From the Autonomy Framework towards Networks and Systems Approaches for 'Autonomous' Weapons Systems

Preprint · May 2019

DOI: 10.1163/18781527-01001010

1 author:

Hin-Yan Liu
University of Copenhagen
28 PUBLICATIONS **72** CITATIONS

SEE PROFILE

BRILL
NIJHOFF

JOURNAL OF INTERNATIONAL HUMANITARIAN LEGAL STUDIES 10 (2019) 89-110

INTERNATIONAL HUMANITARIAN LEGAL STUDIES

brill.com/ihls

From the Autonomy Framework towards Networks and Systems Approaches for 'Autonomous' Weapons Systems

Hin-Yan Liu
Associate Professor and Coordinator, Artificial Intelligence and Legal Disruption Research Group, Faculty of Law, University of Copenhagen; Visiting Research Professor, Faculty of Law, University of Otago; Member of the Editorial Board of the Journal
hin-yan.liu@jur.ku.dk

Abstract

The legal debate surrounding the development and deployment of autonomous weapons systems (AWS) has stagnated in recent years, having arguably hit the hard limits of legal doctrine. At the heart of this impasse lies the focus upon autonomy as both the innovative and defining feature of AWS. Thus, the autonomy of the weapons system places it in a legally liminal zone between agent and object, revealing a set of legal problems that revolve around issues of control, influence, responsibility and liability, and questions of legal compliance that follow from the prospect of autonomous lethal decision-making.

This paper seeks to explore alternative framings to the same underlying technology as a means of escaping the limits imposed by the autonomy framework that has dominated the debate to date, and to examine the consequences that flow from pursuing these approaches from legal and regulatory perspectives. In particular, emphasis is placed upon the networks approach, and the systems approach, which this paper sets out and differentiates from the orthodox emphasis upon autonomy. These alternative approaches suggest that the legal problems arising from the autonomy framing are the easiest set of issues to address, insofar as these frame legal problems, while the networks and systems approaches seem to touch upon legal mysteries to which no ready legal or regulatory responses can be made. Rather than dismiss the network and systems approaches, however, this paper suggests that appropriate, adequate and robust legal and regulatory responses must consider the insights and challenges that these approaches pose, and that pursuing these approaches will lead to powerful converging arguments supporting a moratorium on the deployment of AWS.

 | DOI:10.1163/18781527-01001010

Keywords

autonomous weapons systems – killer robots – legal mysteries – legal problems – networks approach – systems approach – sociotechnical change

1 Introduction and Rationale

Recent technological advances are setting the stage for the deployment of weapons systems that are capable of operating in the battlespace and beyond as supporters, partners, coordinators and commanders of human combatants, rather than as the mere implements of violence. The interface between the technological capacities that these technologically advanced weapons systems introduce, and the legal, ethical, and policy environment which seeks to govern their use and application, has led to a lively debate in recent years. Centred upon the notion of *autonomous* weapons systems (AWS),[1] these discussions have sought to address the regulatory challenges posed by the prospect of advanced weapons systems operating autonomously in the battlespace. This framework poses a range of legal questions which the ensuing debate has sharpened into several points of unresolved tension. As these debates continue to drill down into the legal problems posed, these problems are beginning to hit the limits of legal doctrine and are thus beginning to stagnate.

To reinvigorate the debate, it is necessary to take a step back and reassess the allegedly unprecedented developments in military affairs that such technologically advanced weapons systems might introduce. While the issues that have been debated to date are by no means unimportant in understanding the nature and scope of the challenge, they articulate and address only a portion of the issues that are introduced. This creates wobbly foundations for the regulatory regimes that are built in response, potentially creating new regulatory gaps or even jeopardising the regulatory endeavour more generally.

A useful framework upon which to structure such a reassessment is the researcher's approach to ignorance.[2] Noam Chomsky famously distinguished ignorance into the categories of *problems* and *mysteries*:

1 The de-emphasis of autonomy will become clear in the argument set out below. For terminological consistency with the existing literature, however, I will continue to use the 'autonomous weapons system' or AWS label, to denote technologically advanced weapons systems even if the purpose of this article is to foreground other key characteristics.

2 Stuart Firestein, *Ignorance: How It Drives Science* (OUP 2012).

> Our ignorance can be divided into problems and mysteries. When we face a problem, we may not know its solution, but we have insight, increasing knowledge, and an inkling of what we are looking for. When we face a mystery, however, we can only stare in wonder and bewilderment, not knowing what an explanation would even look like.[3]

While the world of linguistics research appears remote from the legal, regulatory and policy issues raised by the prospect of technologically advanced 'autonomous' weapons systems, Chomsky's distinction can be deployed to expand those debates beyond their current confines.

Adapting the problem-mystery distinction to the context of this paper, I would like to suggest that the contemporary debates around AWS involve exclusively legal *problems*, and that the realm of what might be termed legal *mysteries* has been entirely ignored or overlooked. What I mean is that AWS are perceived to raise questions that are familiar to jurists and lawyers, and to ethicists and policy-makers. These are questions that may be articulated in common terminology, discussed with reference to shared frameworks, and examined through established disciplinary methods and methodologies, and which may be subject to rational professional disagreement. Foremost among these are questions surrounding whether AWS would be capable of complying with the obligations imposed by international humanitarian law (IHL) or if AWS can fulfil the requirements of human rights law. Other familiar questions revolve around the apportionment of responsibility for AWS among proximate human beings at a general level,[4] or how to assign liability for the negative or unlawful consequences of AWS deployment specifically.[5] These are examples of legal *problems* because, even though there may not be agreement as to what the solutions would be at the moment, there are agreed methodologies and acceptable arguments that can be pitted against each other in a rational manner to generate responses that may become plausible solutions.

