



When stigmatization does not work: over-securitization in efforts of the Campaign to Stop Killer Robots

Anzhelika Solovyeva¹ · Nik Hynek¹

Received: 22 November 2021 / Accepted: 5 December 2022 / Published online: 21 January 2023
© The Author(s) 2022

Abstract

This article reflects on securitization efforts with respect to ‘killer robots’, known more impartially as autonomous weapons systems (AWS). Our contribution focuses, theoretically and empirically, on the Campaign to Stop Killer Robots, a transnational advocacy network vigorously pushing for a pre-emptive ban on AWS. Marking exactly a decade of its activity, there is still no international regime formally banning, or even purposefully regulating, AWS. Our objective is to understand why the Campaign has not been able to advance its disarmament agenda thus far, despite all the resources, means and support at its disposal. For achieving this objective, we challenge the popular assumption that strong stigmatization is the universally best strategy towards humanitarian disarmament. We investigate the consequences of two specifics present in AWS, which set them apart from processes and successes of the campaigns to ban anti-personnel landmines, cluster munitions, and laser-blinding weapons: the complexity of AWS as a distinct weapons category, and the subsequent circumvention of its complexity through the utilization of pop-culture, namely science fiction imagery. We particularly focus on two mechanisms through which such distortion has occurred: hybridization and grafting. These provide the conceptual basis and heuristic tools to unpack the paradox of over-securitization: success in broadening the stakeholder base in relation to the first mechanism and deepening the sense of insecurity in relation to the second one does not necessarily lead to the achievement of the desired prohibitory norm. In conclusion, we ask whether it is not the time for a more epistemically-oriented expert debate with a less ambitious, lowest common denominator strategy as the preferred model of arms control for such a complex weapons category.

Keyword Autonomous weapons systems · Artificial intelligence · Automated military systems · Campaign to Stop Killer Robots · Securitization · Arms control and disarmament

1 Introduction

This article reflects on the process of securitization concerning ‘killer robots’ or, less colloquially and more impartially, autonomous weapons systems (AWS). The focus is on a coordinated response by civil society to the potential development and use of such weapons—the Campaign to Stop Killer Robots, hereinafter often referred to as the Campaign. Formed in October 2012 and publicly launched in 2013, it is modelled on other, previously successful humanitarian disarmament campaigns and has actively lobbied for ‘a

pre-emptive and comprehensive ban on the development, production, and use of fully autonomous weapons’ (CSKR 2013). The Campaign’s message has been clear from the outset: ‘Life and death decisions should not be delegated to a machine’ (CSKR 2022a). However, its agenda was defined much earlier, as marked by the formation of the International Committee for Robot Arms Control (ICRAC), a founding member of the Campaign, in 2009. The Campaign has eventually united almost two hundred non-governmental organizations (NGOs) based in various countries, gained broad public support, and successfully dragged into its agenda a good number of state governments, thousands of experts, over twenty Nobel Peace Prize laureates, as well as elements of the United Nations (UN) and the European Union (EU).

Although the Campaign and its supporters have proved successful in promoting the centrality of the issue, generating a sense of urgency, and winning the support of various actors, their collaborative effort to establish the desired

✉ Anzhelika Solovyeva
anzhelika.solovyeva@fsv.cuni.cz

Nik Hynek
hynek@fsv.cuni.cz

¹ Department of Security Studies, Faculty of Social Sciences, Charles University, Prague, Czech Republic

legal norm has not succeeded. Even after a decade of active policy advocacy, especially at a time when information and communications technologies can speed up the rate of information diffusion to an unprecedented extent, there is still no international regime formally banning, or even purposefully regulating, AWS. Neither do we record any substantial change in the Campaign's message. Our overarching objective is to understand why the Campaign has not been able to advance its disarmament agenda thus far despite all the resources, means and support at its disposal. For achieving this objective, we highlight in a succinct and disciplined manner some of the problems the Campaign and its supporters have encountered in their efforts to securitize AWS. In doing so, we challenge the popular assumption that strong stigmatization is the universally fastest and most efficient way towards humanitarian disarmament. While it has worked with landmines and cluster munitions, as well as blinding laser weapons, even preventively, with each of these weapons ban treaties concluded in much less than a decade, it will not work with AWS. AWS is perhaps the most complex weapons category ever dealt with and, as we show, this complexity, combined with the prominence of pop culture in its circumventing, lies at the very heart of the problem. It is for this reason that the paper begins by comprehensively defining AWS (Fig. 1).

On the one hand, we recognize that perhaps the only way for the Campaign to achieve a swift preventive ban would be to send a clear message to their intended audience about the possible social dangers of AWS. For that sake, they had to produce a definition of AWS that everyone could understand and stigmatize them by drawing public attention to undesirable consequences and risks associated with their development and use. On the other hand, efforts to simplify or manipulate any of the definitional aspects and characteristics of AWS for the sake of strong stigmatization, especially against the background of the cultural impact of *The Terminator*, lead to the public's distorted perception of AWS and make it easier for the other side to lightly dismiss the Campaign's call for disarmament altogether.

Here we focus on two mechanisms through which such distortion has occurred and prevented the Campaign from achieving humanitarian disarmament in the case of AWS: hybridization and grafting. These provide the conceptual basis and heuristic tools to unpack the paradox of what we call over-securitization: success in broadening the stakeholder base (hybridization) and deepening the sense of insecurity (grafting), i.e. generating a strong stigma against AWS, does not necessarily lead to the achievement of the desired legal norm, i.e. a ban on AWS. In a nutshell, our argument is that more is not necessarily better, at least in this particular case.

Our theoretical contribution is to forge an original way of thinking about the process and existing tools of securitization. Yet we do not claim a general contribution to securitization literature because our primary intention is to tailor the concept to our research needs and better grasp the case of AWS. Our contribution to the empirical literature stems from the application of securitization, as a method of understanding the logic of social and political construction of threats, to the case of AWS. However, it lies not only in developing a theoretically informed understanding but also in presenting a detailed empirical analysis of epistemic (related to knowledge production) and political (related to knowledge utilization) perspectives on 'killer robots' generated by the Campaign and its supporters. One caveat is necessary. While we may sound critical sometimes, we do not counter the argument for banning lethal machine autonomy, or question the Campaign's ethical goals. Rather, we seek to raise awareness of the problems connected with—and unintended consequences of—their efforts to promote genuine progress towards a ban on AWS. Most importantly, our findings signal to both policy makers and policy advocates, not only academics, that strong stigmatization is not necessarily the best universally applicable disarmament strategy.

2 'Killer robots': complexity, ambiguity, and problems of delimitation

Defining AWS with great precision is crucial for understanding the challenges the Campaign and its supporters have encountered in their efforts to ban this category of weapons. To accurately represent the emerging dominant discourse, we often refer to AWS by the Campaign's generally preferred term 'killer robots'. However, we point out the flaws in the perceived interchangeability of these two terms for reasons outlined below. To be more precise, this section demonstrates that 'killer robots', often cited by campaigners as the emerging category of weapons that should be banned, exist in a *space of ambiguity*. Particular attention is paid to the difference between scripted lethal autonomy, AI-equipped weapons, and lethally capable AI, the inconsistency between the Campaign's definition of 'killer robots' and the realities of AWS research, development, and production (R&D&P), as well as the gap between reality and fiction. The following sections will show how different definitional aspects and characteristics of AWS, comprehensively defined here (Fig. 1), have been distorted in the securitization efforts of the Campaign and its supporters and with what implications.

We are convinced that representatives of the Campaign, especially experts, are well aware of all the differences and nuances and we understand that their definition of AWS is purpose tailored. Opting for impartial consideration of the

technology aspect and respective pros and cons of AWS, the Campaign would lose the sense of urgency. However, strong stigmatization of such a complex, sci-fi-laden weapons category has led to its misrepresentation, as we show later in this paper.

Our starting point is the Campaign's original definition of AWS. One of the Campaign's original calls for action contains the following arguments:

The Campaign ... is a coordinated international coalition of non-governmental organizations concerned with the implications of fully autonomous weapons, also called 'killer robots'. ... [It] calls for a pre-emptive and comprehensive ban on the development, production, and use of *fully autonomous* weapons. ... It is concerned about weapons that *operate on their own without human supervision*. [It] seeks to prohibit *taking a human out-of-the-loop* with respect to targeting and attack decisions on the battlefield. ... [It] believes that humans should not delegate the responsibility of making *lethal* decisions to machines (CSKR 2013).

This definition is rather vague. It has become more sophisticated over time. For example, 'killer robots' are increasingly referred to as 'autonomous', as opposed to 'fully autonomous', weapons and their definition has become more circumstantial, which brings it much closer to reality, as we show below:

In these systems, upon activation, there is a *period of time* where the weapon system can apply force to a target without additional human approval. The specific object to be attacked, and the exact time and place of the attack, are determined by sensor processing, instead of an immediate human command. This means the human operator does not determine *specifically where, when or against what* force is applied (CSKR 2021a).

Nevertheless, the key properties of their definition have persisted. The Campaign still warns that 'machines are beginning to *replace* humans in the application of force' and raise concerns about '*handing over life and death decision making* to machines' (CSKR 2021a). The broader discourse also testifies to the fact that of particular concern are lethal autonomous weapon systems (e.g. UNODA 2017), lethal autonomous systems (e.g. Lucas 2010, p. 293), autonomous lethal technologies (e.g. Asaro 2012, p. 693), lethal autonomous weapons (e.g. Scharre 2018, Chap. 17), lethal autonomous robot weapons (e.g. Sharkey 2012, p. 790), fully autonomous weapons (e.g. HRW 2012), fully autonomous armed robots (e.g. Sharkey 2010, p. 370), and fully autonomous robotic weapons (e.g. O'Connell 2014, p. 526). The variety of these terms and definitions allows us to capture

two distinctive features of AWS: lethality and full autonomy, as enabled by advanced robotic capability.

However, it is important to differentiate AWS from other existing systems, exhibiting similar characteristics in these regards. Bode and Watts (2021, p. 6) rightly noted that 'precedents created by the decades-long use of weapons technologies with automated and autonomous features' have to be 'fully explored'. We can distinguish two dimensions of autonomy, according to which different weapons systems, existing, under development, and envisioned, can be distinguished and ordered: their degree of autonomy in the 'kill chain', and the complexity of their autonomous function (Fig. 1).¹

The 'kill chain' stands for the structure of an attack, consisting of the detection, tracking, and engagement of a target, not necessarily a human being. Here we foreground the distinction between supporting autonomous functions and (almost) full autonomy exercised by weapons themselves in completing the 'kill chain'. With respect to the latter, we also differentiate between weapons systems capable of autonomous non-lethal engagement and those capable of autonomous (non-)lethal engagement, i.e. capable of delivering lethal or both lethal and non-lethal effects (Fig. 1).

However, as already mentioned above, one more dimension of weapons autonomy is introduced to capture the complexity of the definition of AWS. It reflects the distinction between the execution of scripted autonomous performance and the capacity for autonomous 'decisions' based on situational awareness, adaptability and learning (Fig. 1). The former is considered to be a suitable term for defining weapons whose autonomous performance, even if it only appears so, is based on a preplanned script (Sharkey 2008, p. 16; 2010, p. 377). Our understanding of the term 'script' in this context goes beyond computer-based scripting, which often involves advanced sensor inputs, and covers also rudimentary forms of scripting such as purely mechanized scripts (e.g. a landmine is triggered by someone merely stepping on it) (Fig. 1). The latter dimension is enabled by advances in artificial intelligence (AI). AI paves the way for machines capable of carrying out tasks that would normally require human intelligence. Machine learning and deep learning algorithms, which may be considered as an integral part of AI, allow such machines to detect patterns in data and orient themselves in a given environment, make 'decisions' and undertake

¹ We wish to thank both anonymous reviewers whose challenges and insights inspired this operationalization. We also acknowledge that our conceptual and visual differentiation is only for heuristic purposes and does not exclude cross-sectional links (e.g. the possibility of systems moving along the spectrum based on the degree of sophistication of their software; shared or prototypical enabling technologies; etc.).

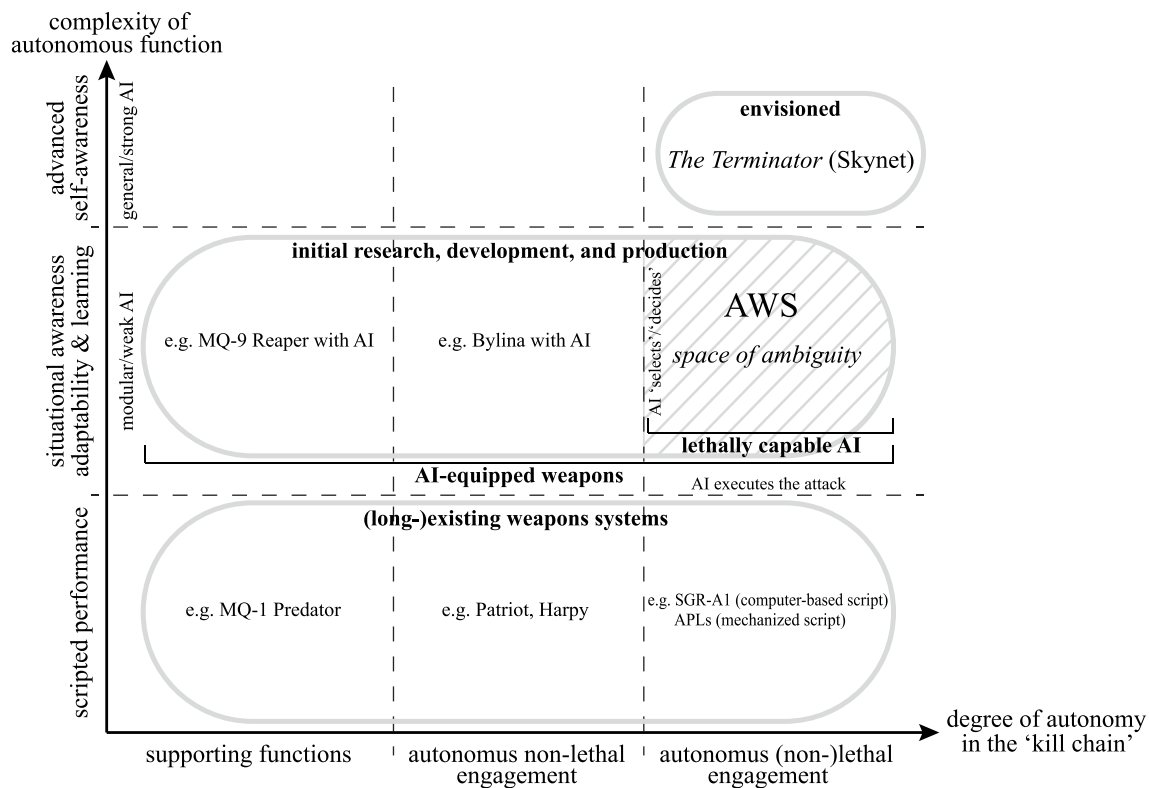


Fig. 1 The spectrum of autonomy in the definition of AWS. Authors' own figure. The figure does not aim to differentiate between different categories of weapons. It distributes them along the spectrum of autonomy. This is why one specific category may appear in several

sections (e.g. drones, i.e. MQ-1 Predator or MQ-9 Reaper) or different categories may be heuristically grouped in one (e.g. Patriot and Harpy)

tasks, as well as dynamically adjust behaviour on the basis of experience without human input (Goodfellow et al. 2016, pp. 95–96, 151; Gadiyar et al. 2019, pp. 167–169). Computer programmers do set the initial parameters of their performance but the outputs are defined by learning algorithms (McFarland 2015, pp. 1327–1329; Layton 2018, p. 7).

