice-based process, we found that practices of designing, training personnel for and operating weapon systems that integrate automated, autonomous and AI technologies have qualitatively changed the roles of human operators and users by simultaneously minimising and making them more complex. Therefore, human-machine interaction in specific use-of-force situations may become meaningless. This finding is based on an empirical deep dive into two types of weapon systems integrating automated, autonomous and AI technologies: air defence systems and loitering munitions. AutoNorms has created qualitative data catalogues and in-depth case studies based on open-source material.

Designing with autonomous and AI technologies increases system complexity. This hinders comprehension of the system's 'decision-making' from the design stage onwards. From the get-go, being able to comprehend the system, therefore, puts a great knowledge burden on human users. Practices of training personnel for operating existing weapon systems integrating automated and autonomous technologies, such as air defence systems, appear to follow common myths about autonomous systems, such as that increasing autonomous features reduces human-machine interaction and makes human operators' jobs 'easier'. The reverse is true, but training reality appears to be inadequate in light of problems of over-trust, the inclination to over-rely on automation and the tendency of automated or autonomous systems' outputs uncritically. Operating practices demonstrate that human users often do not have sufficient situational awareness as they have been relegated to passive supervisors. After all, autonomous and AI technologies execute motor, sensory and cognitive tasks. Human users may be idle until they are called to respond, switching from underload to overload in high-pressure combat situations. It remains unclear how human users can tackle this challenge when they often lack a functional understanding of the system's targeting process and the time to regain situational awareness. Therefore, the norm emerging from the operational practice-based process accepts a diminished, reduced form of human control when interacting with autonomous/AI technologies as 'normal' and 'appropriate'.

AutoNorms research has also studied how this practice-based process that shapes emergent norms on AWS interacts with the public debate at the CCW and other governance forums. We have found that the norm-shaping potential of practices of designing, training personnel for, and using weapon systems integrating autonomous/AI technologies has the potential to undercut the public processes of norm-making at the CCW and beyond. At international forums such as the CCW, states have expressed different perceptions of autonomous/AI technologies, different interests and different visions of measures to be taken at the global level. There is consensus that retaining human control over the use of force is vital, but disagreement regarding the quality of human control required, where it should be exercised, and whether we need new international law to handle the adverse consequences connected with AWS. We can group state positions into three basic groups:

i. those that support starting the negotiation of a new legally binding instrument now (e.g. Austria, Brazil, Pakistan)

ii. those that are sceptical of any new legally binding regulation and argue that current IHL is sufficient (e.g. Israel, Russia)

iii. those that favour a non-binding political declaration, code of conduct, or list of principles either because not all states are 'on board' with a new legal instrument or because such states consider a 'soft' regulatory approaches as preferable (e.g. the US, UK, Australia, Japan, Republic of Korea, Netherlands).

Against this backdrop, AutoNorms research has found three dynamics of interaction between practice-based and public-deliberative processes of norm emergence in the case of human control at the CCW. First, there has been minimal verbalised interaction between the two processes. Existing weapon systems integrating autonomous/AI technologies have rarely been discussed in the debate. Second, states have engaged in distancing, characterising AWS as a future problem and the GGE's debate as an exercise of pre-emptive norm-making. Third, in the few instances where existing weapon systems are mentioned, states affirm that practices attached to them adhere to the principle of human control. Further, such practices have been held by some states as a model to be followed. These arguments preclude a more in-depth consideration of existing systems, which are allegedly already used to meet the needs of the emerging human control norm. States appear to agree that studying existing systems integrating autonomous/AI technologies yields only 'best practices' for how human control may be properly exercised.

Implications of diminished human control

Fundamentally, existing design and use practices undermine international efforts to regulate AWS by codifying an obligation to human control. Current public-deliberative processes do not thoroughly examine the human control norm that emerges from practices performed in relation to current systems. Further, by excluding current weapons from the debate about future AWS, such practices legitimise existing systems because they are *not* AWS. Furthermore, some states and other stakeholders positively acknowledge that present practices of using weapons with autonomous and AI technologies adhere to the principle of human control. However, what is positively acknowledged here is not a high quality of direct human control in specific use-of-force situations.

As AutoNorms research has shown, a detailed examination of some existing systems reveals that direct human control at the point of use does not necessarily imply high quality human control due to the complexities of human-machine interaction. This is hardly surprising, given that human factor analysis research has been demonstrating these concerns for years. It is surprising that these research findings have such little presence in the international debate on AWS.

AutoNorms' mapping of these broad, global developments has been supplemented by our research on national practices in the context of autonomous/AI technologies in the military domain performed by various actors in China, Russia and the US—the three states frequently regarded as key developers of such technologies. To varying degrees and from various perspectives, all three are sceptical of legally regulating AWS. As a result, their positions and practices continue to have a negative impact on the CCW's capacity to serve as a negotiation platform for legally binding regulation.

Overall, the emerging norm of diminished human control represents a significant societal challenge and public policy issue since it has the potential to undercut human agency in the use of force. The proliferation of autonomous and AI technologies in the military domain, extending beyond weapon systems, gives rise to numerous political and legal ambiguities. Using AI technology in military contexts encompasses a wide range of applications, including decision support in a more expansive manner. It is probable that this will lead to an expansion in instances of human-machine interactions in the military domain. The AutoNorms project will closely observe these developments and their implications for human control in decision-making about the use of force.

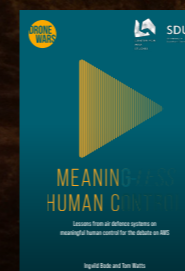## Academic publications

European Journal of International Relations
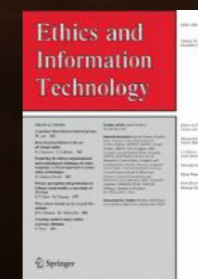
Ethics and Information Technology

Policy reports

Meaning-less Human Control

Loitering Munitions and Unpredictability