3 Noam Chomsky, 'Problems and Mysteries in the Study of Human Language' in Asa Kasher (ed), *Language in Focus: Foundations, Methods and Systems: Essays in Memory of Yehoshua Bar-Hillel* (Springer Netherlands 1976). Quote drawn from Steven Pinker, *How the Mind Works* (Penguin 2015) xvii.

4 See, for example, attempts to apportion responsibility to different proximate humans in Marcus Schulzke, 'Autonomous Weapons and Distributed Responsibility' (2013) 26 Philosophy & Technology 203.

5 See generally, Hin-Yan Liu, 'Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems' in Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics Policy* (Cambridge University Press 2016).

A different way of approaching my claim that contemporary AWS debates address legal problems is by looking at the proposed responses that have been tabled. One notable response aims at a pre-emptive prohibition of AWS, which consolidated around 2012/2013 with the publication of the Human Rights Watch's influential report,[6] and the formation of the Campaign Against Killer Robots.[7] In the middle ground is Christof Heyn's position of advocating for a moratorium on the deployment of AWS.[8] At the other end of the spectrum, a laissez-faire camp identifies no significant legal problems because extant law is capable of regulating the alleged issues raised by AWS or can be adapted to accommodate the challenges.[9] While these positions appear to run the gamut from one pole to its diametric opposite, these responses illustrate the *problem* nature of the debate. There may not be a ready solution at hand, but these debates and proposals suggest that 'we have insight, increasing knowledge, and an inkling of what we are looking for'. In other words, much of the AWS debate concerns legal problem-solving insofar as familiar recommendations are made that situate the challenge in terms of the broader legal context even in the absence of agreement.

What I propose is to galvanise explorations into the potential legal *mysteries* that AWS may both create and reveal: that we should spend the time to 'stare in wonder and bewilderment' in order to contemplate unconventional perspectives and to entertain unorthodox responses. The imperative is to prepare the broader regulatory sphere to ensure its appropriateness and adequacy with regard to the prospect of technologically advanced weapons systems, currently conceived of as *autonomous* weapons systems. Doing this through an exclusively problem-oriented approach would require the identification of the full set of relevant problems to ensure the sufficiency of that regulatory sphere moving forward, which arguably highlights the vagaries and incompleteness

6 Human Rights Watch, 'Losing Humanity: The Case against Killer Robots' (Human Rights Watch and International Human Rights Clinic 2012).

7 <www.stopkillerrobots.org> accessed 2 April 2019.

8 Christof Heyns, 'Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Christof Heyns' (UNGA 2013) A/HRC/23/47. My own position grounds a moratorium upon ambiguous and inadequate concepts of responsibility, but the recommendation is the same: Liu, 'Refining Responsibility' (n 5).

9 Kenneth Anderson, Daniel Reisner and Matthew Waxman, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems' (2014) 90 International Law Studies 386; Kenneth Anderson and Matthew Waxman, 'Law and Ethics for Robot Soldiers' (*Policy Review*, 1 December 2012) <www.hoover.org/research/law-and-ethics-robot-soldiers>; Michael Schmitt and Jeffrey Thurnher, '"Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict' (2013) 4 Harvard National Security Journal 231.

of the contemporary debates insofar as such comprehensiveness will remain elusive.

Put in different terms, the nature of regulatory challenge lies not with the introduction of new technologies, qua technology, but rather with technological *change* perturbing the regulatory system.[10] In Lyria Bennett Moses' words:

> [T]he primary issue for regulators is not the need to 'regulate technology' but the need to ensure that laws and regulatory regimes are well adapted to the sociotechnical landscape in which they operate, which changes over time… Regulators need to respond to new technologies, not because they are technological per se, but because they are *new* and law and regulation need to be changed to align with the new sociotechnical landscape, including new negative features.[11]

If this orientation is correct, then a fuller assessment of what is *truly novel* about AWS is necessary to better prepare the sociotechnical landscape. In this sense, the prospect for weapons systems *autonomy* is clearly something genuinely new, even if human soldiers can be considered as close analogues or historical precedents.[12] There is obvious value in the discussions centred upon weapons systems autonomy in this framework that should not be discounted.

The remaining question is whether autonomy is the *only* new feature introduced by AWS into the sociotechnical landscape. If it is, then the problem-oriented debates are both necessary and sufficient preparation. But if there are other key factors that are arguably new to this sociotechnical landscape, then it is possible that the problem-oriented debates remain insufficient, even if they are necessary. Furthermore, the failure to consider the impact of these other new characteristics represents vulnerabilities in the sociotechnical landscape that may become the locus for regulatory instability, which in itself merits attention.

This brings us back to the problem-mystery distinction: legal methodologies and approaches yield legal problems that are amenable to legal solutions. These repetitive tautologies underscore the limitations of the disciplinary lens,

10 Lyria Bennett Moses, 'Regulating in the Face of Sociotechnical Change' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford University Press 2016).

11 Ibid 577.

12 Geoffrey Corn, 'Autonomous Weapons Systems: Managing the Inevitability of "Taking the Man Out of the Loop"' in Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics Policy* (Cambridge University Press 2016).

which identify legal problems in a capable and reliable manner,[13] but largely fail to unveil legal mysteries.

Put in different terms, it is possible that the *autonomy* of AWS looms large over the ethical, legal and policy debates,[14] because this framework presents relatively clear problems in these domains that are centred upon issues raised by control, influence, responsibility and liability. Framed as 'autonomous' it becomes possible to have ethical, legal and policy discussions over the impact of these technologies in these domains.