Three caveats are necessary. The word ‘decides’ is in quotation marks (Fig. 1) to avoid the illusion that AI algorithms are capable of independent decision-making, at least at the current stage of their development. Since computers run whatever software is installed on them, behaviour of even the most intelligent machines originates ‘not in the machines themselves, but in the minds of their developers’ (McFarland 2015, p. 1329). When it comes to the fact of AI ‘selecting’ targets, Nilsson (2010, p. 105) described the process of how an aircraft is trained to locate and identify targets in photographs: there was a ‘training sample’ of 50 images containing tanks and 50 samples of terrain not containing tanks; using these boundaries, the system was then tested on a different set of 50 images containing tanks and 50 images not containing tanks, and its performance reportedly ‘exceeded all expectations’. The word ‘selects’ is also in

quotation marks (Fig. 1). Finally, we concur with McFarland (2015, p. 1316) that the distinction between ‘automated’ and ‘autonomous’ systems is artificial in this context. We inquire into the spectrum of autonomy to fine-tune our definition of AWS.²

Our findings clearly indicate that AWS is a category of weapons that can (1) ‘select’ a human target and ‘decide’ to engage it, as well as (2) execute this attack without a human interface (Fig. 1). In this regard, AWS needs to be differentiated from simple systems that possess a (seemingly) high degree of autonomy in terms of lethal engagement such as anti-personnel landmines (APLs) or missiles that can be used—and have been used for years—in a ‘fire-and-forget’ arrangement. APLs are closer in their characteristics to AWS. First, they are considered to be ‘indiscriminate weapons that lie dormant until triggered, be it by a soldier, or a civilian, a friend or a foe, an adult or a child’ (Doswald-Beck et al. 1995). A similar argument has been put forth against AWS. Gubrud (2014, p. 35), for example,

² We wish to thank one of the anonymous reviewers for suggesting us to inquire into the spectrum of autonomy to fine-tune our definition of AWS.

argued for a ban on AWS stressing that weapons not capable of distinguishing between civilians and combatants, and between civilian objects and military objectives are ‘indiscriminate’. Second, an APL does not need to be activated by a human operator; it is triggered by someone merely stepping on it. However, an APL does not ‘decide’ to explode because it has only the most rudimentary forms of sensor inputs (it is designed to blow when triggered by pressure, meaning it follows a purely mechanized script of action) and humans choose where it is placed (Asaro 2008, p.51). Therefore, while falling into the category of systems capable of autonomous lethal engagement, even though with reservations, APLs differ fundamentally from AWS in the complexity of their autonomous function (Fig. 1). A fire-and-forget missile on an aircraft locks onto a target identified by the pilot (meaning it is neither autonomous to the same extent nor indiscriminate) and only then does it attack this target, allegedly without human involvement (Schmitt 2013, p. 5).

AWS also needs to be distinguished, along the same lines, from other robotic weapons, including (almost) fully autonomous ones and those outfitted by AI. More generally, a robot is a powered machine that (1) senses, (2) reasons (in a deliberative, non-mechanical sense), and (3) acts (Lin et al. 2008, p. 4). Existing robotic weapons include unmanned aerial, ground, underwater, etc. weapons systems (e.g. MQ-1 Predator, MQ-9 Reaper, IAI Harpy, Talon SWORDS), counter-rocket, artillery, and mortar systems (e.g. Iron Dome), missile defence systems (e.g. Patriot and Aegis), anti-aircraft systems (e.g. S-300), and close-in weapons systems (e.g. the Phalanx CIWS). Although different terms have been used to characterize these systems (e.g. ‘human-in-the-loop’ or ‘semi-autonomous’ for unmanned combat vehicles, mainly drones; ‘human-on-the loop’ or ‘supervised’ for (almost) fully autonomous defence batteries, etc.), their autonomy is no different, in principle, from that of a landmine or a ‘fire-and-forget’ missile: all these weapons systems carry out a pre-planned, more or less complex, script of action (Fig. 1). For example, unmanned weapons systems such as drones can navigate autonomously toward targets specified by GPS coordinates, which is a scripted operation (Lin et al. 2008, pp. 12, 14–15). Defence batteries also perform ‘pre-programmed’ (meaning scripted) actions within ‘tightly set’ parameters, even in their autonomous mode, and, being ‘stationary’ in addition to this, operate in ‘comparably controlled’ environments (Altmann and Sauer 2017, p.118). The principle of scriptness differentiates them from AWS (Fig. 1).

At the same time, the script performed by these robotic weapons differs in terms of their autonomous engagement in the ‘kill chain’, and this is another dimension of the principal difference between most of the existing systems and AWS (Fig. 1). While unmanned systems such as drones that have

been deployed by the moment of writing this paper may have some autonomy, mainly in navigation, a human operator is still remotely watching a computer screen and making the final decision on when and what to fire upon (Johnson and Axinn 2013, p. 130). For this reason, such systems are identified as performing only supporting autonomous functions in the ‘kill chain’ in which key decisions still have to be made by a human operator.

Israel’s Harpy is one of the most advanced unmanned weapons which is considered to be an exception and even a precursor to AWS (HRW 2012). It is a suicide drone which can select a target based on radar signals and engage with it thereafter (Horowitz 2016, p. 91). It can loiter for hours before detecting, locking onto, and destroying its target. However, the parameters of its autonomous mode are pre-determined and humans decide on its target area (Vallor 2016, p. 212; Brenneke 2018, p. 65). At the same time, it is designed for use against hostile radars rather than against humans (Finn and Scheduling 2010, p. 178). The above indicates that this system falls within the range of scripted weapons but its autonomy in the ‘kill chain’ is greater than otherwise expected of drones: it is supposedly capable of autonomous non-lethal engagement (Fig. 1). For the very same reasons, other (almost) fully autonomous systems, particularly defence batteries (e.g. Patriot), are excluded from our definition of AWS. Not only are they stationary and fixed in their parameters (meaning their performance is scripted too), but they are also designed for defence against inanimate targets (Altmann and Sauer 2017, p. 118). It means their main purpose consists of autonomous non-lethal engagement (Fig. 1). The situation is slightly different with anti-aircraft systems. If employed against manned aircraft, they are potentially capable of autonomous lethal engagement, besides autonomous non-lethal capabilities (Fig. 1). For example, an advanced Russian S-300 air defense system engaged Israeli fighter jets in Syria. Reportedly, no one died in the attack (*The Times of Israel* 2022). There have also been reports based on limited and unconfirmed evidence that Russia’s Su-35 was destroyed by a Ukrainian-owned S-300 (Axe 2022). However, not only are such cases rare and ambiguous in terms of confirmed lethal effects, but anti-aircraft systems also have similar limitations to those of (almost) fully autonomous counter-rocket and anti-missile systems. All these weapons systems, at least for now, are locked into performing a pre-planned script of action defined by fixed programmed procedures, and humans decide where to deploy them, when to activate their autonomous mode, and can override their operation at any time (Walsh 2015a, p. 2; Horowitz 2016, pp. 89–90).

The most significant of all the precursors to AWS is the SGR-A1 (PAX 2021). It is a stationary weapon system designed to guard the demilitarized zone between North and South Korea. This system is distinct in that it classifies

human beings detected in this zone as targets, meaning it is capable of autonomous lethal engagement (Fig. 1). Some sources claim its software is capable of ‘pattern recognition’, allowing it to distinguish humans from animals or other objects, and ‘voice recognition’, supposedly allowing it to distinguish between friends and foes through the provision of a proper ‘access code’ (Kumagai 2007; Etzioni 2018, p. 260). However, these capabilities are rather primitive yet. The SGR-A1 reportedly uses movement detectors and thermal imaging to lock on to human-sized targets (Johnson and Axinn 2013, pp. 137–138). Most importantly, the SGR-A1 is placed in a controlled environment to which human access is ‘categorically prohibited’ (Tamburrini 2016, p. 126). Therefore, its performance is scripted and its script is constrained not so much by its design as by the environment in which it operates, with ‘only legitimate targets present’ (Arkin 2009a, pp. 167, 171). In addition to this, its autonomous ‘mode’ is ‘optional’ (Kallenborn 2021) and humans supervise its operation via camera links (Wakefield 2018).

AI is being actively adapted for further sophistication of autonomous functions in weapons systems but the very fact of this happening does not make the so-called ‘AI arms race’ a race for AWS (Fig. 1). Even though AI, itself not a weapon, can be adapted for military purposes and even integrated into weapons systems, it does not necessarily power the actual application of force. For example, the US F-35’s advanced sensor fusion algorithms will acquire, distill, and organize otherwise disparate pieces of intelligence into a single integrated picture for the pilot (Osborn 2017). Another example is the US MQ-9 Reaper. Its sensing capabilities intended to assist the human operator will be enhanced with object recognition algorithms based on AI (Defense Post 2020). China’s PLAN nuclear submarines will be equipped with decision support systems to enhance commanding officers’ thinking skills and reduce their workload and mental burden (Chen 2018; Kania 2018). These examples illustrate that supporting functions performed by weapons systems themselves in the ‘kill chain’ are being elevated to the next level of complexity with the help of AI. But this does not necessarily change their role in the ‘kill chain’ (Fig. 1).

Even in cases where AI is designed to power the application of force, its function may not necessarily be the application of lethal force. This is demonstrated by the sophistication of autonomous non-lethal engagement capabilities in weapons systems (Fig. 1). The Russian Aerospace Forces have, for instance, tested an automated control system with elements of AI. It will combine air defence systems (S-300 s, S-400 s, Pantsirs) and early warning radars into a single ‘armoured fist’, perform realtime situation analysis, and issue recommendations for the use of weapons (Kruglov et al. 2018). AI algorithms are also being adapted to fully autonomize the most complex processes of electronic warfare. For example, Russia’s brand-new electronic warfare

system Bylina will be able to establish communications with electronic warfare stations, higher headquarters, and command posts without human intervention. Among its autonomous functions will also be to analyse the situation in real time, find and recognize different sorts of targets (e.g. enemy radio stations, communication systems, radars, early warning aircraft, satellites), choose the best means to suppress them, give orders to individual electronic warfare stations, and control their operation (Ramm et al. 2017). Therefore, the category of AI-equipped weapons is much broader than it may seem in the context of an extensive international debate about lethally capable AI (Fig. 1).

Now we come to the most important part of our discussion and focus on the most advanced kind of AI-equipped weapons—lethally capable AI. In addition to the above, AI also paves the way for weapons systems that can supposedly ‘select’ a human target, ‘decide’ to engage it, and execute this attack without a human interface (Fig. 1). Perhaps the best example is the autonomous fire control module presented by the Kalashnikov Group. Keller (2017) called it ‘a real life Terminator’. This module will be able to recognize, illuminate, and track targets and will be compatible with all combat modules produced by the Kalashnikov Concern. Whenever installed onboard, it will be a valuable asset to the human operator. However, it can also be switched to an autonomous mode in which it will reportedly scan the operational space, detect hostile objects, distinguish between humans and machines, determine priorities in the sequence of defeat, decide on the required number of shots for guaranteed destruction of each, and open fire. If unarmed, and therefore harmless, people—or civilians—appear in the operational space, the module will supposedly steer the fire aside. One caveat, as we also discuss below, is that such reports should be treated with caution. Pattern recognition in complex contexts is still a challenge for software engineers, at least for now, when it comes, for example, to distinguishing between a man carrying an AK-47 and a man carrying a walking stick; between a non-combatant carrying an AK-47 and a combatant carrying the same weapon; between combatants and non-combatants in situations when insurgents pose as civilians; or between active combatants and wounded ones who are unable to fight or those who have surrendered (Lin et al. 2008, p. 76; Kastan 2013, p. 60). Since artificial neural networks modeled on the human brain will reportedly be used to structure the software of this fire control module, it will also be able to learn in the process of operation. Kalashnikov Media, a media platform that reports on the whole variety of products and services of the Kalashnikov Group, has released a video that presents the module’s autonomous mode behaviour (Kalashnikov Media 2018; TASS 2018). Another example is the Turkish-built attack drone called STM Kargu-2. There have been numerous reports

on the March 2020 accident³ in which this drone supposedly hunted down a human target without being instructed to do so, while operating in an autonomous mode that required no human controller (Froelich 2021; Hambling 2021; Zitser 2021). ‘The news raises the spectrum of [T]erminator-style AI weapons killing on the battlefield without any human control,’ as subsequently reported by Moran (2021).

At the same time, the development of humanoid robots equipped with AI has been gaining momentum too as illustrated by Boston Dynamics’ Atlas and the Russian-made robot nicknamed Fedor. RIA Novosti (2017) even reported, citing Dmitry Rogozin, that the latter ‘learned to shoot with two hands’.

However, the above does not explicitly testify to the existence of fully autonomous, even less so Terminator-style, killing machines. The video released by Kalashnikov Media demonstrates that their new autonomous fire control module will be *activated* by a human and, assuming this report is accurate, our interpretation of the fact is that its deployment will be a tactical decision taken under battlefield conditions (Kalashnikov Media 2018; TASS 2018). Turkish drone maker STM Defence denies the Kargu-2’s autonomous strike capability and claims that this drone ‘keeps a human in the loop during attacks on targets’ (Özberk 2021). The video released by STM Defence on their Youtube channel also demonstrates that this drone will supposedly operate in a ‘customizable’, which we understand as a pre-selected, detonation range (STM 2020). One caveat, as we discuss in more detail below, is that these reports should be treated with caution. Russia’s Fedor is a space robot built to assist space station astronauts, not a weapons platform. Teaching Fedor to shoot with both hands was reportedly meant to improve its motor skills and decision-making abilities (Grishchenko 2017). Furthermore, this robot did not live up to its task and will be used as a platform for testing new technologies on Earth and developing another Russian anthropomorphic robot and the successor to Fedor for outer space (Apazidi 2022).

