EU Non-Proliferation and Disarmament Consortium

Promoting the European network of independent non-proliferation and disarmament think tanks

NON-PROLIFERATION AND DISARMAMENT PAPERS

No. 65 December 2019

THE QUESTION OF SWARMS CONTROL: CHALLENGES TO ENSURING HUMAN CONTROL OVER MILITARY SWARMS

MAAIKE VERBRUGGEN

I. INTRODUCTION

Swarms represent a futuristic use of autonomy in weapon systems. A swarm is a group of individual systems that interact and operate as a collective with a common goal. Swarms are a novel type of weapon system in which there is great military interest because of their potential for enabling new types of missions. However, because swarms operate as collectives, there are real limitations to the extent of control that humans can exert over them. This paper covers the various challenges to human control, especially where relevant to European Union (EU) research on swarms.

The goal of this paper is therefore twofold. First, it aims to increase policymakers' understanding of how swarms work based on the current state of the art in the literature of swarm robotics, and to highlight the challenges posed to human control over swarms. Second, it aims to provide clarity on ongoing EU defence research on swarms, and to inform the defence cooperation communities about the potential problems that swarms might pose. Unfortunately, although the fields of swarm robotics and defence swarms are highly related, there is little interaction between them. Integration of the issues in this paper may help to guide the policy decisions of the new European Commission (EC) that took office on 1 December 2019.

This paper commences by contexualizing these developments in section II, and providing a brief introduction to swarms in section III. This is followed by a review of the technical literature on swarm robotics discussing human control over machines in section IV and challenges to human control and human–swarm interaction in section V. Section VI considers the policy implications of the challenges raised in the previous sections. The final two main

SUMMARY

The European Union (EU) and EU member states are increasingly investing in military swarm research, despite the significant challenges that exist in establishing human control over swarms. These challenges include the high cognitive demands on human operators; interface and control design choices; disrupted communications between operator and swarm, whether from in-built technological limitations or environmental factors; and the inherent unpredictability of certain kinds of swarms. These challenges create tactical risks and increase the chances of undesired outcomes, such as conflict escalation and violations of international humanitarian law and ethical principles. EU-funded swarm research programmes should take steps to address these issues.

ABOUT THE AUTHOR

Maaike Verbruggen (Netherlands) is a doctoral researcher at the Institute for European Studies at the Vrije Universiteit Brussel. Her research focuses on the intersection between emerging technologies, military innovation and arms control. She is currently working on a PhD thesis on military innovation in artificial intelligence.

discussions, section VII and section VIII respectively, assess the status of human control as an aspect of swarm research in EU defence cooperation and provide recommendations for ensuring human control is considered under EU-funded swarm research programmes.

II. BACKGROUND

The international response to autonomous weapons systems and concepts of 'control'

Improvements in autonomous capabilities and increased global military interest in swarms have sparked discussions about their impact and the appropriate response by the international community. The prime forum for these discussions is the series of debates on lethal autonomous weapon systems (LAWS) that take place under the auspices of the 1981 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 (CCW Convention). LAWS are difficult to define, but they can generally be viewed as weapon systems that can select and engage targets without human intervention.

Within the discussions on LAWS, the concept of 'meaningful human control' (MHC) has gained popularity as a potential tool for regulation. For example, MHC could entail adopting a positive obligation to maintain a significant human role in the selection and engagement of targets. However, this concept also has fierce opponents, and even among its proponents there is no agreement on either the substance of the concept or how it should be used. To avoid confusion over the definition—and because this paper covers the nature of the human–machine relationship at large, including in regard to non-critical and non-lethal functions, as explained below—the emphasis here is on 'human control' rather than MHC. In this paper 'human control' is used as a catch-all term for the 'mechanism for achieving [a human] commander's intent' and refers to the extent of human influence on the outcome of the mission.[1]

The issue of human control is wider than LAWS, as autonomy poses fundamental questions about the changing relations between humans and machines.[2] If not thought out carefully, this new human–machine configuration could lead to a loss of human control over weapon systems. The loss of control is a direct result not only of automating certain functions, but also of a deterioration of critical independent judgements by humans over information provided by machines. As operators lose control, they lose the tools to ensure that missions are executed according to their intent.

Similarly, the challenges of human control over swarms in particular—while relevant to the discussions on the CCW Convention—also go beyond LAWS. The dispersed nature of a swarm means that control over it functions differently from controls over most existing autonomous systems, and the international community needs to consider how to translate existing concepts and frameworks onto such distributed and networked military systems. This paper focuses both on armed and unarmed swarms because control challenges are relevant to all types of swarms and also, since current research on unarmed swarms can aid in the development of armed swarms. For example, swarms that merely provide intelligence, surveillance and reconnaissance (ISR) functions can still provide critical information that ultimately leads to deadly engagements or causes accidents, while misperceptions about their use can escalate conflicts. It is thus important to apply a broad lens when assessing how swarms work, and not limit the analysis to swarms that could be defined as LAWS.

The role of the European Union

The EU plays a dual role with regard to LAWS. The first role is regulatory in nature: the European Parliament (EP) has called for a ban on LAWS; in Geneva the EU diplomatic mission aims to ensure that all weapon systems are developed, deployed and used in compliance with international humanitarian law (IHL); and in Brussels the European External Action Service (EEAS) actively searches for solutions to the issue of LAWS and how to consolidate the opinion of member states, since their opinions on these issues are widely divergent.[3] The second role is that of defence technology investment and development: in 2015

[1] Moyes, R., *Key Elements of Meaningful Human Control* (Article 36: Geneva, Apr. 2016), p. 4.

[2] Huelss, H., 'Deciding on appropriate use of force: human–machine interaction in weapons systems and emerging norms', *Global Policy*, vol. 10 (2019), pp. 354–58.

[3] Kayser, D. and Beck, A., *Crunch Time: European Positions on Lethal Autonomous Weapon Systems: Update 2018* (PAX: Utrecht, Nov. 2018).

the EU began directly funding defence research and development (R&D), with key priority areas being artificial intelligence (AI) and robotics in general, and swarms in particular. This paper investigates the state of swarm research and regulation in the EU, and makes recommendations for EU-funded research programmes to ensure that these two roles do not undermine each other.

The state of swarms technology

Currently there is only a small body of literature on swarms robotics that covers in-depth investigations of their working and use.[4] This apparent oversight is partially fuelled by a lack of information, as military operational swarms are still a few decades away. However, the literature on swarms robotics provides more information about how swarms work and the challenges to human control over swarms. This paper translates these technical findings to a policy audience.

Swarms are still in their infancy. A 2016 report from the Australian Department of Defence, Defence Science and Technology Group suggested that R&D into military swarms is only at the stage of developing proofs of concepts, designing components and demonstrating prototypes.[5] Most swarms are only tested in simulations or laboratories under highly optimized conditions. The few outdoor demonstrations minimize the problem, assume good communications, use off-board global positioning systems (GPS) and motion-capture infrastructure, and either simplify the environment or downsize the swarm.[6] It is thus too early to say what operational swarms would be capable of. This does not mean this paper is premature. The money, time, prestige and reputations involved in defence research projects mean that they are highly path-dependent. If problems are only brought up at the end, it is unlikely that projects will change course or be abandoned. Furthermore, international law requires legal reviews of new weapons and of new means and methods of warfare (so-called Article 36 reviews) while they are under study or development. Critical analysis at an early stage of whether the weapons under development do not pose major strategic, ethical and legal risks is therefore essential.