The same cannot be said if, instead, we pivot our approach on the same core technologies and the capabilities that these enable to focus upon the features of interconnectedness or of their networked character, or to emphasise their systems dynamics or their complex adaptive nature. This is where the *mystery* part of our ignorance re-enters: it is far from obvious what the ethical, legal or policy issues are through these approaches, and they are difficult enough to broach, let alone discuss intelligently. Put differently, my claim here that we need to have a regulatory debate centred upon AWS not as *autonomous* weapons systems, but rather as interconnected or networked weapons systems, and as complex adaptive weapons *systems* may leave the reader in 'wonder and bewilderment', or less charitably, leave the reader cold and befuddled. There are no ready problems to be posed and debated, and it is far from obvious what suitable responses might look like. Indeed, it is difficult to get a foot in the door, let alone know what the utility would be of pursuing these different approaches or where this might lead.

The first reason to explore the legal mysteries of networked weapons systems, and weapons *systems*, is to round out considerations as to what is new for the sociotechnical landscape as discussed above. Even if we are not in a position to frame intelligible problems out of these approaches at the moment, engaging with the novelty that these approaches raise is necessary to complete the regulatory endeavour. This process of engagement might itself break off pieces of the mystery and upgrade them into clusters of problems that can then be addressed.

The second reason is to expose hitherto unchallenged presumptions lurking in legal doctrine, and to reveal latent inconsistencies and incoherencies that pursuing these different approaches enables. These are unpacked in greater detail below, but it is worth acknowledging legal fictions are heuristics and that

13 The old adage, 'If all you have is a hammer, then everything starts to look like a nail' may provide a loose analogy.

14 Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press 2016).

heuristics function well under a set of stable circumstances but may not fail gracefully when exposed to another environment. Processes of sociotechnical change destabilise the environment in which legal fictions operate, and in doing so give us insight into the their presumptions, biases and limits.[15] This may in turn reveal unexamined settings in the regulatory configuration that may then be more intentionally manipulated.

The third, and perhaps most important, reason is that many alleged objections to AWS actually fail upon closer inspection and the approaches championed provide converging arguments to support those objections. The following sections set out the limitations of the autonomy framework and why, in particular, IHL objections and control and responsibility issues fall short of the mark (or, for example, are at least insufficiently strong to ground a pre-emptive prohibition at present).

This is incidentally where a piece of the legal mystery might be upgraded into a cluster of legal problems because both the networks and the systems approaches provide reinforcing arguments for a moratorium or even a pre-emptive prohibition of AWS.[16] From the networks approach, the fundamentally dispersed and seemingly uncoordinated nature of the weapons system suggests an absence of control and concomitant responsibility that is legally necessary for the use of organised force. Furthermore, the systems approach suggests that organised violence may be emergent outcomes of complex adaptive systems which not only echo the absence of control and responsibility for that violence, but creates new challenges where such violence maps onto breaches of IHL or violations of human rights law. The systems approach also imports the possibility for 'normal accidents',[17] a small subset of predictable, yet inevitable, accidents to be discussed in the context of AWS. If advocates for a pre-emptive prohibition are seeking arguments as to why AWS are inherently incapable of adhering to the requirements imposed by IHL, for example, the 'normal accidents' case may be just what they are looking for.

This article first discusses the autonomy framework and its limits, with reference to the sense-think-act paradigm, before moving to consider why

15 Analogous to the insights garnered into the visual system by studying the effects of optical illusions, Pinker (n 3).

16 The possibility to propose a response suggests that there is 'an inkling of what we are looking for', suggesting that we are no longer confronted with a mystery, and this shows progress in gaining higher quality ignorance.

17 Charles Perrow, *Normal Accidents: Living with High Risk Technologies* (Princeton University Press 2011); Matthijs Michiel Maas, 'Regulating for "Normal AI Accidents": Operational Lessons for the Responsible Governance of AI Deployment', *Association for the Advancement of Artificial Intelligence* (2018).

objections to AWS on the grounds of IHL fail because of an unacknowledged alignment between technological trajectories and legal progress. The article then considers the limits of control as framing devices for AWS and the persistence of responsibility gaps, which in turn suggests that the prospects for meaningful human control are likely to remain elusive. After drawing the limitations of the autonomy framework together, the paper launches into explorations of the networks and the systems approaches to AWS. It seeks to demonstrate both how these approaches are resilient to the problem-oriented legal debates, but also that core presumptions and principles of legal doctrine might break down as a result of pursuing these approaches. This article concludes by recognising that it is caught by the mysterious nature of the networks and the systems approaches to AWS: in not being able to articulate a cogent set of questions beyond reframing the problems posed by control and responsibility, we remain for the time being in a position of 'not knowing what an explanation would even look like'.

2 Autonomy and Its Limits

The common definition of an autonomous weapons system focusses exclusively on its *autonomous* character. According to UN Special Rapporteur Christof Heyns, lethal autonomous robotics 'refers to robotic weapon systems that, once activated, can select and engage targets without further intervention by a human operator', further clarifying that '[t]he important element is that the robot has an autonomous "choice" regarding selection of a target and the use of lethal force'.[18] This definition establishes the primacy of autonomy as the defining characteristic of AWS and situates subsequent discussions within the autonomy framework. While subsidiary questions concerning the type of autonomy might be raised,[19] these are minor details that confirm that dominant framework. Yet, the monopoly of autonomy is neither helpful nor justified.

Underlying all of these discussions is a simple technology that can be expressed as the 'sense-think-act' paradigm. In such an arrangement, apparently autonomous action is the outcome of inputs impinging upon a sensor, which

18 Heyns (n 8) para 38. Both the US Department of Defense and Human Rights Watch converge on this definition, as Heyns notes. While Heyns refers to 'lethal autonomous robotics', this paper will use 'autonomous weapons systems' for consistency with the literature.

19 Noel Sharkey, 'Staying in the Loop: Human Supervisory Control of Weapons' in Nehal Bhuta and others (eds), *Autonomous Weapons Systems: Law, Ethics Policy* (Cambridge University Press 2016). See also, Liu, 'Refining Responsibility' (n 5) 327–330.

sends a signal to a processor, which in turn sends a signal to an actuator which then has the capacity to produce an output which often involves manipulating the physical world.