It is important to remember, however, that the development of the most advanced AI-equipped weapons, especially lethally capable AI, takes place in conditions of the utmost secrecy. Therefore, we argue that AWS exists, even at the level of reporting and R&D&P, in a space of ambiguity (Fig. 1). Assuming all of the above is true, we can concur with the roboticist Ronald Arkin (2009b, p. 32), arguing that the highest degree of weapons autonomy that is being developed for autonomous lethal engagement is still ‘bounded’ and applicable ‘for very narrow tactical situations’. It is not about ‘replacing a human soldier one-for-one’, Arkin

(2009b, p. 32) added. This challenges the Campaign’s argument on lethal weapons operating without human supervision and points out the gap between the Campaign’s definition of ‘killer robots’ and the realities of AWS R&D&P. This gap is utilized by the other side, represented mainly by the world leaders in the development of related technologies, for bypassing normative pressure and maintaining flexibility in R&D&P. For example, in 2018, Russia proposed to define fully autonomous weapons as weapons ‘designed to carry out combat and support tasks without *any* participation of an operator’ (Country Statement 2018a). However, our understanding that some of the above information may be misleading or incomplete leads us to consider two manifestations of strategic ambiguity: the capabilities of such weapons may be downplayed by their manufacturers for normative reasons; or, on the contrary, they may be exaggerated for business reasons. The former may have been the case with the aforementioned lethal attack by the Kargu-2. The UN Panel of Experts on Libya (2021) indicated that ‘the lethal autonomous weapons systems such as the STM Kargu-2 ... were programmed to attack targets *without requiring data connectivity* between the operator and the munition’. What needs to be kept in mind, however, is that it is difficult to retrospectively verify the actual mode of engagement.

Noteworthy is the fact that normative pressure on the market may not necessarily be strong in the long run because AWS are arguably neutral in terms of their military effect. Some argue that the delegation of life and death decisions to machines is unethical, immoral, and should be made illegal (Asaro 2012, p. 708). Others, even some roboticists, assume that AWS might in fact be better than humans in satisfying ethical codes and legal principles (Schmitt 2013; Arkin 2009a, b, 2018). This is because AWS may, as some assume, be more precise and accurate in their targeting than any existing weapons and can potentially reduce unnecessary casualties and prevent unwarranted injuries (Wagner 2014, p. 1411; Birnbacher 2016, p. 119). We recognize that algorithms are always biased and contextually embedded, as actively researched and convincingly testified to within STS. However, as McFarland (2015, pp. 1328–1329) convincingly argued in his piece on AWS, the development of ‘intelligent [enclosed in quotation marks in the original source too] machines is ‘in fact just an exercise in software development’ and that their subsequent behaviour ‘originates not in the machines themselves, but in the minds of their developers’. All of this adds to the gap between the Campaign’s normatively-oriented, preventive ban-motivated representation of ‘killer robots’ and the dynamics of AWS R&D&P.

This ambiguity, combined with the Campaign’s terrifying image of ‘killer robots’, gives rise to delusional fantasies about Terminator-like killer robots and misrepresentations by mass media, as even illustrated above. The following

³ We wish to thank one of the anonymous reviewers for drawing our attention to this accident.

message appeared, for example, in an article published by *The Guardian*:

As Ray Kurzweil speaks to the Observer New Review about the impending advances in artificial intelligence, it seems a good time to heed the warning of such screen classics as *Alien*, *The Terminator* and *Blade Runner* and look back at the rogue computers, robots and replicants that have brought death, disquiet and destruction to humankind. Enjoy, before it's too late (Whitmore 2014).

Such references do not appear so odd to the general public, especially as humanoid robots are already being equipped with AI. However, the dividing line between fiction and reality should be drawn, even if it is increasingly blurred. The notion of self-awareness lies at the heart of this distinction. According to the plot of the movie, the program that ran the Terminator (Skynet) achieves self-awareness and decides to destroy humanity. The concept of self-awareness is not an all-or-nothing phenomenon but a spectrum ranging from the simplest forms (stimulus-awareness) to the most complex ones (meta-self-awareness). While an AI-equipped weapon can be stimulus-aware, interaction-aware and possibly even time-aware (any combination of which implies a certain degree of situational awareness), it is difficult to imagine, at least for the moment, that it will be goal-aware (in the sense of being able to reason about its goals) and meta-self-aware (in the sense of having a clear understanding of its own self-awareness) (Lewis 2014, pp. 275–278). The most complex forms of self-awareness embedded in machines, including those featured in *The Terminator*, are often associated with general or strong AI which is a distant prospect (Ayoub and Payne 2016, p. 812). Bhuta et al. (2014, p. 263) also discussed the possibility of ‘the choice of the enemy [not the target]’ falling upon weapons themselves as a scenario that originated in science fiction. AI-equipped weapons discussed above in this paper are the examples of modular or weak, i.e. domain- and problem-specific, AI (Ayoub and Payne 2016, p. 795). This is the line between fiction and reality yet (Fig. 1).

The above lines lead us to five major findings. First of all, these are not only AI-equipped weapons that are capable of lethal action without a human being directly involved in the initiation and execution of this attack, as illustrated by the discussion on APLs and ‘fire-and-forget’ missiles, as well as the SGR-A1. The only difference is that AI-equipped weapons can themselves ‘decide’ to do so (i.e. ‘select’ their targets), with the word ‘decide’ and the word ‘select’ both being in brackets for three reasons. First, according to the existing open-source data, a human operator will remain in or at least on the loop. Second, AI will be trained to select its targets. Third, we are still far from highly advanced forms of computational self-awareness, which makes us

rather sceptical about the future in which independent decisions will be taken and independent choices will be made by general or strong AI. Another key finding is that there is a difference between AI-equipped weapons and AWS. AI may fulfill mere supporting functions and does not necessarily decide on the use of force. Even if it does, it does not necessarily engage in lethal decision-making. These two findings point at a much more complex relationship between AI and AWS than it is often assumed. One more finding of great significance is that AWS exist in a space of ambiguity as there is a high degree of ambiguity regarding their military effect and respective R&D&P. This leads us to our next major finding, that is the gap between the Campaign’s partisan, preventive ban-motivated representation of ‘killer robots’, further amplified and simplified through Terminator-inspired fantasies of ‘killer robots’ spread by mass media, and the much more complex realities of AWS R&D&P. Our last and closely related point concerns the maintenance of the boundary between fiction and reality.

3 The paradox of over-securitization

We identified the apparent contradiction between the existing ‘killer robots’ discourse and the much more complex realities of military R&D&P. Respective securitization efforts by the Campaign and its supporters add an additional layer of problems that ultimately direct them away from their desired goal. We seek to understand their continued lack of political success through the lens of securitization theory and, in particular, what we call over-securitization. Through this prism, we unpack the problem of strong stigmatization with respect to such a complex weapons category as AWS, additionally embedded in pop culture and often associated with *The Terminator*.

The concept of securitization and the original logic of the process was introduced by the Copenhagen School, mainly Barry Buzan and Ole Wæver. Their intention was to move away from the state as the central referent object in all security sectors towards a multisectoral approach to security allowing referent objects other than the state into the picture, as well as to question the primacy of the military element in the definition of security and broaden it to other possible referent objects such as the individual, the international community, the environment, the economy, etc. (Buzan et al. 1998, pp. 1, 8). The same authors (1998, pp. 23–24) defined securitization as a process of when an issue is ‘presented as an existential threat, requiring emergency measures’. They (1998, p. 30) also argued that securitization is an ‘intersubjective process’ because ‘[i]t is not easy to judge the securitization of an issue against some measure of whether that issue is “really” a threat; doing so would

demand an objective measure of security that no security theory has yet provided’.

We go beyond to inquire into the process of over-securitization. The Copenhagen School admitted the possibility of under- and over-securitization (Buzan et al. 1998, p. 30). The former was, inter alia, associated with ‘political choice’. In their view (1998, p. 86), ‘actors might choose to ignore major causes for political or pragmatic reasons and therefore may form a security constellation that is different from what one would expect based on one’s knowledge of effects and causes’. Over-securitization was, according to the same authors (1998, p. 211), linked to the traditionalists’ ‘objectivist, externally determined’ definition of security, with too much focus on one sector (the military) and one actor (the state). We use these two definitions as the starting point for our argument. We generally agree with Buzan et al. (1998, p. 211) that it is a ‘choice’ to phrase certain things in security terms, not an ‘objective’ feature of the issue in question. However, we comprehensively illustrate throughout this paper that there may be a gap between two different inter-subjective structures: what we call an epistemically-oriented expert debate on a given security issue, moving us closer to the supposedly objective understanding of the issue (not necessarily a scientific consensus), and normatively-oriented principled understandings of the same. Therefore, we define over-securitization in a different way, reversing the Copenhagen School’s definition of under-securitization: actors might choose to manipulate knowledge about the problem for it to appear far more urgent and far more important than one would expect within an expert debate. This is why we devote particular attention to the blurring line between political activists and experts, with the latter often choosing to foreground certain scientific facts and omit others. However, the process is not necessarily intentional. Further contributing to the existing literature, we identify and theorize two mechanisms facilitating over-securitization in the studied case: hybridization and grafting. Both mechanisms, as we show, contribute to the strong stigmatization of AWS but both have counterproductive effects when such a complex, sci-fi-laden weapons category is concerned. The former captures the following argument in a nutshell: the more actors are involved in policy advocacy and the more scientific impartiality is compromised, even if for good cause, the higher the probability that knowledge may become fragmented, inconsistent, and emotionally saturated, thus leaving room for interpretation and manipulation. The latter captures a prominent technique of knowledge manipulation through which a given security issue is grafted onto other security issues of immediate importance, legal precedents and even science fiction imagery, often on a selective basis, to highlight multifaceted security risks and the urgency of action. However, the more

issues are brought on board, the higher the probability of reductionist oversimplifications and loss of focus. Here lies the paradox of over-securitization or strong, strictly one-sided and sci-fi-laden stigmatization of a complex weapons category: success in broadening the stakeholder base and deepening the sense of insecurity does not necessarily mean the success of securitization.

Our theorization of both mechanisms builds extensively on the existing literature, as the following two sections demonstrate, but goes beyond. The existing literature approaches the problem of over-securitization from the perspective of referent objects of security. Hammerstad (2008, pp. 1–2) was the first to point out that a security issue can ‘become over-securitised to the point where it is in danger of creating threats [to the referent object] where before there were none’. Ihlamur-Öner (2019, p. 210) concurred: ‘The securitization of irregular and forced migration has reached to the point that it can be described as over-securitization, which creates more threats where there were none while putting the lives of migrants and refugee protection at risk’. We take a different approach and explore the dynamics of over-securitization from the perspective of securitizing actors and securitizing moves. The Copenhagen School originally defined a securitizing move as a specific rhetorical structure or discourse—or, more precisely, a ‘speech act’—that frames an issue as an existential threat, i.e. a security issue, while a person or a group that performs such a move as a securitizing actor (Buzan et al. 1998, pp. 25–26, 40). Securitization literature has been further developing ever since. The general tendency in such literature has been the redefinition and broadening of our understanding of the identity of securitizing actors and the means of securitization. Stritzel (2012, p. 553) summarized the general tendency as the development away from static understandings of the authority to securitize and single speech acts to more complex processes of authorization and more dynamic representations of existential threat. It is where we depart from the Copenhagen School and where our theorization of hybridization and grafting, respectively, takes root.

3.1 Hybridization: the diffusion of authority and knowledge

Only a few actors and groups do have ‘the power to define security’, according to the Copenhagen School. Among them are political leaders, governments, bureaucracies, lobbyists, and pressure groups (Buzan et al. 1998, pp. 31, 40). Wæver (1995, p. 57) insisted specifically, however, that ‘security is articulated only from a specific place, in an institutional voice, [typically] by elites’. Buzan et al. (1998, p. 21) reaffirmed the principle in their joint book: ‘Traditionally, by saying “security”, a *state representative* declares an emergency condition’.

Other, especially more recent, supplements to securitization theory contributed to the development of a more sophisticated set of assumptions regarding the authority of securitizing actors. Foreign politicians can, in turn, either provide or withhold external legitimation for one's securitization efforts, according to Floyd (2020, p. 10). The Paris School, mainly represented by Didier Bigo and Thierry Balzacq, put particular emphasis on expert security knowledge (Bigo 2006). Security professionals and security agencies, the ones who routinely collect and analyse data, were recognized as having the authority to determine what exactly constitutes security (Bigo 2000, p. 176). Attention was also drawn to bureaucracies that serve as an 'intermediary' with the central government and are directly involved in the provision of security services (e.g. military and police services, border guards and customs agents, intelligence services, risk assessment experts, etc.) (Bigo 2006). Bigo (2002, p. 83) specifically highlighted that, even if NGOs intervene, 'they can do so only by turning professional'. Berling (2011, p. 386) also argued that science co-determines the status of the securitizing actor. She (2011, p. 392) particularly assumed that 'scientific capital' co-determines 'the *hierarchy* in the field of security and the chances of winning'. Brauch (2009, p. 94) noted the significance of scientific 'reputation'. Floyd (2020, p. 10) also stressed that media outlets can prioritize certain issues over others, decide how information is relayed, and, therefore, control what becomes public knowledge. Vultee (2011, pp. 77–93) showed practically how the media 'speak security'. Members employed in the relevant industry can equally facilitate or impede securitization by presenting reasoned arguments for one side or the other (Floyd 2020, p. 11).

There has, at the same time, been greater awareness that the audience can reinforce the authority of the securitizing actor. The Copenhagen School provided initial instruction on how to assess the role the audience plays in constructing insecurity. Success, in their view, depends on the audience being convinced that the issue is an existential security threat. The issue is securitized only if and when emergency measures that go beyond standard political procedures are accepted as justified (Buzan et al. 1998, pp. 23–25). Balzacq (2005, pp. 171–172) carved out a more central role for the audience, focusing on 'the *power* that both speaker and listener bring to the interaction'. Salter (2008, pp. 321–322) conceptualized interactions between the securitizing actor and the audience as 'iterative'. He studied the process of 'audience-speaker *co-constitution* of authority and knowledge'. McInnes and Rushton (2011, p. 117) even introduced the idea that original audiences can, at some points themselves, act as securitizing actors.