III. ABOUT SWARMS

A swarm is a group of systems that operate as a collective. A swarm is not a specific type of system, but a specific type of configuration, namely 'a large group of locally interacting individuals with common goals'.[7] This can be better understood with reference to biology. A school of fish, a flock of birds and a pack of wolves are all swarms: each comprises individuals, but the individuals interact and work as a group to achieve a collective goal. Swarms are different from multi-robot systems, in that operators do not control all individual units separately; rather, an operator of a swarm steers (subsets of) the swarm collectively. Internal communication and coordination is what defines a swarm. Attacks using multiple systems at once, but without internal coordination, thus do not qualify as swarms proper, but can perhaps be more accurately described as proto-swarms.

Swarms are generally envisioned as a large collection of homogeneous simple systems, but heterogeneous swarms also exist, as do both homogeneous and heterogeneous swarms with more complex units. Most commonly, the individual units of a swarm are unmanned systems, but a swarm can also contain manned systems (such as a swarm of one aircraft and many drones) or static sensors. Nonetheless, the individual units, or nodes, are usually not very advanced. The utility of swarms derives from the fact that the whole is better than the sum of its parts. Through coordination and task distribution, swarms can accomplish complex missions, giving them three major benefits. Swarms are: (*a*) scalable, as it is easy to change the size of the swarm depending on the mission; (*b*) adaptable, as they can be used for different types of missions; and (*c*) robust, because if a single node fails, other nodes can take over.[8]

[4] See, e.g., Brehm, M. and de Courcy Wheeler A., 'Swarms', Discussion Paper (Article 36: Geneva, Mar. 2019); Lachow, I., 'The upside and downside of swarming drones', *Bulletin of the Atomic Scientists*, no. 73 (2017), pp. 96–101; Hambling, D., *Change in the Air: Disruptive Developments in Armed UAV Technology* (UNIDIR: Geneva, Nov. 2018); Schmuel, S., 'The coming swarm might be dead on arrival', War on the Rocks, 10 Sept. 2018; Ilachinski, A., *AI, Robots, and Swarms: Issues, Questions, and Recommended Studies*, CNA Research Memorandum (CNA: Washington, DC, Jan. 2017).

[5] Ivanova, K., Gallasch, G. and Jordans, J. *Automated and Autonomous Systems for Combat Service Support: Scoping Study and Technology Prioritisation* (Australian Government Department of Defence, Defence Science and Technology Group: Canberra, 2016), p. 26.

[6] Steinberg, M., 'ONR science of autonomy & swarming challenges', Office of Naval Research, n.d.

[7] Barca, J. and Sekercioglu, Y., 'Swarm robotics reviewed', *Robotica*, vol. 31 (2013), p. 345.

[8] Tan, Y. and Zheng, Z., 'Research advance in swarm robotics', *Defence Technology*, vol. 9 (2013), pp. 20–23.

How are swarms organized?

There are four different command and control structures for swarms, divided into two categories. Centralized swarms have a central planner that coordinates tasks. Under 'centralized control', a leader commands all individual nodes, while under 'hierarchical coordination', task allocation is hierarchical through several subsequent layers. In contrast, decentralized swarms do not have a single leader or central planner. Under 'coordination by consensus', nodes collectively decide how to execute missions and coordinate tasks, for instance through voting or an auction system. Under 'emergent coordination', coordination arises because the individual nodes respond to the nodes surrounding them. Many swarm roboticists do not consider centralized systems to be swarms but general multi-robot systems, but both are included here as this is a convention in discussions of political strategy.[9]

Centralized swarms are better at finding good-enough solutions quickly and their behaviour is easier to plan in advance, but they are sensitive to the loss of their leader, computationally complex and slow to distribute commands.[10] Decentralized swarms are easier to expand in size, have no single point of failure, can operate under low communication bandwidth, are good at coming up with new solutions to problems assigned to them, and can achieve complex results with a simple system design, but the decisions that each node makes is based on localized information rather than on information aggregated at the global (swarm) level.[11]

The military benefits of swarms

Over the past decades, European militaries have prioritized quality over quantity when it comes to complex weapon systems. But as the costs of defence R&D have risen exponentially, per-unit costs have become so high that most militaries can only field a low number of them. In contrast, swarms are predicted to have low per-unit costs, which would allow militaries to field them in high numbers, and return mass to the battlefield.[12]

Decentralized swarms are considered especially promising because they rely less on constant communication with the operator, which reduces the network bandwidth required compared with that required for multi-robot systems or centralized swarms. This is militarily advantageous (as battlefields often lack good communications infrastructure and communications links cannot always be maintained during battle) and also complements the move towards more local and distributed computing, such as 5G networks.[13] As decentralized swarms have no single point of failure, they are more robust against electromagnetic weapons that can disable, alter or take over weapon systems.

Conceivable examples of missions for swarms include overpowering enemy air defences, overwhelming enemy fighter aircraft in dogfights, engulfing warships, reconnaissance over large areas or urban areas, forming nets of underwater mines, and functioning as anti-access/area denial systems (known as A2/AD systems).[14]

The most novel and thus conceptually interesting types of swarms are also the most problematic. Decentralized swarms are the most resistant to communications disruptions but are more unpredictable than centralized swarms. Swarms comprising simple and cheap systems offer the sought-after mass and robustness, but also have less advanced sensors, communications equipment and processing power, which would increase their weight and their cost. This unpredictability and simplicity, respectively, makes human control over the swarm more difficult.

IV. HUMAN CONTROL OVER MACHINES

This section highlights the known challenges to human control over machines, not exclusive to swarms. It combines the literature on human–machine interaction and on MHC to detail the various problems that have been identified in asserting human control over a machine and in finding the right level of control, and the solutions offered.

[9] Brambilla, M., et al., 'Swarm robotics: a review from the swarm engineering perspective', *Swarm Intelligence*, vol. 7 (2013), pp. 2–3.

[10] Although speaking of the 'behaviour' of a machine risks anthromorphization, this terminology is widely used in the swarm robotics literature and will therefore be employed here.

[11] Barca and Sekercioglu (note 7), p. 348; Yogeswaran, M. and Ponnambalam, S., 'Swarm robotics: an extensive research review', ed. I. Fuerstner, *Advanced Knowledge Application in Practice* (IntechOpen: London, 2010), p. 259.

[12] Scharre, P., *Robotics on the Battlefield Part II: The Coming Swarm*, Center for a New American Security (CNAS) report (CNAS: Washington, DC, Oct. 2014).

[13] Schneider, W., 'Transcript: The future of warfare: gaining a tactical edge through cloud and AI', Hudson Institute, 30 May 2019.

[14] Brehm and de Courcy Wheeler (note 4), p. 3.