The *appearance* of autonomy is critical here, because this is just one way of modelling this paradigm.[20] In other words, the autonomy model throws a box around the sensor-processor-actuator 'unit' and considers it an autonomous entity. This is convenient, especially in legal terms, as it enables the direct substitution of natural persons and artificial systems. The autonomy framework minimises legal disruption because this substitution facilitates legal accommodation: essentially it allows us to step into the robot's shoes, and enables the robot to step into our shoes, and allow the law to continue functioning. By presuming notions of singular embodied entities, however, the autonomy approach may actually be misleading.

The claim here is not that the autonomy framework in general is incorrect or necessarily misleading, but rather that it privileges certain problems over others that might hinder full legal responses. In particular, the autonomy framework foregrounds the liminal status of autonomous entities as being between agent and object.[21] When coupled with unpredictability of behaviour and unforeseeability of outcome, questions of control and concomitant responsibility flow freely from the autonomy framework as the dominant legal problems. Before tackling these full on, it is useful to consider the shortcomings of relying upon international humanitarian law (IHL) to regulate the use of autonomous weapons systems. This raises problems associated with capacity for legal compliance, but also illustrates the possibility that solving legal problems might not provide the fuller regulatory responses that might be necessary to re-establish equilibria in the wake of sociotechnical change.

3 Legal Shortcomings: International Humanitarian Law

Given the purpose of autonomous weapons systems in the projection of armed force, recourse to IHL is not only intuitive but takes precedence over

20 In other words, this framing permits *problematisation* of the issues while resisting *mystification*. While generally perceived to be productive, this process may introduce distortions to the research agenda where the problems posed do not model the challenge in its entirety.

21 Hin-Yan Liu, 'Categorization and Legality of Autonomous and Remote Weapons Systems' (2012) 94 International Review of the Red Cross 627.

other bodies of law as *lex specialis*.[22] The IHL questions raised by AWS revolve around their capacity to comply with the cardinal principles of distinction and proportionality. It is not necessary for our purposes to delve into the details of these principles,[23] but merely to observe that the principles prevent 'sloppy' violence that is indiscriminate and excessive while incentivising 'slick' violence that is precisely targeted and carefully calibrated. In sum, IHL has nothing more to say about measured violence that is directed towards a legitimate military target.

The reason why IHL fails to constrain the challenges posed by AWS is twofold. First, the IHL principles of distinction and proportionality are, in fact, technical performance criteria contingent upon field performance, and not true legal challenges. This is borne out in the debate between Human Rights Watch[24] and AWS advocates,[25] where the projected legality of AWS hinges upon the capacity of the systems to comply with IHL obligations in a functional and practical manner. In this sense, the AWS discussion adds nothing new in terms of IHL: either AWS are technically capable of complying with the principles of distinction and proportionality and can be deployed legally into the battlespace, or they do not, and therefore cannot be lawfully deployed.

The second reason is that the justifications undergirding AWS deployment flow in the same direction as IHL. AWS are lauded by supporters to be restrictive in the use of force: targeted and calculated in engaging the enemy. In this sense, AWS may actually become the paragon of military force, and far from merely complying with IHL, may even become mandated in the battlespace as their advocates suggest.[26]

22 But see, David Luban, 'Human Rights Thinking and the Laws of War' in Jens David Ohlin (ed), *Theoretical Boundaries of Armed Conflict and Human Rights* (Cambridge University Press 2016); and Marko Milanović, 'The Lost Origins of *Lex Specialis*' in Jens David Ohlin (ed), *Theoretical Boundaries of Armed Conflict and Human Rights* (Cambridge University Press 2016).

23 The opposing claims made with regard to IHL compliance can be found in Human Rights Watch (n 6), and prominent responses are articulated in Anderson, Reisner and Waxman (n 9), and Michael Schmitt, 'Autonomous Weapons Systems and International Humanitarian Law: A Reply to the Critics' [2013] Harvard National Security Journal Features.

24 Human Rights Watch (n 6).

25 Schmitt (n 23); Anderson, Reisner and Waxman (n 9).

26 Whether the advent of AWS in the battlespace will correspondingly raise the thresholds required for IHL compliance is another discussion. Yet, suggestions along these lines will have the effect of curtailing human participation in armed conflict in the future insofar as those thresholds are raised beyond inherent human capabilities. Such moves could

Both these reasons suggest that reliance upon IHL to determine the legality of AWS misses an important point: that their purposes go in aligned directions. The purpose of IHL was to direct organised armed violence towards military targets and to contain that violence towards pursuing legitimate military objectives. AWS, however, are envisaged to deliver exactly those capabilities, which largely explains why these are unproblematic in terms of IHL.

More radically, IHL compliance might be seen as part of the problem.[27] Insofar as the legality of AWS is collapsed into the question of IHL compliance, AWS will be declared as lawful prematurely and based upon overly narrow grounds. Instead, IHL compliance by AWS should be treated as a sign of caution, precisely because high-performing AWS will be rendered invisible and unobjectionable in IHL terms. This is a different way of saying that AWS raise different types of problems than those that are articulated under IHL provisions. This provides a strong example of legal problem-solving being necessary but insufficient in framing regulatory responses to sociotechnical changes in the conduct of hostilities.

4 Limits of Control and Responsibility

The autonomy framework foregrounds the liminal status of AWS between agent and object, leading towards the core questions of control and responsibility that are overlooked by IHL principles of distinction and proportionality.