Having considered how securitization theory has broadened with respect to theorizing the identity of securitizing actors, we found out that state representatives and bureaucrats, military personnel and scientists, policy advocates and NGOs, themselves collecting, distributing, and efficiently utilizing professional knowledge, the industry, mass media, and even the general public (e.g. through partaking in surveys and voicing their concerns) can actively contribute to securitization efforts. Therefore, complex hybridization of securitizing actors and target audiences takes place. Salter (2012, p. 934) reminds us of the fact that 'securitization is a constant process of struggle and contestation'. In accord with his interpretation (2012, p. 931), the so-called securitizing move consists of 'overlapping ... language security games performed by varying relevant actors'. However, the problem with viewing the securitizing 'actor' as a hybrid construction with many different voices is in underestimating that what results are circulatory, transepistemic, and post-truth configurations of security. The very fact that scientists become willingly and directly involved in policy making has created an open window of opportunity for different types of expert knowledge utilization. On the one hand, it has led to a better comprehension of complex policy issues, as other literature suggests (e.g. see the volume edited by Haas 1992). However, as we show, it can also lead to political appropriation of science, subsequent erosion of its credibility and original purpose. When making this argument, we are inspired by Aradau and Huysmans (2018). They accurately determined that 'transepistemic relations create greater symmetry between various knowledges and dilute the superior authority of science in truth telling and factual knowledge about the world' (Aradau and Huysmans 2018, p. 49). The same authors (2018, p. 54) defined the condition of post-truth as 'a less hierarchical and more horizontal transversal practice of knowledge creation and circulation'. Sismondo (2017, p. 3) admitted that the field of Science and Technology Studies (STS) also suggests 'the emergence of a post-truth era might be more possible than most people would imagine' and occasionally refers to the process of 'epistemic democratization' in this light, recording the role of social media platforms such as Twitter in 'the dissolution of the modern fact'. We have a practical illustration of epistemic democratization and show that, paradoxically, the diffusion of authority and knowledge decreases the likelihood of successful securitization, at least under certain conditions that we identify. These are the complexity of AWS as a distinct category of weapons and the role of pop culture in the securitization process. This is because it becomes difficult to achieve consistency and precision in argument, although one may initially assume that the more actors are involved in spreading the message the better.

3.2 Grafting: a tug between simplicity and excessive complexity

Our theorization of another mechanism of over-securitization, i.e. that of grafting, is inspired by the broadening understanding of the means of securitization. The conceptualization of the securitizing move itself extends beyond the single speech act, as claimed by the Paris School. Bigo (2002, pp. 65–66) stressed the importance of bureaucratic practices performed by security professionals and involved in the creation of administrative knowledge (e.g. population profiling, risk assessment, statistical calculation, category creation, proactive preparation, etc.). Therefore, practical work and expertise are certainly no less important than the discourse (Bigo 2000, p. 194). Here we concentrate less on bureaucratic practices and more on different discursive frames in play. However, we still note, for example, that the Campaign has also drawn quite heavily upon surveys done by the market research company Ipsos. The most significant observation that comes out of it, however, is different: the Paris School broadened the definition of the securitizing move beyond the speech act, as originally maintained by the Copenhagen School. Balzacq (2005, p. 191) also focused on the ‘manner’ in which the securitizing actor makes the case for the point and drew attention to two basic principles ensuring ultimate success: ‘emotional intensity’ and ‘logical rigor’. Both have been explored by the Campaign and we deal with specific arguments in more detail elsewhere (Solovyeva and Hynek 2018). Balzacq (2005, pp. 172, 179) also reminded us of the role of analogies, metaphors, and stereotypes as effective tools of persuasion. This is what serves as the basis for our analysis of the role of *The Terminator*. The increasingly blurred line between fact and fiction is of particular significance to our understanding of the terrifying image of ‘killer robots’.

Numerous studies have sought to develop a more nuanced understanding of the core of securitizing moves. First of all, there has been a general awareness that scientific data and facts ‘can be mobilized strategically’ (Berling 2011, p. 393). We illustrate it here by showing how scientists and researchers align with the Campaign. Stritzel (2012, p. 560) explored the link between power politics and pop culture as ‘a principal background of meaning’. In particular, he inquired into the use of pop culture and cultural myths by securitizing actors. We are going to do the same. Williams (2003, pp. 526–527) suggested that images and other visual representations are also part of ‘a broader performative act’ and do play a significant role in the process of securitization. We show it here too, analysing the role of images from *The Terminator*.

This article shows how scientific facts and cultural imaginaries, visual and discursive frames are all mobilized to construct the threat of so-called ‘killer robots’.

Some scholars have even sought to highlight the supposed links between different security agendas such as the ‘migration-terrorism nexus’ (Ihlamur-Öner 2019), the ‘terrorism-asylum nexus’, or even the ‘terrorism-immigration-asylum nexus’ (Tsoukala 2006, pp. 612, 618). This is an important, yet undeveloped, argument. It is a good starting point for us to properly conceptualize processes of grafting involved in securitizing moves. The concept of ‘grafting [a new norm onto existing norms]’ stands for a well-established legal practice and is borrowed from Price (1998). He (1998, pp. 628–629) defined grafting as ‘the mix of genealogical heritage and conscious manipulation involved in ... normative rooting and branching’. What is of particular importance, according to him, is how well a new norm ‘resonates’ with already established norms. For example, he argued that the effort to delegitimize APLs was grafted onto a viable chemical weapons taboo, the laws of war, and IHL.

With all of the above in mind, the concept of grafting helps us conceptualize how a new security agenda can be grafted, through both discursive and visual representations, themselves drawing on strategically mobilized evidence and saturated with emotional meanings and influences, onto other security issues of immediate importance, legal precedents and even science fiction imagery. The latter stands out as an example of what we call *diagonal grafting* because it involves inter-field grafting (i.e. weapons law on pop culture), rather than normative or legal grafting across different issue areas within the same domain (disarmament). Such grafting techniques may reinforce the sense of insecurity but, paradoxically, can in some cases impede instead of facilitate the securitization process. They are especially problematic in the case of AWS. On the one hand, the very ‘killer robots’ frame, often accompanied by images from *The Terminator* and references to weapons of mass destruction (WMD) (Carpenter 2016, p. 53; FLI 2022a, 2022c), was intended to convey ‘a simple and dramatic message’ (Rosert and Sauer 2021, p. 21). On the other hand, the imprecise focus and reductionist paths, both unintended consequences of deepening the sense of insecurity (grafting), make it more difficult to name the threat clearly, which is a precondition for successful securitization. Salter (2012, pp. 938–940) gave an example of what it means if ‘the threat remains *vague*’. We show it here too.

This paper relies on discourse analysis in its examination of hybridization and grafting, hence over-securitization respectively. Selected texts are segmented into discursive structures, i.e. particular statements of the Campaign and other actors actively contributing to its cause. We particularly search for similar statements to identify actors participating in the production and promotion of the ‘killer robots’ discourse, the key themes holding this discourse together, as well as contradictory statements revealing inconsistencies in this discourse. Visual discourse analysis where images are

treated as arguments is part of it. Our corpus of data is not exhaustive. The range of materials engaging with the topic of ‘killer robots’ and in one way or another contributing to the Campaign’s cause is vast and continually growing. We select representative statements of the Campaign itself and a wide range of actors involved directly or indirectly in the securitization process (experts, bureaucrats, mass media, etc.) and believe this approach is appropriate for a given task, which is to illustrate the multiplicity of voices (hybridization), as well as scattered efforts to forge novel legitimizing links between different security issues and social concerns (grafting). The concept of hybridization is applied first because it helps us to identify the actors involved. Only then is attention drawn to grafting processes and data selection at this stage is informed by the identification of relevant actors at an earlier stage. The core of our attention falls to subsequent contradictions between discursive structures, as well as between discourse and practice to demonstrate the dynamics of over-securitization.

4 Hybridization of expertise: the birth of ‘scientific’ policy advocacy

The Campaign to Stop Killer Robots has thus far spearheaded much of the securitization effort in the issue area of AWS. However, over the last decade, it has de facto grown into a truly global coalition of international, regional, and national NGOs, technology companies, expert communities, governments, administrative officers, and international organizations (IOs). Besides different kinds of actors involved proactively, there are numerous examples of the successful persuasion of audiences and, most importantly, the transformation of the original audience into an even broader coalition of like-minded securitizing actors. Even if not officially affiliated with the Campaign, more and more actors contribute to its cause at least by aligning with the Campaign’s message and goals. This is how the movement transforms into a hybrid construction and how the Campaign loses control of the ‘killer robots’ discourse, itself increasingly losing focus and transforming into an inconsistent body of knowledge composed of circulatory, disjointed, transeptic, and post-truth narratives.

What constitutes the core of the problem is the fact that the line between political activists and experts has become ever more blurred and difficult to define. What we are increasingly observing over the last two or three decades is the birth of ‘scientific policy advocacy’ (Hynek and Chandler 2013). In the studied case, academic researchers and technical experts have, in practice and principle, aligned with the Campaign and contributed to its cause by mobilizing their authority, knowledge, and scientific reputation. AI and robotics researchers wrote an open letter in 2015.

This initiative called for ‘a ban on offensive autonomous weapons beyond meaningful human control’. The letter portrays autonomous weapons as ‘ideal for tasks such as assassinations, destabilizing nations, subduing populations and selectively killing a particular ethnic group’. It has, to date, been signed by 4,502 researchers (FLI 2015). The global health community launched a similar initiative to ‘call for an international ban on lethal autonomous weapons’. Their open letter highlighted that ‘lethal autonomous weapons can fall into the hands of terrorists and despots, lower the barriers to armed conflict, and become weapons of mass destruction enabling very few to kill very many’. It has so far been signed by 90 health professionals (FLI 2022a). In 2017, the leaders of over a hundred AI and robotics companies signed another open letter urging the UN ‘to prevent an arms race in these weapons’. They called attention to the fact that lethal autonomous weapons ‘can be weapons of terror, weapons that despots and terrorists use against innocent populations, and weapons hacked to behave in undesirable ways’. The letter was signed by Tesla’s Elon Musk and DeepMind’s Demis Hassabis and Mustafa Suleyman, among others (FLI 2017a). As of the time of writing, the military personnel’s collective letter calling for ‘a ban on the development, use and deployment of autonomous weapons’ is also being prepared (CSKR 2022c). The gap between political activists and experts is closing even faster as NGOs themselves engage in collecting and distributing professional knowledge, besides mobilizing their social capital. Human Rights Watch reviewed the precursors to fully autonomous weapons and presented a sound legal analysis to justify the call for preventive action in its 2012 report called ‘Losing Humanity: The Case Against Killer Robots’ (HRW 2012). In 2019, PAX released a research report titled ‘Slippery Slope: The Arms Industry and Increasingly Autonomous Weapons’. The report provided an overview of recent developments in unmanned technologies and applications of AI (PAX 2019).

The relevant industries have also contributed, and very heavily, to the Campaign. The initiative was supported by many companies and individuals working in the field of robotics and AI: since 2018, 247 organizations and 3253 individuals pledged to ‘neither participate in nor support the development, manufacture, trade, or use of lethal autonomous weapons’. In their pledge, lethal autonomous weapons are described, inter alia, as ‘powerful instruments of violence and oppression’. The pledge calls upon governments and government leaders ‘to create a future with strong international norms, regulations and laws against lethal autonomous weapons’. Among the signatories are Google DeepMind, Clearpath Robotics, Silicon Valley Robotics, GoodAI, and TeslaVision Corporation (FLI 2022b).

There is also an emerging like-minded bureaucratic coalition spreading at both the regional and global levels. In 2013, the Special Rapporteur of the UN Human Rights

Council (UNHRC), Christof Heyns, called on ‘all states to declare and implement national moratoria on at least the testing, production, assembly, transfer, acquisition, deployment and use of [lethal autonomous robots]’. In his view, allowing such robots to kill people ‘may denigrate the value of life itself’ and ‘seriously undermine the ability of the international legal system to preserve a minimum world order’ (Heyns 2013, pp. 20–21). The UN Secretary-General (UNSG), António Guterres, delivered the following message at the Web Summit in 2018: ‘machines that have the power and the discretion to take human lives are politically unacceptable, are morally repugnant, and should be banned by international law’ (Guterres 2018). In its resolution of 12 September 2018, the European Parliament (EP) urged the member states ‘to work towards the start of international negotiations on a legally binding instrument prohibiting lethal autonomous weapon systems’. The same document cautioned that ‘lethal autonomous weapon systems have the potential to fundamentally change warfare by prompting an unprecedented and uncontrolled arms race’ (EP 2018). All these actors contribute their formal institutional power to the Campaign’s cause.

Dozens of countries have also called for an outright ban on fully autonomous weapons: Pakistan (2013), Ecuador (2014), Egypt (2014), the Holy See (2014), Cuba (2014), Ghana (2015), Bolivia (2015), the State of Palestine (2015), Zimbabwe (2015), Algeria (2016), Costa Rica (2016), Mexico (2016), Chile (2016), Nicaragua (2016), Panama (2016), Peru (2016), Argentina (2016), Venezuela (2016), Guatemala (2016), Brazil (2017), Iraq (2017), Uganda (2017), Austria (2018), Djibouti (2018), Colombia (2018), El Salvador (2018), Morocco (2018), Jordan (2019), and Namibia (2019). China’s ban call (2018) is limited to the use of fully autonomous weapons only and does not cover their development and production (CSKR 2020a). These countries have themselves taken active part in the securitization of the threat of AWS. For example, in November 2015, Iraq associated fully autonomous weapons with ‘an arms race which could have catastrophic results’. Colombia called such weapons ‘a military and legal *threat*’ in December 2016. Brazil warned in November 2017 that certain weapons with autonomous capabilities ‘will prove to be incompatible with international humanitarian law and international human rights law’. In October 2018, El Salvador stated that ‘a machine that has the responsibility to decide about a person’s life ... raises great ethical and legal challenges’ (cited in HRW 2020a). In April 2018, the Holy See put emphasis on the moral side of the problem in the following statement: ‘[a]n autonomous weapons system could never be a *morally responsible subject*’ (cited in PAX 2018, p. 16). The ‘moral tutelage’ by the Holy See is crucial, given its unique and symbolic ‘moral authority’ (Eyffinger 1999, pp. 77–88).