Factors affecting human control over a machine

Human control over a machine can be eroded by many different factors. To start, operating a complex military system can be cognitively demanding and many missions involve long working hours. When attention drops, so does readiness and critical thinking. On top of that, operators are subject to stress, exhaustion, fear and boredom, any of which alone or in combination degrades their performance.[15] Control systems for automated weapons need to facilitate active deliberate reasoning by operators to avoid the risk of the operator controlling the weapon on a 'mental autopilot' by neglecting ambiguity, doubt and conflicting evidence.[16] Moreover, humans are susceptible to both undertrust and overtrust in machines which, especially in the heat of the battle, can lead to errors in judgement.[17]

Design choices also significantly influence the opportunity for control. This obviously includes which functions are autonomous, but control goes beyond that: the physical set-up is also critical. For instance, the Aegis naval weapon system requires a human to insert a physical key before a missile can be fired, and deployment on each new mission involves commanders creating a specific doctrine with possible operational modes, setting limits on how the Aegis system can be used in specific situations. In contrast, operators controlling the Patriot surface-to-air missile system can choose between several pre-programmed modes on a display, and have thus much less flexibility in tailoring their options to the mission.[18] It has been argued that the Aegis system aims to capture the mission commander's intent, while the Patriot system embodies the intent of the designers and testers because the latter allows for much less control.[19]

Another factor is the interface between the human and the machine. R&D is ongoing into interfaces that use natural language, gestures and touch to display information and receive instructions, making it easier and more intuitive for humans to control the machines. However, fundamental challenges remain as machines are currently not capable of semantic understanding, and it is disputed whether they will ever become able to grasp an operator's goals or intent in the foreseeable future.[20]

The right level of control

The systems engineering literature differentiates between two types of control: direct and supervisory control. Supervisory control requires less intensive attention and provides more contextual awareness, so operators can better control multiple systems at once. While supervisory control decreases operator fatigue, it can increase operator boredom and attention loss. Switching the level of control if an emergency arises is especially problematic, as it takes time to assess the state of the system and the problem in detail. This can lead to rushed decisions and errors if time is not available. Moreover, a fully autonomous or fully manually controlled system is easier to design than a system with intermediate or mixed levels of control.[21]

An important aspect of the discussions on LAWS is exactly what role humans should have when operating systems with autonomous functions; the question is more complex than whether humans or machines push the button to release the payload. Even when human operators might make the executive decision to strike, there are risks: they may not be meaningfully engaged in the operation, they may lose situational awareness, or they may not critically assess whether they should take a machine-recommended action. Such sitiuations could lead to violations of IHL and erode the legal and moral responsibility of operators.[22] Integrating a positive obligation to ensure MHC over a weapon has therefore arisen as a potential solution in the discussions on LAWS.

MHC is an intuitive concept that is broadly appealing.[23] Unresolved, however, is the question of how MHC should be accomplished, and what its

[15] Cummings, M. L., et al., 'Boredom and distraction in multiple unmanned vehicle supervisory control', *Interacting with Computers*, vol. 25 (2013), pp. 34–47.

[16] Sharkey, N., 'Staying in the loop: human supervisory control of weapons', eds N. Bhuta et al., *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press: Cambridge, UK, 2016), pp. 30–34.

[17] Cummings, M., 'Automation and accountability in decision support system interface design', *The Journal of Technology Studies*, vol. 32 (2006).

[18] Scharre, P., *Army of None: Autonomous Weapons and the Future of War* (W. W. Norton: New York, 2018), pp. 163–72.

[19] Scharre (note 18), p. 165

[20] Chen, J. and Barnes, M., 'Human–agent teaming for multirobot control a review of human factors issues', *IEEE Transactions on Human-Machine Systems*, vol. 44 (2014), p. 14.

[21] Mindell, D., 'Driverless cars and the myths of autonomy', Huffington Post, 14 Oct. 2015 (updated 6 Dec. 2017).

[22] Moyes (note 1).

[23] Crootof, R., 'A meaningful floor for meaningful human control', *Temple International and Comparative Law Journal*, vol. 30, no. 1 (2016), pp. 53–62.

exact goal should be.[24] Not only is there no agreement about what exactly constitutes 'meaningful' and what exactly should be controlled, but the concept itself is used very differently in the literature. For example, the term MHC is used varyingly by scholars: to describe whether humans are in, on, or out of the decision-making loop; to describe at which stages of the targeting cycle humans can interfere; to mean a legal mechanism for ensuring humans will maintain legal responsibility for the mission outcomes; or to refer to a design principle for maintaining moral responsibility.[25] The disagreement over the ontological nature of the concept makes it difficult to work with. Additionally, key states, including Russia and the United States, oppose concentrating on MHC in the discussions on LAWS; Russia considers it politicizing and the USA considers it divisive.[26] Nonetheless, the rising prominence of the concept highlights the growing consensus that human control is more than merely 'pressing the red button', and that operators need to remain critically engaged with the operation.

Ways of ensuring human control over weapon systems

Human control over a weapon system can be ensured across its life-cycle: from initial planning, through R&D and certification, to deployment. To start, countries should assess, through legal reviews of new weapons and new means and methods of warfare, whether the level of human involvement in the planned design would violate IHL.

Additionally, human control can be maximized in the R&D stage through different approaches to design. Control-by-design suggests systems should be devised in such a way that operators are sufficiently informed about the state of the system and its context before making decisions, and they can provide input during specific moments in the targeting cycle.[27] Value-sensitive design commences with stakeholder engagement to identify the values most important to them, and building those values into the system from the earliest stages.[28] A function allocation approach ensures that each step in the targeting cycle is identified and specifically assigned to either a human or a machine.[29] Machines should be tested extensively during the R&D phase, with frequent feedback from actual users in the field, in near-realistic conditions.

There should be an extensive verification and validation (V&V) process to ensure that machines meet the requirements and specifications and that they fulfil their intended purpose. For example, only systems that have been proven to perform better than humans at complex safety-critical tasks could be certified.[30]

Once the weapon has become part of the arsenal, military planners can decide on unique limits on which modes to use (with which levels of automation), or the geographic or temporal scope of the weapon.[31] Before launch, legal advisers should assess the legality of use in the specific context to ensure it does not violate IHL. While a system is operational, it could potentially log every decision it makes, to allow for a post-use assessment by its operators.[32] The opportunities for control thus cover a broad spectrum.

V. CHALLENGES FOR HUMAN CONTROL OVER AND INTERACTION WITH SWARMS

This section presents the challenges to human control for swarms specifically. It starts by setting out how swarms are controlled, followed by the challenges to human–swarm interaction, and the unpredictable nature of emergent swarms. Above all, there are multiple ways to design swarms, and not all problems

[24] Ekelhof, M., 'Moving beyond semantics on autonomous weapons: meaningful human control in operation', *Global Policy*, vol. 10, no. 3 (Sept. 2019).

[25] Neslage, K., 'Does "meaningful human control" have potential for the regulation of autonomous weapon systems?', *University of Miami National Security and Armed Conflict Law Review*, vol 6, (2015), pp. 164–167; Sharkey (note 16); Chengeta, T., 'Defining the emerging notion of "meaningful human control" in weapon systems', *International Law and Politics*, vol. 49 (2017), p. 838; Santoni de Sio, F., and van den Hoven, J., 'Meaningful human control over autonomous systems: a philosophical account', *Frontiers in Robotics and AI*, vol. 5, no. 15 (2018), pp. 11–12.

[26] Reaching Critical Will, 'News in brief', *CCW Report*, vol. 6, no. 8 (29 Aug. 2016), p. 4; Acheson, R., 'Effectuating *our* intention', *CCW Report*, vol. 6, no. 10 (31 Aug. 2016), p. 1.