In a previous work,[28] I have argued that there are two different types of responsibility at play in attempts to allocate responsibility for AWS outcomes to proximate human beings, and that this problem is one of legal doctrine that cannot be fixed by improving technical performance. The conceptual issues of responsibility arise from conflating two separate notions of responsibility – causal responsibility and role responsibility – within the same terminology. Causal responsibility actively connects behaviour with consequences, but may do so in a factual sense devoid of blame or praise.[29] Role responsibility involves the discharge of obligations attendant to an office or position, and

make IHL compliance synonymous with AWS use, or conversely with human exclusion, in future armed conflicts.

27 David Kennedy, *The Dark Sides of Virtue: Reassessing International Humanitarianism* (Princeton University Press 2005).

28 Liu, 'Refining Responsibility' (n 5).

29 HLA Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford University Press 2008) 213–215.

is omission-based.[30] Attempts to ascribe responsibility for AWS outcomes to proximate humans do so by collapsing this important conceptual difference.[31] Thus, the conceptual responsibility gap inheres in the space between the different types of responsibility that is invoked, and represents a different type of responsibility gap than the circumstantial responsibility gap that arises from the unpredictability and unforeseeability that is introduced by autonomous entities. The conceptual gap between role and causal forms of responsibility illustrates why it is not possible to substitute causal responsibility attached to the AWS with a patchwork of role responsibilities attached to proximate, and in the temporal sense preceding, human beings. Indeed, attempts to do so will always scapegoat those human beings by over-ascribing responsibility to them because they would be held causally responsibility for something where they only appropriately have role responsibility obligations to discharge.

In other words, there are really two different legal problems lurking within the responsibility gap: one that is created by AWS as liminal entities falling between the categories of agent and object; and another that is revealed by AWS in rupturing flows of causality that in turn split the very concept of responsibility. While the former is a prototypical legal problem, in that it can be articulated and addressed through standard legal methodologies, the latter is an example of a legal mystery that has been upgraded into a legal problem. It was not obvious that there were fundamentally different conceptions of responsibility that were effectively collapsed together in the ensuing discussions of responsibility in relation to AWS. Yet, the refusal to treat the issues unidimensionally through apportion and distribution among proximate human beings,[32] yielded further and more refined legal problems that require further research. Furthermore, excavating the concept of responsibility has practical implications relating to proposed governing criteria for containing AWS: that of 'meaningful human control'.

5 Prospects for Meaningful Human Control?

The criteria of 'meaningful human control' was first advanced by the NGO Article 36,[33] and has gathered significant momentum as the controlling

30 Ibid 212–214.

31 Notably Schulzke (n 4).

32 Ibid.

33 Article 36, 'Killer Robots: UK Government Policy on Fully Autonomous Weapons' (Article 36 2013) Policy Paper 1 <www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf>.

criterion for the legal deployment of AWS. The difficulty, however, is that while there is widespread international consensus as to its importance, there is next to nothing regarding the actual content of the concept as applied to AWS.[34]

If control and responsibility are two sides of the same coin, then the litmus test for 'meaningful human control' would be the unambiguous presence of 'meaningful human responsibility'. This suggests that there are at least two forms of control that are relevant to the human-AWS relationship that mirror the concepts of causal and role responsibilities. Taking this approach would suggest that meaningful human control would be limited to due diligence dimensions (role responsibilities) involved in fulfilling prescribed obligations attaching to an office or function, and would not be appropriately applied to active causal outcomes. In other words, meaningful human control, like meaningful human responsibility, would remain circumscribed to discharging role responsibilities.

Drawing the conceptual distinction between role and causal responsibilities thus leads to insights as to the useful limits of what meaningful human control might achieve. While it may be true that meaningful human responsibility may be secured, it will more likely reflect the doctrine of command responsibility[35] and the limits inherent within it. Thus, meaningful human control over the behaviour of, and outcomes precipitated by, AWS will remain elusive while it may simultaneously be possible to state that meaningful human control (delimited to due diligence obligations) has been secured.

6 Shortcomings of the Autonomy Framework

Thus far, I have sought to demonstrate that the legal debate on AWS has hit the hard limits of legal doctrine – that the legal problems generated by AWS have been solved to the point of reasonable disagreement – and that further discussion along this vein will yield diminishing returns. This is because the autonomy framework privileges an approach involving direct human substitution,

34 But see Rebecca Crootof, 'A Meaningful Floor for "Meaningful Human Control"' (2016) 30 Temple International & Comparative Law Journal 53.

35 Beatrice Bonafé, 'Finding a Proper Role for Command Responsibility' (2007) 5 Journal of International Criminal Justice 599; Chantal Meloni, 'Command Responsibility Mode of Liability for the Crimes of Subordinates or Separate Offence of the Superior?' (2007) 5 Journal of International Criminal Justice 619; Volker Nerlich, 'Superior Responsibility under Article 28 ICC Statute: For What Exactly Is the Superior Held Responsible?' (2007) 5 Journal of International Criminal Justice 665.

and foregrounds questions of control and responsibility as the primary legal problems posed by AWS. But through the autonomy framework, control fades into mere influence, and meaningful human responsibility is eroded not only by unpredictability and unforeseeability but also by the distance between different conceptions of responsibility. These intractable problems are features, not bugs, of the autonomy framework. These are the direct consequences of the autonomy approach and arise because of fundamental incompatibilities between the technology and the impact of AWS on the one hand, and of legal doctrine and ordinary legal processes on the other hand. While the autonomy framework facilitates legal accommodation, and represents the easiest type of legal challenge, there are clear limits to the insights it yields and the adequacy of responses it can offer.