Even those countries whose governments have not necessarily been supportive of the Campaign’s disarmament call have often contributed to the securitization of AWS. For example, in May 2014, the Czech Republic warned that lethal autonomous weapons ‘could pose a serious threat for civilians’. In November 2014, Ireland also expressed concern at the ‘eventual use of these technologies outside of traditional combat situations, for example in law enforcement’. Kuwait stated in October 2015 that such weapons systems ‘pose moral, humanitarian, and legal challenges’. In October 2017, Myanmar explicitly characterized lethal autonomous weapons as ‘*a security issue*’. Finland concluded in November 2017 that the ‘development of weapons and means of warfare where humans are completely out of the loop would pose serious risks from the ethical and legal viewpoint’ (cited in HRW 2020a).

Now we turn our attention to the special status and role of the International Committee of the Red Cross (ICRC). The ICRC has international legal personhood, i.e. a ‘legal status vis-a-vis states in international law’. It is, at the same time, a recognized ‘humanitarian expert’ (Mathur 2011, p. 182, 2017, p. 19). First and foremost, however, it is the guardian of international humanitarian law (IHL), or the so-called Geneva Conventions (Forsythe 2005, p. 13; Mathur 2011, p. 182; 2017, pp. 4, 95). Since this is the case, the ICRC has long been at the ‘intersection’ of disarmament and IHL (Mathur 2017, p. 15). Normative and diplomatic facilitation by the ICRC, mobilizing its unique symbolic position, legal expertise and trust in support of the Campaign, has been observed consistently in the case of AWS. It has, since 2015, urged states ‘to establish internationally agreed limits on autonomous weapons systems to ensure civilian protection, compliance with international humanitarian law, and ethical acceptability’ (ICRC 2021a). Peter Maurer, the President of the ICRC, stressed in his recent speech that ‘the use of autonomous weapons to target human beings should be ruled out’. He explained his motive in saying so as follows: ‘The potential humanitarian consequences are concerning for the ICRC. These weapons systems raise serious challenges for compliance with international humanitarian law’ (cited in ICRC 2021b).

The media, traditionally perceived as the medium between securitizing actors and their audiences, have generously and actively contributed to the ‘killer robots’ discourse and spreading the sense of political urgency. Such titles abound: ‘Terminator or Robocop?’ (*The Economist* May 2013); ‘The Rise of the Killer Robots—And Why We Need to Stop Them’ (CNN, October 2015); ‘Killer Robots: New Reasons to Worry About Ethics’ (*Forbes*, January 2016); ‘Is “Killer Robot” Warfare Closer Than We Think?’ (BBC, August 2017); ‘Killer Robots Must Be Banned But “Window to Act is Closing Fast,” AI Expert Warns’ (*The Independent*, November 2017); ‘Killer Robots Are Coming:

Scientists Warn UN Needs Treaty to Maintain Human Control Over All Weapons’ (*The National Post*, November 2017); ‘Stop the Rise of the “Killer Robots,” Warn Human Rights Advocates’ (*The Washington Post*, November 2017); ‘Killer Robots Will Only Exist If We Are Stupid Enough to Let Them’ (*The Guardian*, June 2018); ‘Why We Need a Pre-emptive Ban on “Killer Robots”’ (*The Huffington Post*, August 2018); ‘Killer Robots Aren’t Regulated. Yet.’ (*The New York Times*, December 2019); ‘“Killer Robots” and AI Could Wipe Out Humanity, Report Warns’ (*The Telegraph*, August 2020). Broad media coverage has indeed been identified by the Campaign as one of the contributing mechanisms (CSKR 2018b, pp. 26–42).

It is also interesting to observe how ordinary people, the audience in its most traditional sense, become actively engaged in the process of securitization as well. They participate in polls and surveys which are then used by pro-ban advocates to publicly reinforce their position and the view that their policies are justified. Human Rights Watch announced in 2019, based on a survey conducted in December 2018 by Ipsos, that 61 percent of respondents from 26 countries are opposed to the development of killer robots (HRW 2019). Based on the new survey conducted by Ipsos in December 2020, the Campaign publicly declared in 2021 that opposition to killer robots remains strong and that more than three in five people responding to a new online survey in 28 countries oppose the use of fully autonomous weapons (CSKR 2021b). Human Rights Watch revealed, rather surprisingly, in the report titled ‘Children Vote to Stop Killer Robots’ that interest in the Campaign is ‘growing across the world, especially among children’ (HRW 2020b). The organization clearly sent the message to further open the gateway to a new trend of recognizing children as political contributors, a trend set by Swedish climate activist Greta Thunberg.

The previous paragraphs clearly demonstrate that different actors did participate at different points in the creation of baseline knowledge about the threat of AWS. The relationship between a large number of experts in different fields, bureaucrats, policy advocates, mass media, and the general public have become more symmetrical. Science is increasingly hijacked by political activists of whom many are scientists themselves, hence unbiased scientific research and an epistemically-oriented expert debate are compromised in an effort to deliver a political message. It means the hierarchy of knowledge, based on the primacy of scientific fact finding, is undermined. This gives a wide-open-window of opportunity for different actors with multiple interests, positions and strategies to engage in knowledge production. We call it the diffusion of authority and knowledge. Therefore, the securitizing ‘actor’, originally the Campaign, evolves into a hybrid construction, intentionally or unintentionally, bringing together everyone actively contributing to its cause.

Eventually, there is a repeated circulation of the same or similar arguments, iterated and reiterated continually. At the same time, facts become more relational and difficult to trace and juxtapose, as we will illustrate below. As seen from today’s perspective, circulatory, transepistemic, and post-truth knowledge underlies the understanding of what constitutes a security threat in the case of AWS. This causes problems when such a complex weapons category, especially if it deeply resonates with sci-fi imaginaries, is concerned, as we show below.

In fact, one may argue that the case is not so different from many other securitization processes in the field of arms control and disarmament. We can indeed cite numerous examples of when a wide range of actors were involved directly or indirectly in the securitization process (Hynek and Solovyeva 2020). However, the case we analyse here is almost unique (except perhaps cyber weapons, cf. Stevens 2019, p. 284). While it is relatively easy to define nuclear, biochemical, and laser weapons, for instance, the term ‘killer robots’ is notoriously difficult to define. So the diffusion of authority to different actors and a systematic departure from the kind of unbiased expertise necessary to design a workable security regime, reflecting all the pros and cons, reduce the likelihood of success. Not only is it getting more difficult to name the threat clearly, but it also getting harder to put across a consistent, scientific proven message to the target audience. The other side of the problem is that one-sided, selective, imprecise or incomplete information can be easily challenged.

Below are selected illustrative examples of what happens when too many actors are actively involved in securitization efforts when such a complex, sci-fi-laden weapons category is concerned. We identify the three most important challenges at the basic definitional level which result from overlapping security language games: (1) no clear definition of the so-called killer robots; (2) no certainty as to whether (or to what extent) they already exist; (3) no grasp of the relationship between AI and AWS. Another key point of contention is the line between fiction and reality, drawn differently by different actors involved in the process, but it will be discussed in the next section.

- (1) The following definition is given by Human Rights Watch: ‘Fully autonomous weapons, also known as “killer robots,” would be able to select and engage *targets* without meaningful human control’ (HRW 2021). AWS are similarly defined by the ICRC as those that ‘can independently select and attack *targets*, i.e. with autonomy in the “critical functions” of acquiring, tracking, selecting and attacking *targets*’ (ICRC 2014, p.7). An expert from the ICRC also drew attention to ‘serious concerns about allowing the decision to kill a human *or* apply violent force to be delegated to

autonomous weapons systems (AWS)—systems that, once activated, can track, identify and attack *targets* without further human intervention’ (Sharkey 2017, p. 178). However, all these definitions are broad enough to cover weapons systems capable of autonomous non-lethal engagement (Fig. 1). Although they may exhibit similar characteristics, they do not necessarily represent AWS. Amnesty International (2015) provides a more specific definition: ‘Killer robots are weapons systems which, once activated, can select, attack, kill and injure *human targets* without a person in control’. Another expert from the ICRC defined AWS even more precisely as systems ‘capable of targeting and initiating the use of *potentially lethal* force without direct human supervision and direct human involvement in *lethal* decision-making’ (Asaro 2012, p. 690). These definitions relate specifically to weapons systems capable of autonomous lethal engagement (Fig. 1). Paradoxically, all these actors call for urgent action against ‘killer robots’ but some definitions leave confusion as to whether one has to (be able to) kill a human to be a ‘killer’.

- (2) An article published in *The Guardian* indicates that ‘[f]ully autonomous weapons *do not yet exist*’ (Busby 2018). PAX fully agrees stating basically the same: ‘[k]iller robots *do not yet exist*’ (PAX 2021). In a section on ‘killer robots’ on their website dedicated to humanitarian disarmament, Harvard Law School’s Armed Conflict and Civilian Protection Initiative develops a sense of urgency, insisting that ‘killer robots’ are ‘currently under development [and] moving rapidly closer to reality’ (ACCPI 2021). There apparently remains confusion as to whether the Campaign is taking a ‘preventive’ (FLI 2015) or ‘preemptive’ (HRW 2016) action against AWS. In an article published by *The New York Times*, it is even reported that ‘[t]here *are* not many verified battlefield examples’ of the use of AWS (Satariano et al. 2021). Another source of direct contribution to the Campaign appears to have more definite information: ‘In reality, weapons which can autonomously select, target, and kill humans are *already here*’ (Stop Autonomous Weapons 2022a). Therefore, it remains unclear which weapons systems are covered by the definition and whether the Campaign’s objective is, in fact, a preventive, preemptive, or perhaps even an ex-post ban on AWS.
- (3) The BBC summarized one of the key initiatives in this issue area as follows: ‘A group of scientists has called for a ban on the development of weapons *controlled* by artificial intelligence’ (Ghosh 2019). An expert from the ICRC similarly envisioned AWS ‘*directed* by a sophisticated artificial intelligence’ (Sparrow 2007, p.66). However, PAX defined ‘killer robots ... as weap-

ons which, once activated, using sensors *and/or* artificial intelligence, will be able to operate without meaningful human control over the critical functions’ (PAX 2018, p. 6). One of the ICRC’s (2019, p. 5) reports also indicated that AWS ‘*do not necessarily* incorporate AI’. While problems and fears associated with AI have been foregrounded in the killer robots discourse, as discussed in detail below, there apparently remains considerable confusion as to whether AWS should necessarily possess AI. If the answer is ‘not’, the types of weapons concerned may arguably cover both scripted lethal autonomy (e.g. ‘fire-and-forget’ missiles, APLs) and lethally capable AI (Fig. 1). These are different categories, however, and the former has long been seen as legitimate. Without limited AI capabilities at least, a weapons system cannot ‘select’ targets, as discussed above, it can only engage targets pre-selected by humans or locate them based on some pre-selected criteria set at the time of programming. On the contrary, there also remains uncertainty as to whether the adaptation of AI for military purposes should necessarily be associated with AWS. Cited on the Campaign’s website, state representatives of Ecuador made the following statement, referring particularly to ‘lethal autonomous weapons’, at the UN General Assembly in 2020: ‘The *militarization* of artificial intelligence presents challenges for international security, transparency, control, proportionality, and accountability’ (CSKR 2020b). However, such a broad reference covers not only weapons but also other uses of AI in military systems and networks (e.g. intelligence, surveillance, and reconnaissance purposes, swarming capability, etc.). The UNSG, António Guterres, formulated the problem differently at the Web Summit in 2018: ‘The *weaponization* of artificial intelligence is a serious danger’ (Guterres 2018). This statement is more precise in that it focuses on the adaptation of AI for weapons systems in particular. But the word ‘weaponization’ can possibly cover AI-equipped weapons and lethally capable AI. As explained above, these are also different categories (Fig. 1). An expert who signed the open letter (FLI 2015) clarified that the focus should be on ‘lethal artificial intelligence’ (Garcia 2018, p.335).

The previous paragraphs illustrated how different interpretations of the same were provided by one NGO and another NGO, one expert and another expert, one media outlet and another media outlet, an NGO and a media outlet, a media outlet and an expert, a representative of an IO and a state representative, and so on. These are only selected examples but they do indicate incoherence, ambivalence, and uncertainty which result from the presence of too many voices in defining such a complex category of weapons

strongly resonating with sci-fi imaginaries. Increasingly difficult to trace, facts become more relational and source-dependent. The very fact that experts are getting directly engaged in policy advocacy, and therefore become unavoidably biased, undermines their scientific integrity, erodes the hierarchy of knowledge, and puts their arguments in line with the statements by policy advocates, state representatives, etc. Therefore, the above is a perfect illustration of how transepistemic and post-truth knowledge is produced, leaving room for interpretation and manipulation. Yet it is important to define the threat as precisely as possible to close interpretive loopholes. Differences in nuance and emphasis may be detrimental to the Campaign's ability to succeed when such a complex, sci-fi-laden weapons category is concerned. Although one may initially assume that the more actors are involved in spreading the message the better, the opposite may be the case. Illustrative of how inconsistency and imprecision hamper progress is China's statement submitted to the Group of Governmental Experts (GGE) on lethal AWS (LAWS) in 2018:

LAWS still lack a clear and agreed definition and many countries believe such weapon systems do not exist. ... Therefore we support discussions first on technical characteristics (specifications, perimeters) of LAWS and on such a basis seeking a clear definition and scope (Country Statement 2018b).

In 2017, Russia came up with a similar statement at the GGE, more explicitly rejecting the idea of already existing AWS:

[T]he lack of working samples of such weapons systems remains the main problem in the discussion on LAWS. Certainly, there are precedents of reaching international agreements that establish a preventive ban on prospective types of weapons. However, this can hardly be considered as an argument for taking preventive prohibitive or restrictive measures against LAWS being *a by far more complex* and wide class of weapons of which the current understanding of humankind is rather approximate (Country Statement 2017).