[27] Dahlmann, A. and Dickow, M. (eds), *Focus on Human Control*, International Panel on the Regulation of Autonomous Weapons (iPRAW) working paper (iPRAW: Aug. 2019), p. 12.

[28] Verdiesen, I., 'How do we ensure that we remain in control of our autonomous weapons?', *AI Matters*, vol. 3, no. 3 (2017), pp. 49, 51–52.

[29] Canellas, M. and Haga, R., 'Toward meaningful human control of autonomous weapons systems through function allocation', 2015 IEEE International Symposium on Technology and Society (ISTAS), 11–12 Nov. 2015, Dublin, Ireland.

[30] Cummings, M., 'Lethal autonomous weapons: meaningful human control or meaningful human certification?', draft paper, n.d.

[31] Roorda, M., 'NATO's targeting process: ensuring human control over (and lawful use of) 'autonomous' weapons', eds P. Scharre and A. Williams, *Autonomous Systems: Issues for Defence Policymakers* (NATO Allied Command Transformation: Norfolk, VA, 2015).

[32] Feldman, P., Dant, A. and Massey, A., 'Integrating artificial intelligence into weapon systems', *ArXiv*, 10 May 2019.

apply to all swarms. The most pressing concerns are about decentralized and especially emergent swarms. Nonetheless, there are several overarching challenges commonly mentioned in the swarm robotics literature that are worth highlighting.

Controlling a swarm

Controlling a swarm is complicated. There are four main avenues to control: switching between the algorithms that specify the behaviour of the swarm; changing the parameters of a swarm control algorithm; remote control of specific nodes (leaders); and altering the environment to influence the swarm's behaviour.[33] Each of these four options is discussed below.

First, operators can switch between different pre-programmed packages of algorithms that will lead to the desired swarm behaviour.[34] Notably, the US Naval Postgraduate School is developing 'playbooks' with predefined tactics from which operators can choose different specified behaviours. Using such playbooks will likely be the most common method for operators to control swarms on the battlefield.[35]

Second, an operator can change the parameters of a swarm control algorithm. Such parameters might include the radius of the swarm, the maximum or minimum distance between nodes, or the 'personality' of nodes to influence voting behaviour of swarms that operate under consensus. By changing how the nodes interact with each other and the environment at an individual level, different behaviour patterns are achieved on a group level.[36] However, parameter setting is mostly done during the R&D phase because of the unpredictability involved.[37]

Third, operators can control the actions of a selected swarm member through continuous input (tele-operation) or by sending intermittent messages. The selected node can send commands to the other nodes, or the other nodes might respond to its actions or signals. Tele-operation is only possible when there are few to no delays in the signal transmission (low latency) and with continuous updates about the state of the swarm.[38]

Finally, operators can alter the environment that swarms operate in to influence their behaviour. This is generally achieved through the introduction of virtual agents such as virtual 'pheromones' that tell nodes to (not) further explore an area; beacons that send spatially bound signals to direct a subset of the swarm; and proxies, attractors or predators, which are simulated nodes or adversaries to which the nodes respond.[39] Altering the environment requires an intimate understanding of how a swarm might respond and is less suited for novice operators.[40]

Critically, many of the avenues to control rely on indirect effects. The collective entity of the swarm is an abstract concept that does not exist on the physical level where only the nodes truly exist. An operator cannot thus command or program the swarm as such. An operator using tele-operation, parameter changes, or proxies, attractors or predators, does not directly send commands at the group level to display a specific type of behaviour. Instead, this behaviour results from changes in how nodes act, or responses to other nodes or the environment. Even libraries are not created by formally modelling a direct command at the swarm level; rather, swarm engineering is typically done through endless experiments in which programmers make small iterative changes in the algorithms to find the settings that lead to the desired group behaviour.[41]

Interacting with a swarm

Swarms are difficult to operate in practice. Foremost, the cognitive complexity of operating a swarm is high, and increases the greater the swarm size. With each added node it becomes harder to keep track of the swarm, and the amount of possible interactions within the swarm and with the environment increases

[33] Kolling, A., et al., 'Human interaction with robot swarms: a survey', *IEEE Transactions on Human–Machine Systems*, vol. 46, no. 1 (2016), p. 8.

[34] Kolling et al., 'Human interaction with robot swarms: a survey' (note 33), p. 8.

[35] Giles, K. and Giammarco, K., 'Mission-based architecture for swarm composability (MASC)', *Procedia Computer Science*, vol. 114 (2017), pp. 59–60.

[36] Kira, Z. and Potter, M., 'Exerting human control over decentralized robot swarms', 2009 4th International Conference on Autonomous Robots and Agents, 10–12 Feb. 2009, Wellington, New Zealand, pp. 566–69.

[37] Kolling et al., 'Human interaction with robot swarms: a survey' (note 33), pp. 8–9.

[38] Kolling et al., 'Human interaction with robot swarms: a survey' (note 33), pp. 9–11.

[39] Kolling et al., 'Human interaction with robot swarms: a survey' (note 33), p. 9.

[40] Kolling, A., et al., 'Human–swarm interaction: an experimental study of two types of interaction with foraging swarms', *Journal of Human–Robot Interaction*, vol. 2 (2013), p. 125.

[41] Brambilla, M., 'Formal methods for the design and analysis of robot swarms', PhD thesis (Universite Libre de Bruxelles, Berlin, Heidelberg, 2014), p. 14.

exponentially (the so-called 'state space explosion problem' in the technical literature). This can lead to a loss of situational awareness when the workload becomes too high—or too low, when too much work is automated.[42] Using commands that rely on indirect effects by controlling individual nodes also becomes more difficult as the size of the swarm increases.[43]

Operators do not always know the state of the swarm at any particular moment, and it is challenging to predict exactly how a swarm will respond to commands. To control a swarm effectively and choose the right control method at the right time, the operator must have a very good mental picture of the state of the swarm, what different swarm behaviours look like, and the likely outcome of the command they are about to give. In practice operators struggle to accurately perceive such a picture.[44] They often base their decisions on the physical characteristics of the swarm, but the swarm's non-linear dynamics means that these do not always indicate its actual performance.[45] This lack of understanding makes it very difficult to predict the exact impact of a command.[46] To further complicate matters, the local interactions among the nodes in a consensus-based swarm cannot be readily perceived by humans during its formation, so it is challenging for operators to measure the swarm's progress towards its goal.[47]

Additionally, communications constraints are common. Bandwidth is more limited, and latency and asynchrony (when signals between the swarm and the operator are out of sync) are higher in swarms than in other types of systems.[48] Real-time updates of all individual nodes in a large swarm demands a lot of bandwidth, while communicating the status through a leader takes time and causes latency, especially in hierarchical set-ups. It is therefore not possible to expect 'reliable communication with each swarm entity in most situations, due to environmental and technological limitations'.[49]

This latency is significant because timing of commands can be vital. Swarms can take quite some time to reach the desired state following a command, as they need to coordinate internally and take the desired position.[50] Ill-timed commands can have different or even adverse effects depending on the state of the swarm. Commands thus need to go out at exactly the right time to have the desired effect. For optimal control, an operator needs to embrace the concept of 'neglect benevolence', which is the idea that it can be more beneficial to wait and not give instructions to (neglect) the swarm.[51] For example, multiple instructions to join together sent to a semi-fragmented swarm could lead to further fragmentation, and it is better to not take action sometimes. This is a highly counterintuitive method of operation.[52] Swarms thus may not be suitable for time-critical applications requiring constant control or applications that require time-critical decisions.[53] Many military applications would fall into these categories.