7 The Networks Perspective

An alternative framework to the same underlying technology is that provided by the network, which emphasises the interconnections and communications between the constituent nodes. This perspective can be very roughly divided into the internal and the external network, which very roughly corresponds to introspective and extrospective orientations.[36]

The internal network represents an expansion of the 'sense-think-act' paradigm discussed above. While this was simplistically linear, adding a couple of sensors and actuators into the mix rapidly complicates matters.[37] The first point to note is that removing linearity increases the possible permutations between input and output. This suggests that part of the opacity of the network may arise from the mere complicated nature of these permutations,[38] and that this opacity is in addition to any opacity deriving from the 'black-box' of machine-learning processes residing in the processors. In other words, there are at least two different mechanisms underlying unpredictability and unforeseeability in this perspective, and an appropriate regulatory regime must be capable of accounting for at least both of these effects.

An important contribution of the networks perspective is that it is also possible to zoom out and focus on the network within which these sensor-processor-actuator units are subsumed. It is easiest to think of this as a swarm: relatively simple units that, through their communication and coordination

36 Furthermore, this admittedly a false dichotomy for illustrative purposes only.

37 Note also that there might be a connection between the two processors illustrated here.

38 Note that this is complicated, and not complex.

with each other, are capable of exhibiting behaviour that appears more sophisticated and complex than their components or behavioural rules would suggest.

What is curious here is that the autonomy framework is a viable model: it is possible to throw a box around the network and treat it as an individual entity. In other words, such a swarm can still fall within the common definition of an AWS because it can still be considered as 'robotic weapon systems that, once activated, can select and engage targets without further intervention by a human operator'.[39]

Arguably, this is what the AWS debate to date has done. Yet, from the networks perspective it becomes clear that the common AWS definition fails to capture the characteristics of interconnectedness and communication that can lead to much more dynamic capabilities than is implicit with the 'fire-and-forget' orientation of the existing definition. Thus, while it may be possible to model such networks as autonomous entities, it is necessary to be mindful of the limitations inherent within that model.

The networks approach, however, may be anathema to the law. This claim can be substantiated by attempts to *problematise* the networks approach in legal terms. In other words, what are the legal problems that arise from the functioning of a network, what legal methodologies can be deployed to address those questions and what might legal answers to network functioning look like?

Insofar as the law privileges agents and actions, the networks approach undercuts both of these fundamental presumptions, thereby marginalising legal principles and processes from the discussion. Put differently, how can the law articulate the legal challenges raised by networks and nodes? Can a network be the cause of an impugned effect? Can a network be held accountable, and perhaps more problematically how might reparations, retribution and repeat violations be addressed? At this stage at least, it appears as though the networks approach emphasises features that are ordinarily invisible to the law, rendering the networks approach *mysterious* to the law.

This creates significant issues in terms of regulatory responses towards AWS. Unless it is possible to isolate and individuate AWS, which seems remote given that networks flow both introspectively and extrospectively, modelling AWS as a network appears to be both accurate and appropriate.

Foregrounding the network dimension of these technologically advanced weapons systems thus represent a core *change* to the sociotechnical landscape that is independent of the change introduced by the prospect for these

39 Heyns (n 8) para 38.

weapons systems to function autonomously. In other words, there are at least two different and independent changes to the sociotechnical landscape that need to be considered in order to generate robust regulation against the introduction of AWS.

8 Emphasising the 'System' in Autonomous Weapons Systems

The systems perspective is an alternative, yet compatible, approach to the networks approach discussed above. Rather than focus on the nodes and interconnections within a network, the system perspective emphasises the unity of a collective in terms of behaviour and outcome. As Donella Meadows puts it: 'A system is a set of things … interconnected in such a way that they produce their own pattern of behaviour over time … *The system, to a large extent, causes its own behaviour!*'[40] Curiously then, the systems approach converges strongly with the autonomy framework that has been largely dismissed in the preceding discussion, insofar as it throws a box around a complex set of factors and models it as an independent entity. Put differently, a system that causes its own behaviour can be treated as an autonomous system in some situations.

Yet, taking the systems approach to AWS seriously provides three distinct insights. First, it raises significant questions concerning the appropriate 'level' of the system that should be treated as an autonomous entity. Secondly, the emphasis upon the system blurs doctrinal boundaries that separate the military from the civilian, and which distinguish cyberspace from the physical world. Thirdly, it opens up the possibility to talk about system behaviours, emergent effects and normal accidents in the context of AWS deployment.

Implicit within the AWS debates to date is the treatment of an autonomous weapons system as a substitute for the human combatant.[41] While an individual human being might be the logical analogy for an AWS through the autonomy framework, there is no reason to treat an AWS only at this 'level'. It may be overstretching the fractal metaphor provided by complex adaptive systems,[42]

40 Donella H Meadows, *Thinking in Systems: A Primer* (Chelsea Green Publishing 2008) 2 (emphasis added).

41 'No amount of training or supervision can eliminate a very basic reality of human operatives: they are, and have always been, "autonomous" weapons systems, because all soldiers must exercise cognitive reasoning in the execution of their battlefield tasks', Corn (n 12) 211–212.

42 James Gleick, *Chaos: Making a New Science* (Vintage 1997); Melanie Mitchell, *Complexity: A Guided Tour* (OUP USA 2011).

but the general idea is that defining features and patterns are replicated across scales. Thus, not only is the individual combatant an AWS, but so is the platoon, company, battalion, and brigade that he is subsumed within: each level of organisation constituting its own autonomous capacity to project military force. A curious conclusion to this line of thinking is that the nation State itself might be considered as an autonomous weapons system, an insight that converges with Max Weber's definition of the State as 'a human community that (successfully) claims the *monopoly of the legitimate use of physical force* within a given territory'.[43] And the fractal patterns may extend beyond the State to international organisations such as NATO, which might also be considered as an autonomous weapons system.