5 Grafting: deepening the sense of insecurity

There have been further problems at the level of threat construction. We interpret and problematize them under the rubric of grafting. Efforts to ban AWS clearly build on IHL and international human rights law (IHRL) (HRW 2016). At the same time, the use of emotive terms such as 'killer robots' was selected as an appropriate strategy from the

outset as indicated by the very name of the Campaign. The words 'danger' or 'dangerous' have also been in common use to create a sense of urgency and promote the centrality of the issue. Campaigners have described the emergence of increasingly autonomous weapons as a 'dangerous development' (HRW 2020a), a 'destabilizing robotic arms race' (CSKR 2022b), and 'a serious global threat' (Amnesty International 2015). Frightening videos have been created to alert people about the imminent danger and make sure they feel a true sense of urgency (e.g. Stop Autonomous Weapons 2017; CSKR 2018a; FLI 2019; PAX 2021). Contributing to the growing sense of urgency are statements of this kind: 'China, Israel, Russia, South Korea, the United Kingdom, and the United States are *investing heavily* in the development of various autonomous weapons systems' (HRW 2020a). But, to reinforce the sense of political urgency, the Campaign and its supporters went further: their agenda is also being grafted, both discursively and visually, onto other security issues of immediate importance, legal precedents and even science fiction imagery.

First of all, the fear of killer robots is stoked by their extremely stereotyped presentation in science fiction. Media coverage of the Campaign and AWS has broadly featured images of terrifying humanoid military robots, as seen in science fiction films such as *The Terminator* (e.g. Walsh 2015b; Devlin 2018). The Terminator has indeed become a central metaphor in the killer robots discourse. In one of the BBC reports, virtually the same visual representation was even accompanied by the following words: "Killer robots" may seem like something from a sci-fi film, but reality is catching up' (Smith 2017). In a *Forbes* report, a very similar image was captioned as '[t]he reality of the rise of autonomous weapons systems' (Pandya 2019). These examples clearly illustrate the importance of visuals in threat construction and presentation. It is also a clear fact that mass media has actively grafted the image of killer robots on horrifying visual representations of dangerous cinematic robots, therefore featuring diagonal grafting, i.e. an unprecedented kind of grafting technique through which weapons law is being grafted onto pop culture. What needs to be stressed, however, is that policy advocates have repeatedly tried to dispel the illusion of killer robots that look like the Terminator (Mary Wareham cited in Ghosh 2019; PAX 2021). Even one of the Campaign's earliest statements cited roboticist Noel Sharkey as saying that '[k]iller robots are not self-willed "Terminator"-style robots' (CSKR 2013). AI-equipped weapons, even if still under development, do not really have anything in common with the Terminator, as also shown and discussed above (Fig. 1). Stuart Russell, a professor of computer science at UC Berkeley and an expert contributor to the Campaign, even tried to discourage mass media from using images from *The Terminator*: 'I've tried

to convince journalists to stop using this image for every single article about autonomous weapons, and I've failed miserably' (cited in CBC Radio 2022). This is an unintended consequence of hybridization, manifesting itself at the level of grafting. While mass media has helped the Campaign to spread its message and promote the centrality of its agenda, misrepresentations of the threat—be they intentional or unintentional—do have an adverse effect on the securitization process because they make it even easier to challenge the Campaign's message at a very basic level. For example, Christopher A. Ford, as US Assistant Secretary of State for International Security and Nonproliferation, ironically noted:

[A]ctivists concerned about the possibility of LAWS have built their public messaging around evocative 'Skynet' imagery of 'killer robots' precisely because this presses the kind of emotive buttons [...] If that's what is meant by 'killer robots', who wouldn't be opposed to them? [...] [But] we are hardly now on an inexorable slippery slope to 'Skynet' (Ford 2020).

It is also fair to note that not only mass media has contributed to blurring the line, at least in the eyes of the general public, between fact and fiction in relation to AWS. For example, Amnesty International (2015) also highlighted the possible link between the two: "'Killer Robots' will not be a thing of science fiction for long'.

Besides misleading cultural references, the argument for banning AWS is also grafted onto the general public anxiety about AI, especially AI that can possibly turn against human beings. As some studies show, the fear of an AI uprising is deeply ingrained in Western thought (Cave and Dihal 2019). The pre-eminent scientist Stephen Hawking warned that the development of 'full' AI could indeed 'spell the end of the human race' (Cellan-Jones 2014). The technology entrepreneur Elon Musk also branded AI as 'a fundamental existential risk for human civilisation' (Sulleyman 2017). Against this background, the purposeful development of AI-equipped killing machines, even as a possibility, seems like a grim prospect. So arguments of this sort have surfaced in relation to AWS: 'Starting a military AI arms race is a bad idea, and should be prevented by a ban on offensive autonomous weapons beyond meaningful human control' (FLI 2015). The Asilomar AI Principles, proposed by AI experts in 2017, are also formulated in a way that does not explicitly distinguish (especially for non-professionals) between AI and AWS: 'AI Arms Race Principle: An arms race in lethal autonomous weapons should be avoided' (FLI 2017b). It is, however, a reductionist trap. Such a narrow definition of the race puts aside other uses of AI in weapons systems and does not accurately reflect the difference between AI-equipped weapons and lethally capable AI (Fig. 1). Incomplete or inaccurate representations of current R&D&P efforts make

it easier for the other side, represented mainly by the world leaders in the development of related technologies, to dismiss pro-ban arguments. For example, the US drew attention to the possible benefits of equipping weapons with AI:

Emerging autonomy-related technologies, such as artificial intelligence (AI) and machine learning, have remarkable potential to improve the quality of human life with applications such as driverless cars and artificial assistants. The use of autonomy-related technologies can even save lives, for example, by improving the accuracy of medical diagnoses and surgical procedures or by reducing the risk of car accidents. Similarly, the potential for these technologies to save lives in armed conflict warrants close consideration. ... AI could help commanders increase their awareness of the presence of civilians and civilian objects on the battlefield.... Automated target identification, tracking, selection, and engagement functions can allow weapons to strike military objectives more accurately and with less risk of collateral damage (Country Statement 2018c).

In addition to the fear of losing control over AI, the call for a total ban on AWS has also been grafted onto gender and racial biases in AI. Chandler (2021) commented on how these concerns are linked to AWS: 'The criteria that will inform who is and is not a combatant—and, therefore, a target—will likely involve gender, age, race, and ability'. Likely aware of the same, the Campaign publicly stated that 'achieving a ban on fully autonomous weapons or killer robots is a feminist issue' (CSKR 2020c). It was also stressed that lethal autonomous weapons 'increase the risk of targeted violence against classes of individuals, including ethnic cleansing and genocide' (Stop Autonomous Weapons 2022b). Another trusted source provides virtually the same information: 'Autonomous weapons are ideal for ... selectively killing a particular ethnic group' (FLI 2015). Sharkey (2018) explained that, although 'societal values and norms are constantly evolving ... most of the old values are locked into the internet where much of the training data for machine learning algorithms are derived'. However, such a representation of virtually inherent and almost unavoidable biases in AI is inaccurate. Its selective guiding focus is on negative aspects, while positive aspects are omitted. This brings us back to the point that algorithms eventually reflect human intent, as explained above in this article. Commenting specifically on complaints about gender and racial biases in AI, Sunstein (2022, p. 1189) rightly remarked: 'When we find algorithmic bias, or something close to it, the reason lies in emphatically human decisions, not in artificial intelligence as such'. 'AI develops or gives the output based on what we want, and it will optimize what we program it for', according to the professor of MIS/Statistics Gaurav Bansal (cited in Matta et al. 2022). On the contrary, Ayoub and Payne

(2016, p. 799) identified two phases of biased flow in human decision-making: humans ‘*subjectively* analyse ... already biased data’. AI, in their view, may allow for ‘data-driven, bias-free analysis of the biased data, resulting in only one block of bias in the process flow’. The ability of AI to collect its ‘own’ data via sensors further reduces its susceptibility to human bias, they added.

All of the above creates the impression of a two-sided coin and leads to counter-arguments, regardless of their real motivation, similar to that made by China:

[E]ven though emerging technologies such as AI are the basic technologies in the area of LAWS, they have already been widely applied in the economic and social development of many countries, and have greatly advanced human progress. China believes that the impact of emerging technologies deserve objective, impartial and full discussion. Until such discussions have been done, there should not be any *pre-set premises or prejudged outcome* which may impede the development of AI technology (Country Statement 2018b).

The argument for banning AWS is also grafted onto legal precedents in disarmament law but with less success. Parallels are drawn between the case of blinding laser weapons and that of AWS (HRW 2015). However, the Protocol on Blinding Laser Weapons (1995) did in fact target a specific weapon with a well-defined harmful effect: permanent blindness (Sivakumaran 2012, p. 399). It was precisely the reason for strong stigmatization to succeed in this case. The same cannot be said about AWS. As discussed in more detail above, their military effect is neutral-to-beneficial under certain conditions. The ‘killer robots’ rhetoric is also grounded in the discourse on WMD. For example, health professionals highlighted in their aforementioned open letter that ‘lethal autonomous weapons can ... become weapons of mass destruction enabling very few to kill very many’ (FLI 2022a). The Future of Life Institute also defined lethal autonomous weapons as ‘a new class of weapon of mass destruction’ (FLI 2022c). The arms control advocacy viral video *Slaughterbots* visually demonstrated how killer drone swarms can be turned into robotic WMD (Stop Autonomous Weapons 2017). This is an attempt to re-create the emotional appeal of WMD. WMD-based language is, in most cases, pervasive enough to stigmatize the weapons in question as immoral and unacceptable, especially as biological and chemical weapons are already prohibited by the Biological Weapons Convention (1972) and the Chemical Weapons Convention (1993). But there has been little tangible progress in banning nuclear weapons and the situation is even more complicated with AWS. There are many proposed benefits to the actual use (as opposed to the benefits that deterrence yields in the case of nuclear weapons) of AWS. This

is, at the very least, what distinguishes AWS from all other WMD. Enemark (2011) argued, and in this we concur, that the term WMD is ‘misleading from a technological viewpoint’. He stressed that it ‘obscures the paramount threat of nuclear weapons, exaggerates the destructive power of chemical weapons, and is unhelpful or counterproductive when used in the context of biological weapons’. Falling into the reductionist trap of such umbrella terms bears the risk of further over-simplification in the case of AWS.

The above lines illustrate that the Campaign’s steps towards deepening the sense of insecurity about AWS through grafting this security issue onto other issues of immediate importance, legal precedents and, with the help of mass media, even science fiction imagery may, in fact, be counter-productive. In an effort to bring the prospect of lethally capable AI as close as possible to the fears of regular human beings and in search for simple and clear messages that may resonate as broad as possible with the general public, the Campaign achieves the opposite effect. It ends up with a vague definition of the threat because a multi-layered threat construction based on one-sided evidence, problematic over-simplifications, flawed generalizations, and sci-fi imaginaries—rather than anything close to what can be seen as a tangible, comprehensible, and unequivocal threat—becomes subject to securitization. This gives rise to reasonable counter-arguments, as we also illustrated above.

6 Concluding remarks

Our findings leave a big question mark on whether not to rethink the popular assumption that strong stigmatization is the best universally applicable disarmament strategy. We came to the same conclusion elsewhere in relation to a different weapons category, yet for a different set of reasons (Solovyeva and Hynek 2022). We understand that opting for impartial consideration of the technical nuances and respective pros and cons of AWS, the Campaign would have to deal with an endless chain of arguments and counter-arguments and lose the sense of urgency. The Campaign chose the path of normatively-oriented, preventive ban-motivated strong stigmatization and has recently reported success: ‘a stigma is already becoming attached to the prospect of removing meaningful human control from weapons systems and the use of force’ (CSKR 2019). However, their preferred strategy has serious problems too. We showed it by pointing out the paradox of what we called over-securitization. In particular, we identified and theorized two mechanisms facilitating it, namely hybridization and grafting, and demonstrated that success in broadening the stakeholder base and deepening the sense of insecurity, respectively, does not necessarily mean the success of securitization. Through these lenses, we explained why strong stigmatization has not translated

into a ban on AWS. This is due to the complexity of AWS as a distinct category of weapons, combined with their popular image inspired by *The Terminator*. To be more precise, this is due to the ease with which one may eventually dismiss one-sided evidence, problematic over-simplifications, flawed generalizations, and sci-fi imaginaries circulated by a wide range of actors involved directly or indirectly in the securitization process. Eventually, there is no international regime banning, or even purposefully regulating, AWS. Stevens (2019) discussed the ontological complexity of cyber-weapons as a significant barrier to their effective regulation. The prominence of pop culture in creating a public image of AWS adds another layer of complexity, as we showed here, and makes this case even more challenging than that of cyberweapons.

With respect to the role of pop culture, we demonstrated an unprecedented practice of diagonal grafting performed mainly by mass media, contrary to the Campaign's efforts to disclaim the connection between AWS and *The Terminator*. This leads us to assume that there has, in fact, been one more mechanism of over-securitization: two-phased translation of what AWS represent, i.e. preventive ban-motivated simplification of their definition by the Campaign, further simplified and eventually discredited by mass media spreading delusional fantasies about Terminator-like killing machines. We illustrated this mechanism in action throughout this paper, but we did not engage with it conceptually or systematically due to space limitation. We recommend that further research is conducted in this direction. One unintended consequence of an otherwise beneficial relationship between the Campaign and mass media has been overlooked: while mass media helps the Campaign to spread its message and promote the centrality of its agenda, it simultaneously undermines the Campaign's chances for success by reducing its agenda to easily dismissible arguments about the nexus between AWS and *The Terminator*.

It is possible that, unlike it could be expected based on previous disarmament campaigns, an institutionalized and epistemically-oriented expert debate with a less ambitious, lowest common denominator strategy may well constitute the preferred model of arms control for such a complex, sci-fi laden weapons category as AWS.

In theoretical terms, we forged an original way of thinking about the process and existing tools of securitization, especially as the problem of over-securitization has heretofore received scant attention. However, our theorization of (over-)securitization was primarily motivated by our intention to better understand the case of AWS. Although it may serve as an inspiration for further efforts to tailor the concept to practical problems, we do not claim its broad applicability.