Predicting and testing emergence

What makes swarms unique is their capacity for 'emergence'—the complex collective behaviour that arises from the behaviour of the individual nodes. An example in nature are the flocking patterns that birds can form. Emergent patterns are not programmed at the individual level, nor can they be readily explained

[42] Hussein, A. and Abbass, H., 'Mixed initiative systems for human–swarm interaction: opportunities and challenges', 2018 2nd Annual Systems Modelling Conference, 4 Oct. 2018, Canberra, Australia, pp. 2–3.

[43] Brown, D., Kerman, S. and Goodrich, M., 'Human–swarm interactions based on managing attractors', *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction* (Bielefeld: ACM Press, 2014), p. 90.

[44] Roundtree, K. A., Goodrich, M. A. and Adams, J. A., 'Transparency: transitioning from human–machine systems to human–swarm systems', *Journal of Cognitive Engineering and Decision Making*, vol. 13, no. 3 (2019), pp. 187–91.

[45] C. Nam, et al., 'Models of trust in human control of swarms with varied levels of autonomy', *IEEE Transactions on Human–Machine Systems*, early access, 25 Feb. 2019, p. 10.

[46] Walker, P., Lewis, M. and Sycara, K., 'Characterizing human perception of emergent swarm behaviors', 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 9–12 Oct. 2016, Budapest, p. 2436.

[47] Nagavalli, S., et al., 'Bounds of neglect benevolence in input timing for human interaction with robotic swarms', *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction – HRI '15* (ACM Press: Portland, OR, 2015), pp. 197–98.

[48] Kolling et al., 'Human interaction with robot swarms' (note 33), p. 6.

[49] Harriott, C. E., et al., 'Biologically-inspired human-swarm interaction metrics', *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting*, vol. 58 (2014), p. 1472.

[50] Kolling et al., 'Human interaction with robot swarms' (note 33), p. 12.

[51] Kolling et al., 'Human interaction with robot swarms' (note 33), p. 12.

[52] Walker, P., et al., 'Neglect benevolence in human control of swarms in the presence of latency', 2012 IEEE International Conference on Systems, Man, and Cybernetics, 14–17 Oct. 2012, Seoul, pp. 3009–14.

[53] Ahmed, H. and Glasgow, J., *Swarm Intelligence: Concepts, Models and Applications*, Technical report (Queen's University: Kingston, 2012), p. 34.

at the system level. Emergence results from the internal interactions between nodes within the swarm, and the interactions between the nodes and the environment.[54] It is one of the major benefits of using a swarm, as highly complex tasks can be achieved using relatively simple programming and technology.[55]

However, in most cases it is hard to determine what individual nodal behaviour corresponds to specific emergent behaviour. Emergence 'arises on a higher level of abstraction' and is hard to predict or control.[56] Some researchers claim that it might even be difficult to effectively control a swarm once it starts operating, or to abort a mission if a swarm acts in a dangerous or unpredicted way.[57] This is even more problematic in scenarios where emergent swarms would interact with other emergent swarms.

Engineers model swarms in two ways: top-down, designing the behaviour of the swarm as a whole from which the system derives rules to achieve this; and bottom-up, programming the behaviour of the swarm at the level of the individual node. It is possible to validate and verify that all individual nodes meet the specifications and fulfil their purpose. However, much programming is done ad hoc, and formal methods to mathematically ensure that a node will act in the prescribed way are not so common in swarm engineering.[58] Moreover, the behaviour of the swarm as a collective cannot be validated and verified based on the programming of the individual nodes.[59] However, classic control models designed to prove the properties of a swarm at a collective level often do not hold up in practice for reasons of asynchronicity, the element of randomness (stochasticity) programmed in all nodes to ensure that they do not all take the same actions at the same time, and the lack of global information on the state of the swarm.[60] Nonetheless, V&V is not fundamentally impossible; the field is actively researching this problem, so it might be solved in the future.[61]

In addition, debugging an emergent swarm is inexpedient. It is difficult to detect and recreate errors because of the inherent state space explosion in emergent swarms.[62] Most swarms are tested in either simulations or highly controlled laboratory environments—without the noise and environmental conditions that drive much of the unpredictability in emergent behaviour.[63] Additionally, good metrics and testbed applications are lacking.[64]

VI. POLICY IMPLICATIONS OF SWARM CONTROL CHALLENGES

The technical literature has clearly set out why swarm control is problematic from an engineering standpoint. This section sets out the strategic, legal and ethical implications for the military use of swarms.

Strategic implications

Swarms pose several strategic problems. First, swarms might not be fully predictable and reliable, and cannot be depended on to execute missions in accordance with their operator's intentions. Playbooks provide an array of actions to take, but they reflect the intention of the designers, not operators. This is tactically undesirable and limits opportunities to change course during the mission if the situation changes. Second, if operators already struggle to parse the state of a swarm, this is even harder for adversaries. The risks of misperception are high, leading to conflict escalation. Unintentional engagements would be the worst outcome, but the risks are not limited to combat swarms. Non-combat swarms might accidentally intrude into foreign airspace, which adversaries might perceive as an attack. Adversaries likely also cannot assess whether a swarm is armed, conducting ISR or performing another non-combat function. Third, the limited opportunities to alter the behaviour of a swarm after launch, and the reliance

[54] Liu, Q., et al., 'A mechanism for recognizing and suppressing the emergent behavior of UAV swarm', *Mathematical Problems in Engineering* (2018), p. 6.

[55] Harvey, J., 'The blessing and curse of emergence in swarm intelligence Systems', eds H. A. Abbass, J. Scholz, and D. J. Reid, *Foundations of Trusted Autonomy* (Springer International Publishing: Cham, 2018), p. 119.

[56] Kolling et al. 'Human–swarm interaction' (note 41), p. 104.

[57] Brambilla et al. (note 9), p. 36.

[58] Lopes, Y., et al., 'Supervisory control theory applied to swarm robotics', *Swarm Intelligence*, vol. 10 (2016), p. 66.

[59] Kumar, M. and Ramakrishnan, S., *Modeling and Analysis of Stochastic Dynamics and Emergent Phenomena in Swarm Robotic Systems Using the Fokker-Planck Formalism* (Defense Technical Information Center: Fort Belvoir, 29 Oct. 2010).

[60] Brambilla et al. (note 9), p. 14.

[61] Winfield, A. F. T. and Nembrini, J., 'Safety in numbers: fault-tolerance in robot swarms', *International Journal of Modelling, Identification and Control*, vol. 1 (2006).

[62] Rouff, C., et al., 'Verification of emergent behaviors in swarm-based systems', *Proceedings: 11th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems, 2004*, 27 May 2004, p. 444.

[63] Brambilla et al. (note 9), p. 14.

[64] Brambilla (note 41), p. 46.

on indirect control methods, especially in the case of emergent swarms, can lead to adverse incidents—crashes, fratricides (so-called 'friendly fire') and even civilian casualties.