Indeed, while the very notion of an AWS seems to be integrally connected to tactical systems deployed directly in the battlespace when viewed through the autonomy framework, it seems counterintuitive to limit the application of technologically advanced systems to merely substituting the troops on the ground. If true creativity is the hallmark of an artificial intelligence application,[44] and since these have proven capable of defeating human champions at the most strategic and complex of games, it seems logical that AWS would be deployed at least at Brass level. In other words, it makes little sense to squander these brilliant technologies at the tactical level when they can be more effectively deployed to oversee military organisation at the strategic level. This raises the prospect for command responsibility to inhere within the AWS, or at a minimum influence the role and capacity of commanders, thereby blurring further the conceptual boundaries, established practices and perceived flows of responsibility, liability and accountability.

Thus, the systems approach illustrates the shortcomings of the autonomy framework. First, if there are multiple possible boxes that can frame autonomous entities then which is the appropriate legal framing of an AWS? An unambiguous and exclusive level of autonomy seems to underlie notions of legal treatment, and this nested conception of AWS undercuts this fundamental requirement. Secondly, the systems perspective highlights the fact that an AWS is not only a technologically advanced system, but necessarily involves human elements and organisational dynamics. Isolating the technological components of an AWS, as the commonly agreed definition does, artificially

43 Max Weber, 'Politics as a Vocation' in HH Gerth and C Wright Mills (eds), *From Max Weber: Essays in Sociology* (Routledge and Kegan Paul 1946) 78 (emphasis original).

44 Yuval Noah Harari, 'Why Technology Favors Tyranny' *The Atlantic* (October 2018) <www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330> accessed 23 October 2018.

simplifies a complex system and undermines the utility of models that take this approach.[45]

The systems perspective also blurs the fundamental distinction between the mutually exclusive categories of privileged belligerent and civilian.[46] While the trend towards this blurring has been evident with the rise of the private military company (PMC),[47] the systems perspective removes the boundaries altogether for the simple reason that all components of a system function together to produce the behaviour of that system.

The contribution of the systems perspective in this regard is that the discussion is no longer about civilians performing military functions, but rather that the distinction between civilian and combatant is no longer relevant where each is participating in a larger system involved in projecting martial force. This is more than merely reaching 'backwards' to ascribe responsibility to civilian programmers, manufacturers and users of AWS:[48] the claim here is that such distinctions may no longer do useful regulatory work. A different way of expressing this is that the tip-of-the-spear typology is no longer an adequate model, since this presumes graduated distances from where military force is ultimately deployed. Instead, the systems perspective might suggest that the battlespace is everywhere simultaneously. This is an implication arising from the observation that the system as a whole is involved in projecting, rather than the extant notion that only certain sub-components of a system are responsible for utilising force while other sub-components of the same system are somehow relegated to supporting roles. This obviously poses more general problems for IHL, founded as it is upon the fundamental distinction between privileged combatants and civilians.

Finally, the systems perspective opens discussions of the legal treatment of system behaviours. If the system is responsible for its own behaviour in an important and meaningful sense, then the law will have to grapple with outcomes which were not only unpredictable and unforeseen, but also to an important degree unintended, undirected and uncontrolled. This implication of this line

45 Reflecting the shift away from treating 'technology' as a regulatory target, Gregory N Mandel, 'Legal Evolution in Response to Technological Change' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (Oxford University Press 2016); Bennett Moses (n 10).

46 The eroding distinction between cyberspace and the physical world runs along similar lines and will not be discussed further in the interests of space.

47 PW Singer, *Corporate Warriors: The Rise of the Privatized Military Industry* (Cornell University Press 2004). For the legal problems that PMCs raise, see Hin-Yan Liu, *Law's Impunity: Responsibility and the Modern Private Military Company* (Hart 2015).

48 Schulzke (n 4).

of thought is that death and destruction may be nothing more than emergent effects of autonomous weapons system behaviour. This poses problems for processes that that ascribe meaning to military force and hold the projection of force to account. Victims of AWS might therefore have their harm categorised as mere damage, with the legal system failing to recognise the injury at stake.[49]

A strong example would be where AWS are causally responsible for outcomes that look like human rights violations, and would be categorised as such if those arose from the activities of conventional armed forces. What does it mean to notions of human rights if these are violated by system behaviour? In other words, what role might the law play if human rights violations are merely the emergent outcomes of complex adaptive system behaviour? Even if, doctrinally, a solution can accommodate for human rights being violated by system behaviour, this would alter the very meaning of what a human right is, and what it means to breach its protections.

In this vein, it should be noted that artificial intelligence applications are highly susceptible to 'normal accidents'.[50] 'Normal' or 'system' accidents are 'accidents' that can be expected to occur as a result of utilising high risk technologies, and importantly these can neither be designed out nor prevented.[51] While case studies in this area have involved high visibility catastrophes such as the *Challenger* disaster and Three Mile Island, it is not necessary to delimit normal accidents to physical disasters. Thus, the systems perspective builds a conduit between AWS debates and normal accident theory to suggest that human rights violations may be endemic to AWS deployment, where such systems are susceptible to normal accidents occurring and assuming that those accidents invoke outcomes that map onto human rights laws.[52]

---

49 On the difference between harm and injury, see Scott Veitch, *Law and Irresponsibility: On the Legitimation of Human Suffering* (Routledge-Cavendish 2007) 85–92.

50 Maas (n 17).

51 Perrow (n 17).

52 Hin-Yan Liu, 'The Digital Disruption of Human Rights Foundations' in Mart Susi (ed), *Human Rights, Digital Society and the Law: A Research Companion* (Routledge 2018). In other work, I have highlighted the insufficiency of human rights law to articulate the harms posed by artificial intelligence applications, so there may be severe shortcomings in deploying human rights law as the litmus test for AWS and normal accidents, Hin-Yan Liu and Karolina Zawieska, 'From Responsible Robotics Towards a Human Rights Regime Oriented to the Challenges of Robotics and Artificial Intelligence' [2017] Ethics and Information Technology; see also Hin-Yan Liu and Karolina Zawieska, 'A New Human Rights Regime to Address Robotics and Artificial Intelligence' (*JusLetter IT*, 23 November 2017).