Acknowledgements The authors gratefully acknowledge funding from Charles University and its part in the collaborative research project OPSEC “Automated System for Critical Infrastructure Protection Using Cyber-Physical Technologies” (VK01030216). We are especially grateful to the journal editor and anonymous reviewers for their valuable comments and suggestions for improving the paper. Any remaining omissions and errors are our own.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- ACCPI [Armed Conflict & Civilian Protection Initiative] (2021) Humanitarian Disarmament: Killer Robots. <https://humanitariandisarmament.org/issues/killer-robots/>. Accessed 21 Sep 2021
- Altmann J, Sauer F (2017) Autonomous weapon systems and strategic stability. *Survival* 59(5):117–142
- Amnesty International (2015) Ten reasons why it's time to get serious about banning 'killer robots'. November 12. <https://www.amnesty.org/en/latest/news/2015/11/time-to-get-serious-about-banning-killer-robots/>
- Apazidi E (2022) Робота Федора Назвали Бесполезным на МКК [Robot fedor called useless on the ISS]. Lenta, March 15. <https://lenta.ru/news/2022/03/15/robotfedotmikomunenujen/>
- Aradau C, Huysmans J (2018) Assembling credibility: knowledge, method and critique in times of 'post-truth.' *Secur Dialogue* 50(1):40–58
- Arkin R (2009a) Governing lethal behavior in autonomous robots. CRC Press, Boca Raton
- Arkin R (2009b) Ethical robots in warfare. *IEEE Technol Soc Mag* 28(1):30–33
- Arkin R (2018) Lethal autonomous systems and the plight of the non-combatant. Chapter 15. In: Kiggins R (ed) *The political economy of robots*. Palgrave Macmillan—International Political Economy Series, Chennai
- Asaro P (2008) How just could a robot war be? Part II. In: Briggie A, Waelbers K, Brey PAE (eds) *Current issues in computing and philosophy*. IOS Press, Amsterdam
- Asaro P (2012) On Banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *Int Rev Red Cross* 94(886):687–709
- Axe D (2022) Ukrainian missile crews may be creeping closer to a key Russian-held city. *Forbes*, July 22. <https://www.forbes.com/sites/davidaxe/2022/07/22/the-russians-lost-another-high-tech-jet-over-ukraine/?sh=725451ca7add>
- Ayoub K, Payne K (2016) Strategy in the age of artificial intelligence. *J Strateg Stud* 39(5–6):793–819
- Balzacq T (2005) The three faces of securitization: political agency, audience and context. *Eur J Int Rel* 11(2):171–201
- Berling TV (2011) Science and securitization: objectivation, the authority of the speaker and mobilization of scientific facts. *Secur Dialogue* 42(4–5):385–397

- Bhuta N, Rotolo A, Sartor G (2014) Awareness and responsibility in autonomous weapons systems. Chapter 17. In: Pitt J (ed) *Computer after me, the: awareness and self-awareness in autonomic systems*. Imperial College Press, London
- Bigo D (2000) When two become one: internal and external securitisations in Europe. Chapter 8. In: Kelstrup M, Williams MC (eds) *International relations theory and the politics of European integration: power, security and community*. Routledge, London
- Bigo D (2002) Security and immigration: toward a critique of the governmentality of unease. *Alternatives* 27:63–92
- Bigo D (2006) Globalized (in)security: the field and the ban-opticon. Chapter 2. In: Bigo D, Tsoukala A (eds) *Terror, insecurity, liberty: illiberal practices of liberal regimes after 9–11*. Routledge, London
- Birnbacher D (2016) Are autonomous weapons systems a threat to human dignity? Chapter 5. In: Bhuta N et al (eds) *Autonomous weapons systems: law, ethics, policy*. Cambridge University Press, Cambridge
- Bode I, Watts T (2021) Meaning-less human control: lessons from air defence systems on meaningful human control for the debate on AWS. February. <https://dronewars.net/wp-content/uploads/2021/02/DW-Control-WEB.pdf>
- Brauch HG (2009) Securitizing global environmental change. Chapter 4. In: Brauch HG et al (eds) *Facing global environmental change*. Springer, Berlin
- Brenneke M (2018) Lethal autonomous weapon systems and their compatibility with international humanitarian law: a primer of the debate. Chapter 3. In: Gill T, Geiß R, Krieger H, Paulussen C (eds) *Yearbook of international humanitarian law*, vol 21. Springer, Berlin
- Busby M (2018) Killer robots: pressure builds for ban as governments meet. *The Guardian*, April 9. <https://www.theguardian.com/technology/2018/apr/09/killer-robots-pressure-builds-for-ban-as-governments-meet>
- Buzan B, Wæver O, de Wilde J (1998) *Security: a new framework for analysis*. Lynne Rienner, Boulder
- Carpenter C (2016) Rethinking the political/-science/-fiction nexus: global policy making and the campaign to stop killer robots. *Am Polit Sci Assoc* 14(1):53–69
- Cave S, Dihal K (2019) Hopes and fears for intelligent machines in fiction and reality. *Nat Mach Intell* 1:74–78
- CBC Radio (2022) Quit using the terminator as an example of AI gone wrong, argues BBC reith lecturer. January 10. <https://www.cbc.ca/radio/ideas/quit-using-the-terminator-as-an-example-of-ai-gone-wrong-argues-bbc-reith-lecturer-1.6309630>
- Cellan-Jones R (2014) Stephen hawking warns artificial intelligence could end mankind. *BBC News*, December 2. <https://www.bbc.com/news/technology-30290540>
- Chandler K (2021) AI is often biased. Will UN member states acknowledge this in discussions of autonomous weapon systems? IPI Global Observatory, September 20. <https://theglobalobservatory.org/2021/09/ai-is-often-biased-will-un-member-states-acknowledge-this-in-discussions-of-autonomous-weapon-systems/>.
- Chen S (2018) China's plan to use artificial intelligence to boost the thinking skills of nuclear submarine commanders. *South China Morning Post*. Published February 4. Updated February 6. <https://www.scmp.com/news/china/society/article/2131127/chinas-plan-use-artificial-intelligence-boost-thinking-skills>.
- Country Statement (2017) Examination of various dimensions of emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the convention, submitted by the Russian Federation. November 10. Group of Governmental Experts of the high contracting parties to the convention on prohibitions or restrictions on the use of certain conventional weapons which may be deemed to be excessively injurious or to have indiscriminate effects, Geneva, November 13–17. CCW/GGE.1/2017/WP.8
- Country Statement (2018a) Proposal by the Russian Federation on the working definition of the 'fully-autonomous systems'. group of governmental experts of the high contracting parties to the convention on prohibitions or restrictions on the use of certain conventional weapons which may be deemed to be excessively injurious or to have indiscriminate effects, Geneva, April 9–13
- Country Statement (2018b) Position paper, submitted by China. April 11. Group of governmental experts of the high contracting parties to the convention on prohibitions or restrictions on the use of certain conventional weapons which may be deemed to be excessively injurious or to have indiscriminate effects, Geneva, April 9–13. CCW/GGE.1/2018b/WP.7
- Country Statement (2018c) Humanitarian benefits of emerging technologies in the area of lethal autonomous weapon systems, submitted by the United States of America. March 28. Group of governmental experts of the high contracting parties to the convention on prohibitions or restrictions on the use of certain conventional weapons which may be deemed to be excessively injurious or to have indiscriminate effects, Geneva, April 9–13. CCW/GGE.1/2018c/WP.4
- CSKR [campaign to stop killer robots] (2013) Urgent action needed to ban fully autonomous weapons: non-governmental organizations convene to launch campaign to stop killer robots. April 23. https://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_LaunchStatement_23Apr2013_fnl.pdf
- CSKR (2018a) No country would be safe from fully autonomous weapons. Youtube, April 5. <https://www.youtube.com/watch?v=qiJTq11kqdw>
- CSKR (2018b) Convention on conventional weapons group of governmental experts meeting on lethal autonomous weapons systems and meeting of high contracting parties United Nations, Geneva. Report on Activities, February 26. https://www.stopkillerrobots.org/wp-content/uploads/2018b/02/CCW_Report_Nov2017_posted.pdf.
- CSKR (2019) Minority of states delay effort to ban killer robots. March 29. <https://www.stopkillerrobots.org/news/minority-of-states-delay-effort-to-ban-killer-robots/>
- CSKR (2020a) Country views on killer robots. March 11. https://www.stopkillerrobots.org/wp-content/uploads/2020a/03/KRC_CountryViews_11Mar2020a.pdf
- CSKR (2020b) 75th UN Assembly. October 30. <https://www.stopkillerrobots.org/news/un-diplomacy/>
- CSKR (2020c) Killer robots, feminism...and a feminist foreign policy? February 13. <https://stopkillerrobots.medium.com/killer-robots-feminism-and-a-feminist-foreign-policy-d55309c60fae>
- CSKR (2021a) Stopping killer robots: a guide for policy makers. november. <https://www.stopkillerrobots.org/wp-content/uploads/2021a/11/211123-A-Guide-for-Policy-Makers-WEB.pdf>. Accessed 21 Nov 2022
- CSKR (2021b) Opposition to killer robots remains strong—poll. January 28. <https://www.stopkillerrobots.org/news/poll-opposition-to-killer-robots-strong/>
- CSKR (2022a) Who wants to ban fully autonomous weapons. Video. <https://www.stopkillerrobots.org>. Accessed 4 Oct 2021
- CSKR (2022b) The problem. <https://www.stopkillerrobots.org/learn/#problem>. Accessed 23 Sep 2021
- CSKR (2022c) Military and Killer Robots. Accessed November 1, 2021. <https://www.stopkillerrobots.org/military-and-killer-robots/>. Accessed 1 Nov 2021
- CSKR (2022d) Less autonomy. More humanity. <https://www.stopkillerrobots.org/>. Accessed 1 Nov 2021
- Defense Post (2020) General atomics to develop smart sensor technology for MQ9 reaper drone. November 27.

- <https://www.thedefensepost.com/2020/11/27/sensor-technology-mq9-reaper-drone/>
- Devlin H (2018) Killer robots will only exist if we are stupid enough to let them. *The Guardian*, June 11. <https://www.theguardian.com/technology/2018/jun/11/killer-robots-will-only-exist-if-we-are-stupid-enough-to-let-them>
- Doswald-Beck L, Herby P, Dorais-Slakmon J (1995) Basic facts: the human cost of landmines. Landmines in Africa: Fact Sheet, January 1. Archived page of the ICRC. <https://www.icrc.org/en/doc/resources/documents/misc/57jmcy.htm>
- Enemark C (2011) Farewell to WMD: the language and science of mass destruction. *Contemp Secur Policy* 32(2):382–400
- EP [European Parliament] (2018) European Parliament Resolution on Autonomous Weapon Systems. Doc. No. 2018/2752(RSP). https://www.europarl.europa.eu/doceo/document/TA-8-2018-0341_EN.pdf?redirect
- Etzioni A (2018) Pros and cons of autonomous weapons systems (with Oren Etzioni). *Happiness is the wrong metric: a liberal communitarian response to populism*. Library of public policy and public administration, vol 11. Springer, Cham
- Eyffinger A (1999) *The 1899 Hague peace conference: the parliament of man, the federation of the World*. Kluwer Law International, Hague
- Finn A, Scheduling S (2010) *Developments and challenges for autonomous unmanned vehicles: a compendium*. Springer, Berlin
- FLI [Future of Life Institute] (2015) Autonomous weapons: an open letter from AI and robotics researchers. Announced July 28, 2015. <https://futureoflife.org/open-letter-autonomous-weapons/>. Accessed 23 Sep 2021
- FLI (2017a) An open letter to the united nations convention on certain conventional weapons. <https://futureoflife.org/autonomous-weapons-open-letter-2017a/>. Accessed 23 Sep 2021
- FLI (2017b) Asilomar AI principles. <https://futureoflife.org/ai-principles/>. Accessed 4 Oct 2021
- FLI (2019) Why we should ban lethal autonomous weapons. Newsletter, March. <https://futureoflife.org/2019/04/09/fli-march-2019-newsletter/?cn-reloaded=1>. Accessed 15 Oct 2021
- FLI (2022a) Lethal autonomous weapons: an open letter from the global health community. <https://futureoflife.org/medical-lethal-autonomous-weapons-open-letter/>. Accessed 23 Sep 2021
- FLI (2022b) Lethal autonomous weapons pledge. <https://futureoflife.org/lethal-autonomous-weapons-%20pledge/>. Accessed 5 Oct 2021
- FLI (2022c) Lethal autonomous weapons systems. <https://futureoflife.org/lethal-autonomous-weapons-systems/>. Accessed 15 Oct 2021
- Floyd R (2020) Securitisation and the function of functional actors. *Crit Stud Secur*. <https://doi.org/10.1080/21624887.2020.1827590>
- Ford CA (2020) AI, human-machine interaction, and autonomous weapons: thinking carefully about taking ‘killer robots’ seriously. April 20. *Arms control and international security papers* 11(2), produced by the Office of the Under Secretary of State for Arms Control and International Security.
- Forsythe DP (2005) *The humanitarians: the international committee of the red cross*. Cambridge University Press, Cambridge
- Froelich P (2021) Killer drone ‘hunted down a human target’ without being told to. *New York Post*, May 29. <https://nypost.com/2021/05/29/killer-drone-hunted-down-a-human-target-without-being-told-to/>
- Gadiyar R, Zhang T, Sankaranarayanan A (2019) Artificial Intelligence software and hardware platforms. Chapter 8. In: Gilbert M (ed) *Artificial intelligence for autonomous networks*. CRC Press, Boca Raton
- Garcia D (2018) Lethal artificial intelligence and change: the future of international peace and security. *Int Stud Rev* 20(2):334–341
- Ghosh P (2019) Call to ban killer robots in wars. *BBC News*, February 15. <https://www.bbc.com/news/science-environment-47259889>
- Goodfellow I, Bengio Y, Courville A (2016) *Deep learning*. MIT Press, Cambridge
- Grishchenko N (2017) Робот-Космонавт Федор Научился Стрелять с Двух Рук [Robot-cosmonaut fedor learned to shoot with two hands]. *RG.ru*, April 13. <https://rg.ru/2017/04/13/robot-kosmonavt-fedor-nauchilsia-streliat-s-dvuh-ruk.html>
- Gubrud M (2014) Stopping killer robots. *Bull at Sci* 70(1):32–42
- Guterres A (2018) Remarks at ‘Web Summit’ by United Nations Secretary-General. Lisbon, November 5. <https://www.un.org/sg/en/content/sg/speeches/2018-11-05/remarks-web-summit>
- Haas PM (1992) (ed.) *Knowledge, power, and international policy coordination*. Special Issue. *International Organization* 46(1)
- Hambling D (2021) Drones may have attacked humans fully autonomously for the first time. *New Scientist*, May 27. <https://www.newscientist.com/article/2278852-drones-may-have-attacked-humans-fully-autonomously-for-the-first-time/>
- Hammerstad A (2008) Securitisation as a self-fulfilling prophecy: refugee movements and the north-south security divide. *Political Science Association Annual Conference*, Swansea, April
- Heyns C (2013) Report of the special rapporteur on extrajudicial, summary or arbitrary executions. Doc. No. A/HRC/23/47
- Horowitz MC (2016) Why words matter: the real world consequences of defining autonomous weapons systems. *Temple Int Comp Law J* 30(1):85–98
- HRW [Human Rights Watch] (2012) losing humanity: the case against killer robots. November 19. <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>
- HRW (2015) Precedent for preemption: the ban on blinding lasers as a model for a killer robots prohibition. November 8. <https://www.hrw.org/news/2015/11/08/precedent-preemption-ban-blinding-lasers-model-killer-robots-prohibition>
- HRW (2016) Making the case: the dangers of killer robots and the need for a preemptive ban. December 9. <https://www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban>
- HRW (2019) Poll shows strong opposition to ‘killer robots’. January 22. <https://www.hrw.org/news/2019/01/22/poll-shows-strong-opposition-killer-robots>
- HRW (2020a) Stopping killer robots: country positions on banning fully autonomous weapons and retaining human control. August 10. <https://www.hrw.org/report/2020a/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>
- HRW (2020b) Children vote to stop killer robots. June 9. <https://www.hrw.org/news/2020b/06/09/children-vote-stop-killer-robots>
- HRW (2022) Killer robots. <https://www.hrw.org/topic/arms/killer-robots>. Accessed 1 Oct 2021
- Hynek N, Chandler D (2013) No emancipatory alternative, no critical security studies. *Crit Stud Secur* 1(1):46–63
- Hynek N, Solovyeva A (2020) *The logic of humanitarian arms control and disarmament: a power-analytical approach*. Rowman and Littlefield, London
- ICRC [International Committee of the Red Cross] (2014) *Autonomous weapons systems: technical, military, legal, and humanitarian aspects*. November 1. <https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014>
- ICRC (2019) *Autonomy, artificial intelligence and robotics: technical aspects of human control*. Geneva, August. <https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control>
- ICRC (2021a) ICRC position on autonomous weapon systems. May 12. <https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems>