Legal implications

Swarms also cause legal problems in regard to compliance with all principles of IHL. The first problem is compliance with the principle of precaution in attack. Ordinarily legal advisers ensure an attack complies with IHL before any weapon is launched. Operators of drones or aircraft assess whether the situation has changed before the final engagement. But the latency and bandwidth challenges and relatively simple sensors that swarms use mean that assessment of the situation might be more difficult for swarm operators. And if a swarm comes up with a strategy to execute a mission itself, it cannot be guaranteed that the mission will comply with all principles of IHL.

The second problem is the risk of violating the principle of distinction, by posing a threat to civilians. During swarm engagements, it is likely that some nodes will be shot down and left on the battlefield. If they contain unexploded ordnance, this could kill civilians.[65] They could also make land unavailable for use by local communities, and pollute the environment if they leak toxic material.

Ethical implications

The final set of problems is ethical. First, the use of swarms risks eroding the moral responsibility of operators. The current understanding of responsibility is centred around being in full control of our actions, as envisioned through the relationship between one human and one machine.[66] Using a high number of systems simultaneously—especially in a networked setting—risks diffusing the sense of human responsibility for the outcomes of a mission. Second, the increased complexity of controlling systems with autonomous functions could place an unfair burden on operators, as they would be held accountable for the outcome of missions.[67] This logic can be easily extended to swarms. Swarms are difficult to control and cognitively demanding, but operators would nonetheless be held accountable for the outcome of a mission if something goes terribly wrong. It is essential that humans will be held completely accountable for mission outcomes of systems with autonomous functions, but this situation raises the ethical dilemma of whether it is just to place that burden on an operator struggling to fully control their weapon.

VII. THE STATE OF MILITARY R&D ON SWARMS IN EUROPE

Background on European defence research

Within the framework of the EU, military swarm R&D falls into three categories: dual-use research, member-states-driven defence research and EU-driven defence research (see table 1.1).

Dual-use research on swarms has been conducted under the 7th Framework Programme for Research and Technological Development (FP7) in 2007–13 and Horizon 2020 (H2020) in 2014–20 of the EU.[68]

EU member states have been jointly financing and executing military-oriented research and technology (R&T) initiatives under the coordination of the European Defence Agency (EDA) since 2005.[69] This has been taken to the next level under the Permanent Structured Cooperation (PESCO) since 2018.[70]

Since 2015 the EU has also funded defence research directly, managed by the EDA, commencing with the Pilot Project (PP) on defence research in 2015–18. EuroSWARM was one of the three projects included in the PP, indicating the importance placed on swarm research. Current EU-funded defence research programmes include the Preparatory Action on Defence Research (PADR) for R&D at lower technology readiness levels (TRLs), in 2017–20; and the European Defence Industrial Development Programme (EDIDP) to cover the capability development phase and support the European defence industry, in 2019–20.

[65] Homayounnejad, M., 'Autonomous weapon systems, drone swarming and the explosive remnants of war', TLI Think! Paper no. 1 of 2018 (King's College London, 6 Feb. 2018).

[66] Coeckelbergh, M., 'From killer machines to doctrines and swarms, or why ethics of military robotics is not (necessarily) about robots', *Philosophy & Technology*, vol. 24, no. 3 (2011), pp. 274.

[67] Johansen, S. R., 'Technological change and international law: amend, implement, or simply manage expectations?', Paper presented at the 2019 Stockholm Security Conference, Stockholm, 3 Oct. 2019.

[68] Some FP7 programmes continue under H2020. For details see European Commission, 'Research and Innovation funding 2014–20', n.d.; see also European Commission, 'Horizon 2020', n.d.

[69] European Defence Agency, 'Research & Technology', n.d.

[70] European Defence Agency, 'Permanent Structured Cooperation (PESCO)', n.d.

PADR's flagship project is Ocean2020, a €35 million undertaking to improve interoperability between manned and unmanned systems, for example by increasing autonomy for swarms. For reference, the total PADR budget is €90 million. PADR and EDIDP will be subsumed under the European Defence Fund (EDF) in the next EU Multiannual Financial Framework, in 2021–27. The EDF will also co-fund PESCO projects if they meet the requirements for transnational cooperation. However, the new EC and EP still need to approve the final budget and scope of the EDF.[71]

The scope of the R&D programmes is determined by the Capability Development Plan (CDP), developed by EU member states, which feeds into the Coordinated Annual Review on Defence (CARD). The CDP sets out the current capability shortfalls and long-term technology trends, and guides the priorities set under PESCO and the EDF.[72] However, the exact decision-making process is relatively opaque. The EP has ensured that LAWS are ineligible for funding by the EDF (as are other technologies prohibited by international law) but beyond that, has waived their right to parliamentary scrutiny of EDF spending, other than through annual reports and a complete evaluation of the project in 2027.[73] PESCO projects, in contrast, fall outside the competence of the EC and EP, and are regulated by the Council of the EU, which represents the executive governments of EU member states.[74]

In the current EC (1 December 2019), a new directorate general (DG) will be created called DG Defence Industry and Space (DG Defence). This will fall under the portfolio of the Commissioner for Internal Market and Services, and subsequently of the Commissioner for the Digital Agenda, not the EEAS.[75] This positioning risks that defence–industrial concerns—such as supporting the European defence industry—will be given priority over strategic or tactical concerns about their utility for Europe, and over humanitarian concerns too.

The prominence of swarm research

Swarm research has been part of every single programme and is a priority for the EDF, as evidenced by the prominent positions of EuroSWARM and Ocean2020. For many programmes, details are scarce about the exact nature of the research conducted. Research on human control is included, although the importance of the issue differs according to the project. LAWS are not eligible for funding and the EU has not directed R&D on armed swarms.[76] Nonetheless, human control is still important to ensure safety and prevent conflict escalation, and swarms could be armed at a later stage. Supplementary documentation on swarm R&D programmes executed either under the auspices of the EU or European countries or defence companies that have been gathered by the author and analysed for the purpose of this paper can be found at the EU Non-Proliferation and Disarmament Consortium (EUNPDC) website.[77] The next paragraph discusses three examples of swarm research projects: ROBORDER, EUROswarm and ASIMUT.

ROBORDER is a H2020 project to develop a heterogeneous and autonomous swarm of unmanned systems for use in border surveillance. The goal is to detect criminal activities and threatening situations.[78] Most of the papers published by the consortium of research organizations deal with communications technology, although one paper researched the use of augmented and virtual reality for swarm control interfaces.[79] A leaked internal ethical review limited itself to the risk of proliferation—which was ultimately dismissed—but did not discuss the ethical considerations of EU member states using this technology.[80] Similarly, the Ocean2020 project

[71] European Defence Agency, 'Our current priorities', n.d.

[72] Fiott, D., *EU Defence Capability Development: Plans, Priorities, Projects,* European Union Institute for Security Studies (EUISS), Issue Brief no. 6 (EUISS: Paris, June 2018), pp. 1–2.

[73] Csernatoni, R. and Martins, B. O., *The European Defence Fund: Key Issues and Controversies*, PRIO Policy Brief no. 3 (PRIO: Oslo, 2019).

[74] Fiott, D., *The Scrutiny of the European Defence Fund by the European Parliament and National Parliaments*, Policy Department for External Relations, European Parliament Directorate-General for External Policies, Study Paper, PE 603.478, Apr. 2019, p. 10.

[75] Von der Burchard, H. and Winfield, M, 'Ursula von der Leyen's actual org chart', Politico, 11 Sept. 2019.