Invoking normal accidents in the AWS debate suggests that a different calculus might need to be considered if human rights violations are to be ordinary, expected, and unavoidable outcomes of their deployment. This observation should weigh heavily against claims that AWS can be lawfully deployed provided that they meet existing legal obligations set forth by IHL, and arguably should set up a rebuttable presumption (moratorium) that prevents AWS deployment on legal grounds.

Thus, the systems approach anchors a third independent set of novelties to the sociotechnical landscape that needs to be addressed, accommodated, or accounted for in regulatory responses. Again, because this approach captures a plausible set of new changes to the regulatory environment, it requires resolution on its own terms and it is no use to say that the *technological* challenges posed by AWS have been answered from an autonomy or networks perspective. In this sense, addressing the change posed by the very approach itself, as an approach, is necessary irrespective of what it is an approach towards provided that it represents a significant change to the sociotechnical landscape. These may constitute greater, more generalisable, considerations for considering the regulation of (technological) change, and their exposition of these ideas in the context of technologically advanced weapons systems provides a special case study in this regard.

9 Concluding Thoughts

While this article has sought to pivot from the autonomy framework to the networks and systems approaches to the same underlying technology, it has arguably failed to complete the move and take full advantage of the insights that these provide. A large measure of this is due to the continued pull of the concepts of control and responsibility that have become integral to legal and ethical responses to AWS. Indeed, it is difficult to fathom what legal problems arise from weaponised networks or militarised systems, let alone craft appropriate legal responses to them. This in itself may indicate the limits of ordinary legal principles and processes: where networks and systems fall outside of these, the legal system is incapable of comprehending both the opportunities and challenges that these approaches provide. In other words, the networks and the systems approaches may constitute legal mysteries that insulate them from law, ethics and policy debates.

In terms of opportunities, it may be that these approaches provide converging grounds for basing a moratorium on AWS deployment. Whereas previous work discussed above suggested that a moratorium could be founded upon

the conceptual responsibility gap – the irreconcilable differences in meaning between different types of responsibility – as a logical consequence of the autonomy framework, both the networks and systems approach could offer similar conclusions that reinforce the call for a moratorium.

The networks approach, being anathema to ordinary legal principles and processes, could ground the idea that AWS are unlawful in the sense of being beyond legal regulation. This is a considerably stronger legal argument than claims that AWS might violate IHL both because this is not a performance metric that is capable of being satisfied, but also because it reinforces the idea that the conduct of hostilities remains amenable to legal regulation. Insofar as weaponised networks are unleashed into the battlespace, and until the point that the law is capable of comprehending and containing their effects, the networks approach suggests that AWS cannot lawfully be used.

A similar conclusion can be derived from the systems approach, as discussed above. Insofar as system behaviour and its emergent outcomes map onto human rights violations, and these cannot be identified, challenged or overturned by the mechanisms provided by human rights law, this would suggest that human rights violations are inevitable outcomes of deploying AWS. Furthermore, the stability of system behaviour and the unpredictability of emergent outcomes deriving from complex adaptive systems imply that such issues involving human rights might both be persistent and catch us by surprise. Finally, there is the small, but potentially ineradicable sphere of normal accidents where some subsets of AWS will inevitably fail. Such failures cut against the grain of legal principles governing the use of force which require the capacities of control and restraint in the conduct of hostilities.

Taken together, these different approaches may provide powerful arguments towards a moratorium on AWS, if not an outright ban. Yet, despite the promise of these approaches, we have barely scratched the surface of the legal ramifications in this article and in the Beyond Killer Robots conference more generally. Again, this was primarily because to ask a legally framed question was to bring with it the baggage of unexamined presumptions, conceptual boxes and procedural thinking that comes with thinking under a lawyerly hat. To fully exploit the insights that are opened up by the networks and systems approaches to AWS would be to break free from questions of control and responsibility, but to do this would require radically different thinking about the law and how it might apply to AWS. Perhaps more important is that these ways of thinking reveal why the law is locked out of the conversation, or where it has nothing to say, as these indicate the conceptual and procedural limits of the law. To ensure adequate and continued legal governance of armed conflict, it is essential to identify these limits, and their underlying rationale, such that the

enquiry about AWS might contribute to a more general reinforcement of the legal control over the conduct of hostilities.

What is clear, however, is that the emphasis upon regulating for *change* in the sociotechnical landscape leads to an imperative that the regulatory endeavour must identify and seek to address the sources of the 'new' that precipitates this change. The features or characteristics that underlie this change may be cogently framed through different organising concepts, each of which must be considered independently. The implication is that, even if the legal problems are solved from the perspective of autonomy, this has no bearing on the changes to the sociotechnical landscape that might arise from alternative approaches, including those of the networks and the systems approaches. Furthermore, in de-emphasising the *technological* nature of changes that impinge upon the sociotechnical landscape, a case may be made that regulatory responses should primarily be oriented towards the very *approaches* through which a technology – or any change in essence – is modelled.[53] In other words, the pivot to the networks and the systems approach for regulatory approaches to AWS may itself be misguided: the change comes not from the technological innovation, how it may be used, or the impacts it may cause, but rather the change comes from the very fact of introducing changes to the sociotechnical landscape through these very approaches themselves.

Acknowledgements

This paper is based on a keynote delivered at the 'Beyond Killer Robots: Networked Artificial Intelligence Disrupting the Battlefield' conference held at the Faculty of Law, University of Copenhagen, on 15–16 November 2018. I would like to gratefully acknowledge the support of the Danmarks Frie Forskningsfond (Independent Research Fund Denmark) Research Stay Abroad, which provided the time and space to develop this article.

53 'All models are wrong, but some models are useful', commonly attributed to George EP Box.