- ICRC (2021b) Peter Maurer: 'we must decide what role we want human beings to play in life-and-death decisions during armed conflicts'. May 12. <https://www.icrc.org/en/document/peter-maurer-role-autonomous-weapons-armed-conflict>
- Ihlamur-Öner SG (2019) Delinking the migration-terrorism nexus: strategies for the de-securitization of migration. *Perceptions* 24(2–3):195–224
- Johnson AM, Axinn S (2013) The morality of autonomous robots. *J Mil Ethics* 12(2):129–141
- Kalashnikov Media (2018) Демонстрация Работы Системы Автоматического Управление Огнем [Demonstration of the automatic fire control system]. October 1. <https://kalashnikov.media/video/technology/demonstratsiya-raboty-sistemy-avtomaticheskogo-upravlenie-ognem>
- Kallenborn Z (2021) Applying arms-control frameworks to autonomous weapons. Brookings TechStream, October 5. <https://www.brookings.edu/techstream/applying-arms-control-frameworks-to-autonomous-weapons/>
- Kania EB (2018) Chinese sub commanders may get AI help for decision-making. *Defense One*, February 12. <https://www.defenseone.com/ideas/2018/02/chinese-sub-commanders-may-get-ai-help-decision-making/145906/>
- Kastan B (2013) Autonomous weapons systems: a coming legal 'singularity'? *J Law Technol Policy* 45(1):46–82
- Keller J (2017) The company that built the AK-47 is developing a real life 'terminator'. *The National Interest*, July 18. <https://nationalinterest.org/blog/the-buzz/the-company-built-the-ak-47-developing-real-life-terminator-21578>
- Kruglov A, Ramm A, Dmitriev E (2018) Средства ПВО Объединят Искусственным Интеллектом [Air defense systems will be integrated with the help of artificial intelligence]. *Izvestia*, May 2. <https://iz.ru/733333/aleksandr-kruglov-aleksei-ramm-evgen-ii-dmitriev/sredstva-pvo-obediniat-iskusstvennym-intellektom>
- Kumagai J (2007) A robotic sentry for Korea's Demilitarized Zone. *IEEE Spectr* 44(3):16–17
- Layton P (2018) Algorithmic warfare: applying artificial intelligence to warfighting. Air Power Development Centre, Canberra
- Lewis PR (2014) Computational self-awareness and learning machines. Chapter 18. In: Pitt J (ed) *Computer after me, the: awareness and self-awareness in autonomic systems*. Imperial College Press, London
- Lin P, Bekey G, Abney K (2008) Autonomous military robotics: risk, ethics, and design. Version: 1.0.8. Prepared for US Department of Navy, Office of Naval Research. Cal Poly: San Luis Obispo, California Polytechnic State University.
- Lucas GR (2010) Postmodern war. *J Mil Ethics* 9(4):289–298
- Mathur R (2011) Humanitarian practices of arms control and disarmament. *Contemp Secur Policy* 32(1):176–192
- Mathur R (2017) *Red cross interventions in weapons control*. Lexington Books, Lanham
- Matta V, Bansal G, Akakpo F, Christian S, Jain S, Poggemann D, Rousseau J, Ward E (2022) Diverse perspectives on bias in AI. *J Inform Technol Case Appl Res* 24(2):135–143
- McFarland T (2015) Factors shaping the legal implications of increasingly autonomous military systems. *Int Rev Red Cross* 97(900):1313–1339
- McInnes C, Rushton S (2011) HIV/AIDS and securitization theory. *Eur J Int Rel* 19(1):115–138
- Moran M (2021) Killer AI drones 'hunted down humans without being told to', warns bombshell UN report. *Daily Star*, May 28. <https://www.dailystar.co.uk/news/world-news/killer-ai-drones-hunted-down-24203431>
- Nilsson NJ (2010) *The quest for artificial intelligence: a history of ideas and achievements*. Cambridge University Press, Cambridge
- RIA Novisti (2017) Человекоподобный Робот Федор Научился Стрелять с Двух Рук [Humanoid robot fedor learned to shoot with two hands]. April 13. <https://ria.ru/20170413/1492217045.html>
- O'Connell ME (2014) 21st century arms control challenges: drones, cyber weapons, killer robots, and WMDs. *Wash Univ Glob Stud Law Rev* 13(3):515–533
- Osborn K (2017) Air Force Chief Scientist confirms F-35 will include artificial intelligence. *Defense Systems*, January 20. <https://defensesystems.com/articles/2017/01/20/f35.aspx>
- Özberk T (2021) Turkish drone maker denies autonomous strike capability. *Shephard*, July 16. <https://www.shephardmedia.com/news/uv-online/turkish-drone-maker-denies-autonomous-strike-capability/>
- Pandya J (2019) The weaponization of artificial intelligence. *Forbes*, January 14. <https://www.forbes.com/sites/cognitiveworld/2019/01/14/the-weaponization-of-artificial-intelligence/?sh=3775f7453686>
- PAX (2018) Crunch time: European positions on lethal autonomous weapon systems. November. <https://paxforpeace.nl/media/download/pax-rapport-crunch-time.pdf>
- PAX [PAX for Peace] (2018) *Where to Draw the Line: Increasing Autonomy in Weapon Systems – Technology and Trends*. Published November 2017. Updated April 2018. <https://paxforpeace.nl/media/download/pax-report-where-to-draw-the-line.pdf>
- PAX (2019) Slippery slope: the arms industry and increasingly autonomous weapons. November. <https://paxforpeace.nl/media/download/pax-report-slippery-slope.pdf>
- PAX (2022) Killer robots. <https://paxforpeace.nl/what-we-do/programmes/killer-robots>. Accessed 28 Sep 2021
- Price R (1998) Reversing the gun sights: transnational civil society targets land mines. *Int Organ* 52(3):613–644
- Ramm A, Litovkin D, Andreev E (2017) В Войска Радиоэлектронной Борьбы Придет Искусственный Интеллект [Artificial intelligence will complement the electronic warfare troops]. *Izvestia*, April 4. <https://iz.ru/news/675891>
- Rosert E, Sauer F (2021) How (not) to stop the killer robots: a comparative analysis of humanitarian disarmament campaign strategies. *Contemp Secur Policy* 42(1):4–29
- Salter MB (2008) Securitization and desecuritization: a dramaturgical analysis of the Canadian air transport security authority. *J Int Relat Dev* 11(4):321–349
- Salter MB (2012) Securitization of US–Canada border in American political discourse. *Can J Polit Sci* 44(4):929–951
- Satariano A, Cumming-Bruce N, Gladstone R (2021) Killer robots aren't science fiction. A push to ban them is growing. *The New York Times*, December 17. <https://www.nytimes.com/2021/12/17/world/robot-drone-ban.html>
- Scharre P (2018) *Army of none: autonomous weapons and the future of war*. Norton, New York
- Schmitt MN (2013) Autonomous weapon systems and international humanitarian law: a reply to the critics. *Harvard Natl Secur J Feature* 4:1–37
- Sharkey N (2008) Cassandra or false prophet of doom: AI robots and war. *IEEE Intell Syst* 23:14–17
- Sharkey N (2010) Saying 'no!' to lethal autonomous targeting. *J Mil Ethics* 9(4):369–383
- Sharkey N (2012) The inevitability of autonomous robot warfare. *Int Rev Red Cross* 94(886):787–799
- Sharkey N (2017) Why robots should not be delegated with the decision to kill. *Connect Sci* 29(2):177–186
- Sharkey N (2018) The impact of gender and race bias in AI. *Humanitarian Law & Policy* (blog), August 28. <https://blogs.icrc.org/law-and-policy/2018/08/28/impact-gender-race-bias-ai/>
- Sismondo S (2017) Post-truth? *Soc Stud Sci* 47(1):3–6
- Sivakumaran S (2012) *The law of non-international armed conflict*. Oxford University Press, Oxford

- Smith M (2017) Is ‘killer robot’ warfare closer than we think? BBC News, August 25. <https://www.bbc.com/news/business-41035201>
- Solovyeva A, Hynek N (2018) Going beyond the ‘killer robots’ debate: six dilemmas autonomous weapon systems raise. *Central Eur J Int Secur Stud* 12(3):166–208
- Solovyeva A, Hynek N (2022) The paradox of success: evolutionary dynamics between human rights and small arms. *J Hum Rights* 21(1):36–55
- Sparrow R (2007) Killer robots. *J Appl Philos* 24(1):62–77
- Stevens T (2019) Global Code: power and the weak regulation of cyberweapons. Chapter 13. In: Hynek N, Ditrych O, Stritecky V (eds) *Regulating global security: insights from conventional and unconventional regimes*. Palgrave Macmillan, Cham
- STM (2020) KARGU—rotary wing attack drone loitering munition system. Youtube, August 26. <https://www.youtube.com/watch?v=auRlh-f2wwQ>
- Stop Autonomous Weapons (2017) Slaughterbots. Youtube, November 13. <https://www.youtube.com/watch?v=9CO6M2HsoIA>
- Stop Autonomous Weapons (2022a) Slaughterbots are here. <https://autonomousweapons.org/>. Accessed 1 Nov 2021
- Stop Autonomous Weapons (2022b) The risks of lethal autonomous weapons. <https://autonomousweapons.org/the-risks/>. Accessed 1 Nov 2021
- Stritzel H (2012) Securitization, power, intertextuality: discourse theory and the translations of organized crime. *Secur Dialogue* 43(6):549–567
- Sulleyman A (2017) Elon Musk: AI is a ‘fundamental existential risk for human civilisation’ and creators must slow down. *Independent*, July 17. <https://www.independent.co.uk/tech/elon-musk-ai-human-civilisation-existential-risk-artificial-intelligence-creat-or-slow-down-tesla-a7845491.html>
- Sunstein CR (2022) Governing by algorithm? No noise and (potentially) less bias. *Duke Law J* 71(6):1175–1205
- Tamburrini G (2016) On banning autonomous weapons systems: from deontological to wide consequentialist reasons. Chapter 6. In: Bhuta N et al (eds) *Autonomous weapons systems: law, ethics, policy*. Cambridge University Press, Cambridge
- TASS [Russian News Agency] (2018) Ростех: Россия Уже Имеет Нарботки по Самообучающемуся Оружию с Искусственным Интеллектом [Rostec: Russia already has experience in self-learning weapons with artificial intelligence]. June 6. <https://tass.ru/armiya-i-opk/5268765>
- The Times of Israel (2022) In first, Russian Military Said to Fire S-300 Missiles at Israeli Jets Over Syria. May 17. <https://www.timesofisrael.com/in-first-russian-military-said-to-fire-s-300-missiles-at-israeli-jets-over-syria/>
- Tsoukala A (2006) Democracy in the light of security. *Polit Stud* 54(3):607–627
- UN Panel of Experts on Libya (2021) Final report of the panel of experts on Libya Established Pursuant to Security Council Resolution 1973 (2011). Doc. S/2021/229
- UNODA [United Nations Office for Disarmament Affairs] (2017) Perspectives on Lethal Autonomous Weapon Systems, UNODA Occasional Papers, No. 30, United Nations Publication, New York
- Vallor S (2016) *Technology and the virtues: a philosophical guide to a future worth wanting*. Oxford University Press, New York
- Vultee F (2011) Securitization as a media frame: what happens when the media ‘speak security’. Chapter 4. In: Balzacq T (ed) *Securitization theory: how security problems emerge and dissolve*. Routledge, London
- Wæver O (1995) Securitization and desecuritization. Chapter 3. In: Lipschutz RD (ed) *On security*. Columbia University Press, New York
- Wagner M (2014) The dehumanization of international humanitarian law: legal, ethical, and political implications of autonomous weapon systems. *Vanderbilt J Transl Law* 47(5):1371–1424
- Wakefield J (2018) South Korean University Boycotted Over ‘Killer Robots’. BBC News, April 5. <https://www.bbc.com/news/technology-43653648>
- Walsh J (2015a) Political accountability and autonomous weapons. *Res Polit* 2(4):1–6
- Walsh T (2015b) The rise of the killer robots—and why we need to stop them. CNN, October 26. <https://edition.cnn.com/2015/10/26/opinions/killer-robots-walsh/index.html>
- Whitmore G (2014) Killer robots in film—in pictures. *The Guardian*, February 22. <https://www.theguardian.com/film/gallery/2014/feb/22/killer-robots-in-film-in-pictures-alien-blade-runner-terminator>
- Williams MC (2003) Words, images, enemies: securitization and international politics. *Int Stud Quart* 47(4):511–532
- Zitser J (2021) A rogue killer drone ‘hunted down’ a human target without being instructed to, UN report says. *Insider*, May 30. <https://www.businessinsider.com/killer-drone-hunted-down-human-target-without-being-told-un-2021-5>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.