[76] European Commission, *European Defence Industrial Development Programme (EDIDP): 2019 Calls for Proposals, Conditions for the Calls and Annex* (European Union: Brussels, 4 Apr. 2019).

[77] See supplementary documentation: Verbruggen, M., 'European research on military swarms', <https://www.nonproliferation.eu/activities/online-publishing/non-proliferation-papers/>.

[78] ROBORDER, 'Aims and objectives', n.d.

[79] Helin, K., et al., 'AR / VR based novel user interface for drone swarms mission control', Poster presented at the European Association for Virtual Reality and Augmented Reality Conference, 22–23 Oct. 2019, London, UK.

[80] Campbell, Z., 'Swarms of drones, piloted by artificial intelligence, may soon patrol Europe's borders', The Intercept, 11 May 2019.

Table 1.1. European Union defence research programmes

Duration	Programme	Description
2007–13	7th Framework Programme for Research and Technological Development	EU funding for dual-use research
2014–20	Horizon 2020	EU funding for dual-use research
2007–	Research and Technology	Joint defence research by EU member states
2015–18	Pilot Project	Pilot project on EU defence research
2017–20	Preparatory Action on Defence	Preliminary EU defence research for lower TRLs
2018–	Permanent Structured Cooperation	Joint defence research by EU member states, co-financed by EDF
2019–20	European Defence Industrial Programme	Preliminary EU defence research for capability development and industry support
2021–27	European Defence Fund	EU defence research programme under new budget cycle; will subsume PADR and EDIDP and co-fund PESCO

EDF = European Defence Fund; EDIDP = European Defence Industrial Development Programme; EU = European Union; PADR = Preparatory Action on Defence Research; PESCO = Permanent Structured Cooperation; TRL = Technology readiness level

Source: Author's own compilation.

has a dedicated page on 'ethical issues', but these only concern human rights and privacy concerns about the research project itself, not how its military systems might be used.[81] EUROswarm, one of three projects in the PP, aimed to test and demonstrate the utility of using swarms for military operations. The project researched a heterogeneous swarm of mobile and static sensors for surveillance purposes, such as observing hostile military camps.[82] The academic output suggests a focus on behaviour monitoring, anomaly detection, communication frequency and task allocation.[83] Human control and human–swarm interaction do not appear to have been pivotal elements of EUROswarm research. However, these elements do feature as a major theme in the ASIMUT project, which was part of the R&T scheme of the EDA. ASIMUT aimed to decrease operator workload during surveillance missions of aerial swarms by enhancing data exploitation by drones.[84]

Human control and human–swarm interaction are thus topics of concern in some swarm R&D programmes and on individual projects, but do not seem to be structurally integrated as essential components of swarm research in all programmes.

None of these EU projects concern armed swarms at this stage, although one consortium has pitched a PADR proposal to develop a simulation of swarms used for suppression of enemy air defences. This could lead to developing a kinetic version of such a swarm in the future. In an interview, a senior adviser at one of the institutes involved stated that there were 'heated discussions in the consortium about how pronounced the ethical and legal side should be', and that 'while the prequisite is to have a human in the loop for pulling the trigger', there is 'a lot of uncertainty' and complexity involved.[85] The adviser's statements were vague and non-committal, and thus raise questions about the decisions that will be made on the required level of human control over the swarm.

Outside the framework of the EU, some European countries are investing in R&D on combat swarms. The two main projects are the Future Combat Air Systems run jointly by France, Germany and Spain, with an important role for commercial companies Airbus and MBDA; and the Tempest project run jointly by Sweden, Italy and the United Kingdom. Both projects aim to develop optionally manned sixth-generation fighter jets, supported by either or both swarms of unmanned

[81] Ocean2020, 'Ethical issues', n.d.

[82] European Defence Agency, 'Pilot project EuroSWARM and SPIDER activities completed', Press release, 23 Feb. 2018.

[83] See, e.g., Lappas, V., et al., 'Autonomous unmanned heterogeneous vehicles for persistent monitoring', Paper presented at the AIAA Scitech 2019 Forum, 7–11 Jan. 2019, San Diego.

[84] Bouvry, P., et al., 'ASIMUT project: Aid to SItuation management based on MUltimodal, MUltiUAVs, MUltilevel acquisition Techniques', *Proceedings of the 3rd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications* (ACM Press, 2017), pp. 17–20.

[85] Sprenger, S., 'Europeans propose siccing self-learning drone swarms on air defenses', *Defense News*, 22 Oct. 2019.

systems and missiles. Details are scarce, but at least Airbus is known to have recognized the importance of the questions that swarms raise and is in discussion with external experts on the ethical implications of developing such a swarm.[86]

Defence R&D on swarms is thus in full swing across Europe. However, few details are made public on the nature of the programmes. With some exceptions, the articles and reports that have been published do not show a strong pre-occupation with questions about human control and human–swarm interaction, let alone with the legal, ethical or strategic implications.

VIII. RECOMMENDATIONS

This section presents two sets of recommendations to the respective EU agencies working on swarms: the EEAS and DG Defence.

European External Action Service

The EEAS should stress the problems related to swarms in the context of the CCW Convention. While swarms have been mentioned incidentally, they have not yet been the subject of focused diplomatic discussion. At this stage of the negotiations, more in-depth discussions on specific systems and contexts would be helpful. Such discussions can be facilitated through presentations by technical and military experts on what swarms are, how they work and what the human role in the swarm would be.

In these discussions the EEAS should keep two fundamental characteristics of swarms in mind. First, swarms should not be seen as munitions, but as platforms or a system of platforms, and should thus not be judged by the standards set on munitions. Not only are there infinitely more actions possible after launch of a swarm, but the subsystems also respond autonomously to each other. Some type of control after launch is necessary. Second, swarms have two relational levels of autonomy: between the human and the swarm, and between the nodes and the swarm as a whole.[87] Even if humans decide which targets to select and engage, the nodes could have a certain level of freedom in deciding how to execute this. Ultimately, this suggests that a 'level of autonomy' approach would not be the ideal solution for dealing with swarms.

DG Defence

DG Defence has plenty of opportunities to ensure human control is a vital aspect of swarm design. First, all PADR grant proposals require self-assessments on ethical, legal and societal aspects (ELSA), in addition to reviews by a group of independent experts. The expert reviews include assessing the proposal against standards for compliance with IHL, arms control treaties, human rights, data protection and environmental impact, among others. The EDA released a guide to the ELSA self-assessments in March 2019, which included a dedicated section on 'autonomy'.[88] Subsequent versions could provide more in-depth guidance on human control and human–swarm interaction, and list all arms and export control treaties ratified by member states which with these systems must comply. It is not currently known whether PESCO requires similar reviews, but if not, DG Defence could also make that a mandatory requirement for receiving funding from the EDF.

Second, all EU member states are obliged to conduct legal reviews for all new weapons, means and methods of warfare (Article 36 reviews). This requirement will thus apply to any project involving co-financing of EU member states in either the EDA or PESCO, but the EDA guide makes no mention of legal reviews. It is not clear whether Article 36 reviews are required for projects that are only indirectly funded by EU member states through the EDF; nor is it clear, in joint development programmes, whether each member state individually conducts an Article 36 review. This is concerning: while member states are legally obliged to conduct these reviews, in practice many states do not have sufficient mechanisms in place to do so.[89] This could lead to the development of technologies in violation of IHL. Greater clarity on the existing

[86] External expert, Interview with author, 3 Sep. 2019.

[87] Cummings, M., 'Human supervisory control of swarming networks', Paper presented at the 2nd Annual Swarming: Autonomous Intelligent Networked Systems Conference, 2004.

[88] European Defence Agency, Preparatory Action on Defence Research (PADR) Programme, *Guidance on How to Complete Your Self-Assessment on 'Ethics, Legal and Societal Aspects (ELSA)'*, version 1.0, 19 Mar. 2019.

[89] Boulanin, V. and Verbruggen, M., *Article 36 Reviews: Dealing with the Challenges Posed by Emerging Technologies*, Conference Report (SIPRI: Stockholm, Dec. 2017); Boulanin, V. and Verbruggen, M., *SIPRI Compendium on Article 36 Reviews*, SIPRI Background Paper (SIPRI: Stockholm, Dec. 2017).

practices are thus needed, and DG Defence would do well to urge states to fulfil their existing obligations.

Third, DG Defence should make human–swarm interaction and human control a central component of its defence research programme. In 2012 it was shown that human–swarm interaction is an undervalued aspect of swarm robotics, and that research findings are often not reflected in the practical design of swarms.[90] This is despite research in 2009 showing that little work has focused on enabling human control after deployment.[91] Although it is not clear how the situation has developed since then, these statements are troubling and DG Defence could take the lead on research mitigating these issues. For example, the effects of cognitive complexity can be mediated with more intuitive interfaces, and virtual reality can improve situational awareness.[92] Operators might also better predict the impact of their commands by simulating their effects before conveying them to the swarm.[93] Research should thus be conducted on improving interfaces, as well as on finding solutions for V&V, and developing realistic testbed applications.

Finally, it is recommended that the EEAS and DG Defence coordinate and collaborate closely, to ensure that the systems developed by DG Defence comply with IHL and ethics, and that not only industrial, but also humanitarian considerations guide EU defence policy.

IX. CONCLUSIONS

A review of the technical literature on swarm robotics has presented several challenges to human control over swarms. The main challenges are problems with the mode of control, for example, operator reliance on indirect methods of control or algorithmic commands developed by designers, and the ability to maintain control; the fact that operating a swarm is highly cognitively demanding; the susceptibleness of swarms to communication disruptions; and the inherent unpredictability of emergent swarms.

From a policy perspective, these challenges to human control over swarms have strategic, legal and ethical implications for their use in a military context. Strategically, there is no guarantee that swarms will execute missions exactly according to an operator's wishes, risking adverse outcomes such as conflict escalation. Legally, an attack using swarms may not comply with all principles of IHL, while unexploded ordnance may pose a risk to civilians as well as cause pollution. Ethically, controlling a swarm erodes an operator's sense of moral responsibility and at the same time may unfairly lay accountability for the outcome on operators.

Since EU defence programmes have made swarm research a priority, the EEAS should push for in-depth discussion of swarms at the CCW Convention, and DG Defence should focus on ensuring human control is an integral component in its research on swarms.

[90] Kolling, A., Nunnally, S. and Lewis, M., 'Towards human control of robot swarms', *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction – HRI '12* (ACM Press: New York, 2012).

[91] Kira and Potter (note 36), p. 571.

[92] Hocraffer, A. and Nam, C., 'A meta-analysis of human-system interfaces in unmanned aerial vehicle (UAV) swarm management', *Applied Ergonomics*, vol. 58 (2017).

[93] Madey, G., et al., 'Applying DDDAS principles to command, control and mission planning for UAV swarms', *Procedia Computer Science*, vol. 9 (2012).

ABBREVIATIONS

AI	Artificial intelligence
CARD	Coordinated Annual Review on Defence
CCW	1981 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
CDP	Capability Development Plan
DG Defence	Directorate General Defence Industry and Space
EC	European Commission
EDA	European Defence Agency
EDF	European Defence Fund
EDIDP	European Defence Industrial Development Programme
EEAS	European External Action Service
ELSA	Ethical, legal and societal aspects
EP	European Parliament
EU	European Union
FP7	7th Framework Programme for Research and Technological Development
H2020	Horizon 2020
IHL	International humanitarian law
ISR	Intelligence, surveillance and reconnaissance
LAWS	Lethal autonomous weapons systems
MHC	Meaningful human control
PADR	Preparatory Action on Defence Research
PESCO	Permanent Structured Cooperation
PP	Pilot Project
R&D	Research and development
R&T	Research and technology
TRL	Technology readiness level
V&V	Verification and validation

This document has been produced with the financial assistance of the EU. The contents are the sole responsibility of the EU Non-Proliferation and Disarmament Consortium and can under no circumstances be regarded as reflecting the position of the EU.

A EUROPEAN NETWORK

In July 2010 the Council of the European Union decided to support the creation of a network bringing together foreign policy institutions and research centers from across the EU to encourage political and security-related dialogue and the long-term discussion of measures to combat the proliferation of weapons of mass destruction (WMD) and their delivery systems. The Council of the European Union entrusted the technical implementation of this Decision to the EU Non-Proliferation Consortium. In 2018, in line with the recommendations formulated by the European Parliament the names and the mandate of the network and the Consortium have been adjusted to include the word 'disarmament'.

STRUCTURE

The EU Non-Proliferation and Disarmament Consortium is managed jointly by six institutes: La Fondation pour la recherche stratégique (FRS), the Peace Research Institute Frankfurt (HSFK/ PRIF), the International Affairs Institute in Rome (IAI), the International Institute for Strategic Studies (IISS), the Stockholm International Peace Research Institute (SIPRI) and the Vienna Center for Disarmament and Non-Proliferation (VCDNP). The Consortium, originally comprised of four institutes, began its work in January 2011 and forms the core of a wider network of European non-proliferation and disarmament think tanks and research centers which are closely associated with the activities of the Consortium.

MISSION

The main aim of the network of independent non-proliferation and disarmament think tanks is to encourage discussion of measures to combat the proliferation of weapons of mass destruction and their delivery systems within civil society, particularly among experts, researchers and academics in the EU and third countries. The scope of activities shall also cover issues related to conventional weapons, including small arms and light weapons (SALW).

www.nonproliferation.eu

EU Non-Proliferation and Disarmament Consortium

Promoting the European network of independent non-proliferation and disarmament think tanks

FONDATION pour la RECHERCHE STRATÉGIQUE

FOUNDATION FOR STRATEGIC RESEARCH

www.frstrategie.org

PRIF HSFK
Peace Research Institute Frankfurt
Leibniz-Institut Hessische Stiftung Friedens- und Konfliktforschung

PEACE RESEARCH INSTITUTE FRANKFURT

www.hsfk.de

iai Istituto Affari Internazionali

INTERNATIONAL AFFAIRS INSTITUTE

www.iai.it/en

IISS

INTERNATIONAL INSTITUTE FOR STRATEGIC STUDIES

www.iiss.org

sipri

STOCKHOLM INTERNATIONAL PEACE RESEARCH INSTITUTE

www.sipri.org

VCDNP
Vienna Center for Disarmament and Non-Proliferation

VIENNA CENTER FOR DISARMAMENT AND NON-PROLIFERATION

www.vcdnp.org